Using sedaDNA from North Sea sediment cores to reconstruct the early Holocene palaeoenvironment

Becky Cribdon

Submitted for the degree of Doctor of Philosophy

University of Warwick
School of Life Sciences
July 2021

Contents

1 Introduction 14
1.1 Current knowledge of the Doggerland environment . . . 15
1.2 Human effects of inundation . . . 18
1.3 Doggerland and Neolithisation . . . 19
1.4 SedaDNA and its application in the North Sea . . . 21
1.5 Scope and aims of thesis . . . 23

2 Main materials and methods 25
2.1 Data collection . . . 25
2.2 Initial data analysis . . . 32
2.3 Authentication . . . 37

3 PIA: more accurate taxonomic assignment of metagenomic data demonstrated on sedaDNA from the North Sea 42
3.1 Introduction . . . 43
3.2 Algorithm . . . 47
3.3 Methods . . . 51
3.4 Results . . . 52
3.5 Discussion . . . 59

4 Taxonomic results 61
4.1 Comparing taxonomic assignment by MEGAN and PIA . . . 61
4.2 Introduction to core summaries . . . 66
4.3 ELF022 . . . 69
4.4 ELF027 . . . 72
4.5 ELF059 . . . 81
4.6 ELF059A . . . 93
4.7 ELF060 . . . 106
4.8 ELF032A . . . 112
4.9 ELF033 . . . 123
4.10 ELF033A . . . 133
4.11 ELF034A . . . 138
4.12 ELF034 . . . 149
4.13 ELF054 . . . 156
4.14 ELF039 . . . 165
4.15 ELF040A . . . 180
4.16 ELF041 . . . 190
4.17 ELF042 . . . 197
4.18 ELF044 . . . 205
4.19 ELF044A . . . 210
4.20 ELF031 . . . 214

4.21 ELF031A 218
4.22 ELF051 233
4.23 ELF045 243
4.24 ELF047 252
4.25 ELF047A 259
4.26 ELF049 265
4.27 ELF046A 271
4.28 ELF050 272
4.29 ELF053 277
4.30 Higher-level analyses 283
4.31 Conclusion 299

5 Authentication results 302
5.1 MetaDamage analysis 302
5.2 Stratification analysis 316

6 Mesophilic taxa and human disturbance indicators 352
6.1 Binomial probability analysis 353
6.2 Comparative results from more confident taxa 354
6.3 Results from mesophilic taxa and human disturbance indicators 358
6.4 Conclusion 370

7 Discussion 372
7.1 Addressing the project aims 372
7.2 Future directions 385

Bibliography 387

Appendices 397

A Full PDF of Cribdon *et al.* (2020) 398

B DNA yields by sample 411

C Read counts at key stages of initial data analysis 413

D Taxa found in the Viridiplantae data 417

E Taxa found in the Metazoa (no Primates) data 439

F MetaDamage plots per sample for Embryophyta 447
F.1 ELF027 448
F.2 ELF059 449
F.3 ELF059A 450
F.4 ELF060 452
F.5 ELF032A 453
F.6 ELF033 454
F.7 ELF033A 457
F.8 ELF034A 458
F.9 ELF034 461
F.10 ELF054 462
F.11 ELF039 465
F.12 ELF040A 467
F.13 ELF041 469

F.14 ELF042 470
F.15 ELF044 471
F.16 ELF044A 471
F.17 ELF031 472
F.18 ELF031A 472
F.19 ELF051 475
F.20 ELF045 476
F.21 ELF047 478
F.22 ELF047A 479
F.23 ELF049 480
F.24 ELF046A 481
F.25 ELF050 481
F.26 ELF053 482

G MetaDamage plots per sample for Metazoa (no Primates) 484

G.1 ELF027 485
G.2 ELF059 485
G.3 ELF059A 487
G.4 ELF060 488
G.5 ELF032A 488
G.6 ELF033 490
G.7 ELF033A 491
G.8 ELF034A 492
G.9 ELF034 493
G.10 ELF054 494
G.11 ELF039 497
G.12 ELF040A 499
G.13 ELF041 500
G.14 ELF042 501
G.15 ELF044 502
G.16 ELF044A 502
G.17 ELF031 502
G.18 ELF031A 503
G.19 ELF051 505
G.20 ELF045 506
G.21 ELF047 507
G.22 ELF047A 509
G.23 ELF049 510
G.24 ELF046A 510
G.25 ELF050 511
G.26 ELF053 512

List of Figures

1.1 Reid (1913) Fig. 4: Estimated coastline for the lowest submerged forest 15
1.2 Walker *et al.* (2020) Fig. 2: North Sea coastline reconstructions 16
1.3 Platt *et al.* (2017): "Evolution of temperature in the Post-Glacial period according to Groeanland ice cores" 17
1.4 Robb (2013) Fig. 1: The spread of the Neolithic in Europe 20

2.1 Bathymetry map of the North Sea showing all cores 26
2.2 Bathymetry map of the North Sea showing the 27 cores in this project . 27
2.3 Bathymetry map of the palaeochannel transect 28

3.1 Database landscape diagram to demonstrate the problem of "oasis" taxa 44
3.2 Example BLAST output structure in format "6 std staxids" 48
3.3 Flowchart illustrating the PIA algorithm 50
3.4 Taxonomic profiles of MEGAN and PIA outputs from ELF039 (Mesangiospermae) 54
3.5 Taxonomic profiles of MEGAN and PIA outputs from ELF039 (Pentapetalae) 55

4.1 Read counts for all samples when assigned by MEGAN or PIA 63
4.2 Ratios between MEGAN and PIA read counts for all samples 63
4.3 Example Krona chart 67
4.4 Key to ecological categories. 67
4.5 ELF022 location map 69
4.6 ELF022 read counts 70
4.7 ELF022 biogenomic mass profiles 71
4.8 ELF027 location map 72
4.9 ELF027 read counts 73
4.10 ELF027 biogenomic mass profiles 74
4.11 ELF027 Pianka similarity scores 79
4.12 ELF059 location map 81
4.13 ELF059 read counts 82
4.14 ELF059 biogenomic mass profiles (A) 83
4.15 ELF059 biogenomic mass profiles (B) 84
4.16 ELF059 biogenomic mass profiles (C) 85
4.17 ELF059 Pianka similarity scores 91
4.18 ELF059A location map 93
4.19 ELF059A read counts 94
4.20 ELF059A biogenomic mass profiles (A) 95
4.21 ELF059A biogenomic mass profiles (B) 96
4.22 ELF059A Pianka similarity scores 105
4.23 ELF060 location map 106
4.24 ELF060 read counts 107

4.25 ELF060 biogenomic mass profiles (A) 108
4.26 ELF060 biogenomic mass profiles (B) 109
4.27 ELF060 Pianka similarity scores 111
4.28 ELF032A location map 112
4.29 ELF032A read counts 113
4.30 ELF032A biogenomic mass profiles (A) 114
4.31 ELF032A biogenomic mass profiles (B) 115
4.32 ELF032A Pianka similarity scores 121
4.33 ELF033 location map 123
4.34 ELF033 read counts 124
4.35 ELF033 biogenomic mass profiles (A) 125
4.36 ELF033 biogenomic mass profiles (B) 126
4.37 ELF033 biogenomic mass profiles (C) 127
4.38 ELF033 Pianka similarity scores 132
4.39 ELF033A location map 133
4.40 ELF033A read counts 133
4.41 ELF033A biogenomic mass profiles (A) 134
4.42 ELF033A biogenomic mass profiles (B) 135
4.43 ELF033A Pianka similarity scores 137
4.44 ELF034A location map 138
4.45 ELF034A read counts 139
4.46 ELF034A biogenomic mass profiles (A) 140
4.47 ELF034A biogenomic mass profiles (B) 141
4.48 ELF034A biogenomic mass profiles (C) 142
4.49 ELF034A biogenomic mass profiles (D) 143
4.50 ELF034A Pianka similarity scores 147
4.51 ELF034 location map 149
4.52 ELF034 read counts 150
4.53 ELF034 biogenomic mass profiles (A) 151
4.54 ELF034 biogenomic mass profiles (B) 152
4.55 ELF034 biogenomic mass profiles (C) 153
4.56 ELF034 Pianka similarity scores 155
4.57 ELF054 location map 156
4.58 ELF054 read counts 157
4.59 ELF054 biogenomic mass profiles (A) 158
4.60 ELF054 biogenomic mass profiles (B) 159
4.61 ELF054 biogenomic mass profiles (B) 160
4.62 ELF054 Pianka similarity scores 163
4.63 ELF039 location map 165
4.64 ELF039 read counts 166
4.65 ELF039 biogenomic mass profiles (A) 167
4.66 ELF039 biogenomic mass profiles (B) 168
4.67 ELF039 biogenomic mass profiles (C) 169
4.68 ELF039 Pianka similarity scores 178
4.69 ELF040A location map 180
4.70 ELF040A read counts 181
4.71 ELF040A biogenomic mass profiles (A) 182
4.72 ELF040A biogenomic mass profiles (B) 183
4.73 ELF040A biogenomic mass profiles (C) 184
4.74 ELF040A Pianka similarity scores 188
4.75 ELF041 location map 190

4.76 ELF401 read counts 191
4.77 ELF041 biogenomic mass profiles (A) 192
4.78 ELF041 biogenomic mass profiles (B) 193
4.79 ELF041 Pianka similarity scores 196
4.80 ELF042 location map 197
4.81 ELF042 read counts 198
4.82 ELF042 biogenomic mass profiles (A) 199
4.83 ELF042 biogenomic mass profiles (B) 200
4.84 ELF042 Pianka similarity scores 204
4.85 ELF044 location map 205
4.86 ELF044 read counts 206
4.87 ELF044 biogenomic mass profiles (A) 207
4.88 ELF044 Pianka similarity scores 209
4.89 ELF044A location map 210
4.90 ELF044A read counts 211
4.91 ELF044A biogenomic mass profiles (A) 211
4.92 ELF031 location map 214
4.93 ELF031 read counts 214
4.94 ELF031 biogenomic mass profiles (A) 215
4.95 ELF031 Pianka similarity scores 217
4.96 ELF031A location map 218
4.97 ELF031A read counts 219
4.98 ELF031A biogenomic mass profiles (A) 220
4.99 ELF031A biogenomic mass profiles (B) 221
4.100 ELF031A biogenomic mass profiles (C) 222
4.101 ELF031A biogenomic mass profiles (D) 223
4.102 ELF031A Pianka similarity scores 232
4.103 ELF051 location map 233
4.104 ELF051 read counts 234
4.105 ELF051 biogenomic mass profiles (A) 235
4.106 ELF051 biogenomic mass profiles (B) 236
4.107 ELF051 Pianka similarity scores 242
4.108 ELF045 location map 243
4.109 ELF045 read counts 244
4.110 ELF045 biogenomic mass profiles (A) 245
4.111 ELF045 biogenomic mass profiles (B) 246
4.112 ELF045 Pianka similarity scores 251
4.113 ELF047 location map 252
4.114 ELF047 read counts 253
4.115 ELF047 biogenomic mass profiles (A) 254
4.116 ELF047 biogenomic mass profiles (B) 255
4.117 ELF047 Pianka similarity scores 258
4.118 ELF047A location map 259
4.119 ELF047A read counts 260
4.120 ELF047A biogenomic mass profiles (A) 261
4.121 ELF047A biogenomic mass profiles (B) 262
4.122 ELF047A Pianka similarity scores 264
4.123 ELF049 location map 265
4.124 ELF049 read counts 266
4.125 ELF049 biogenomic mass profiles (A) 267
4.126 ELF049 Pianka similarity scores 270

4.127 ELF046A location map . . . 271
4.128 ELF050 location map . . . 272
4.129 ELF050 read counts . . . 273
4.130 ELF050 biogenomic mass profiles (A) . . . 274
4.131 ELF050 biogenomic mass profiles (B) . . . 275
4.132 ELF050 Pianka similarity scores . . . 276
4.133 ELF053 location map . . . 277
4.134 ELF053 read counts . . . 278
4.135 ELF053 biogenomic mass profiles (A) . . . 279
4.136 ELF053 biogenomic mass profiles (B) . . . 280
4.137 ELF053 Pianka similarity scores . . . 282
4.138 Boxplot of data yield by dominant ecological category . . . 283
4.139 Proportions of Embryophyta reads by rank for all samples . . . 285
4.140 Ratios of higher to lower taxa against total Embryophyta reads . . . 286
4.141 Proportions of Embryophyta reads by higher or lower rank for all samples 287
4.142 Proportions of Embryophyta biogenomic mass by higher or lower rank for all samples . . . 289
4.143 Cross-core Pianka analysis heatmap . . . 291
4.144 Boxplot of data yield by cluster . . . 292
4.145 Cross-core Pianka analysis heatmap (subset) . . . 294
4.146 Cross-core Pianka analysis heatmap (subset) . . . 295
4.147 Map of cores with summarised sedaDNA results . . . 297
4.148 Cores by starting depth with summarised sedaDNA results . . . 298

5.1 MetaDamage profile for Embryophyta reads across all samples . . . 303
5.2 MetaDamage profile for Embryophyta negative controls . . . 303
5.3 MetaDamage profile for non-European Embryophyta reads across all samples . . . 304
5.4 MetaDamage profile for Metazoa (no Primates) negative controls . . . 306
5.5 MetaDamage profile for Metazoa (no Primates) reads across all samples 306
5.6 MetaDamage profile for Metazoa (no Primates) reads including deep sequencing . . . 307
5.7 MetaDamage profile for pre-PIA Primates reads from cores ELF021-60 . 308
5.8 Example manual interpretations of MetaDamage profiles . . . 309
5.9 Embryophyta position-zero values and manual interpretation . . . 312
5.10 Embryophyta read count and position-zero mismatch values . . . 313
5.11 Embryophyta read count and manual interpretation . . . 314
5.12 Stratification analysis for ELF027 . . . 318
5.13 Stratification analysis for ELF059 . . . 320
5.14 Stratification analysis for ELF059A . . . 321
5.15 Stratification analysis for ELF060 . . . 322
5.16 Stratification analysis for ELF032A . . . 323
5.17 Stratification analysis for ELF033 . . . 325
5.18 Stratification analysis for ELF033A . . . 326
5.19 Stratification analysis for ELF034 . . . 327
5.20 Stratification analysis for ELF034A . . . 328
5.21 Stratification analysis for ELF054 . . . 329
5.22 Stratification analysis for ELF039 part 1 . . . 331
5.23 Stratification analysis for ELF039 part 2 . . . 332
5.24 Stratification analysis for ELF040A . . . 333
5.25 Stratification analysis for ELF041 . . . 334
5.26 Stratification analysis for ELF042 . . . 335

5.27 Stratification analysis for ELF044 335
5.28 Stratification analysis for ELF031A part 1 337
5.29 Stratification analysis for ELF031A part 2 338
5.30 Stratification analysis for ELF051 340
5.31 Stratification analysis for ELF045 341
5.32 Stratification analysis for ELF047 342
5.33 Stratification analysis for ELF047A 343
5.34 Stratification analysis for ELF049 344

6.1 MetaDamage profile for all Embryophyta reads (all data) 354
6.2 MetaDamage profile for Fagaceae (original data) 355
6.3 MetaDamage profile for Fagaceae (all data) 355
6.4 Probability distribution of MetaDamage outcomes for Fagaceae (all data)356
6.5 MetaDamage profile for *Corylus* (original data) 357
6.6 MetaDamage profile for *Corylus* (all data) 357
6.7 Probability distribution of MetaDamage outcomes for *Corylus* (all data) 358
6.8 MetaDamage profile for Panicoideae (original data) 359
6.9 MetaDamage profile for Panicoideae (all data) 359
6.10 Probability distribution of MetaDamage outcomes for Panicoideae (all data) 361
6.11 MetaDamage profile for Tilioideae (original data) 362
6.12 MetaDamage profile for Tilioideae (all data) 362
6.13 Probability distribution of MetaDamage outcomes for Tilioideae (all data)362
6.14 MetaDamage profile for Juglandaceae (original data) 363
6.15 MetaDamage profile for Juglandaceae (all data) 363
6.16 Probability distribution of MetaDamage outcomes for Juglandaceae (all data) 364
6.17 MetaDamage profile for *Posidonia* (original data) 365
6.18 MetaDamage profile for *Posidonia* (all data) 365
6.19 Probability distribution of MetaDamage outcomes for *Posidonia* (all data)366
6.20 MetaDamage profile for Triticinae (original data) 367
6.21 MetaDamage profile for Triticinae (all data) 367
6.22 Probability distribution of MetaDamage outcomes for Triticinae (all data)368
6.23 MetaDamage profile for other cereals (original data) 369
6.24 MetaDamage profile for other cereals (all data) 369
6.25 Probability distribution of MetaDamage outcomes for other cereals (all data) 370

7.1 Map of cores with summarised sedaDNA results 373
7.2 Key to ecological categories. 374
7.3 Bathymetry map of the North Sea showing 25 of the dated cores 377
7.4 Platt *et al.* (2017): "Evolution of temperature in the Post-Glacial period according to Groeanland ice cores" 377
7.5 Dated cores by starting depth with summarised sedaDNA results 378
7.6 Taxonomic summaries of dated cores by time period 379

List of Tables

2.1 Ecological categories for Viridiplantae 36
2.2 Ecological categories for Metazoa 36

3.1 Numbers of European and non-European taxa and reads from benchmarking 53
3.2 Results of the Embryophyta accuracy test 58
3.3 Results of the Mammalia accuracy test 58

4.1 Read counts for all samples before processing and when assigned by MEGAN or PIA, and the MEGAN:PIA read count ratio. After MEGAN or PIA assignment, samples were filtered by negative controls and replicates merged. Columns are shaded with most saturation for the most extreme values. The ratio column has red for ratios below 1 (PIA assigned more reads) and blue for above 1 (MEGAN assigned more reads). 61
4.2 MEGAN and PIA taxonomic profiles for ELF054_356 64
4.3 MEGAN and PIA European and non-European assignments for ELF054_356 65
4.4 Example taxonomic results table 68
4.5 ELF022 biogenomic masses for Embryophyta 71
4.6 ELF022 biogenomic masses for Metazoa 71
4.7 ELF027 biogenomic masses for Embryophyta 75
4.8 ELF027 biogenomic masses for Metazoa 76
4.9 ELF059 biogenomic masses for Embryophyta 86
4.10 ELF059 biogenomic masses for Metazoa 89
4.11 ELF059A biogenomic masses for Embryophyta 97
4.12 ELF059A biogenomic masses for Metazoa 102
4.13 ELF060 biogenomic masses for Embryophyta 110
4.14 ELF060 biogenomic masses for Metazoa 110
4.15 ELF032A biogenomic masses for Embryophyta 116
4.16 ELF032A biogenomic masses for Metazoa 118
4.17 ELF033 biogenomic masses for Embryophyta 128
4.18 ELF033 biogenomic masses for Metazoa 130
4.19 ELF033A biogenomic masses for Embryophyta 136
4.20 ELF033A biogenomic masses for Metazoa 136
4.21 ELF034A biogenomic masses for Embryophyta 144
4.22 ELF034A biogenomic masses for Metazoa 146
4.23 ELF034 biogenomic masses for Embryophyta 154
4.24 ELF034 biogenomic masses for Metazoa 154
4.25 ELF054 biogenomic masses for Embryophyta 161
4.26 ELF054 biogenomic masses for Metazoa 162
4.27 ELF039 biogenomic masses for Embryophyta 170
4.28 ELF039 biogenomic masses for Metazoa 175

4.29 ELF040A biogenomic masses for Embryophyta 185
4.30 ELF040A biogenomic masses for Metazoa 187
4.31 ELF041 biogenomic masses for Embryophyta 194
4.32 ELF041 biogenomic masses for Metazoa 195
4.33 ELF042 biogenomic masses for Embryophyta 201
4.34 ELF042 biogenomic masses for Metazoa 202
4.35 ELF044 biogenomic masses for Embryophyta 207
4.36 ELF044 biogenomic masses for Metazoa 208
4.37 ELF044A biogenomic masses for Embryophyta 212
4.38 ELF044A biogenomic masses for Metazoa 213
4.39 ELF031 biogenomic masses for Embryophyta 216
4.40 ELF031 biogenomic masses for Metazoa 216
4.41 ELF031A biogenomic masses for Embryophyta 224
4.42 ELF031A biogenomic masses for Metazoa 228
4.43 ELF051 biogenomic masses for Embryophyta 237
4.44 ELF051 biogenomic masses for Metazoa 240
4.45 ELF045 biogenomic masses for Embryophyta 247
4.46 ELF045 biogenomic masses for Metazoa 249
4.47 ELF047 biogenomic masses for Embryophyta 256
4.48 ELF047 biogenomic masses for Metazoa 257
4.49 ELF047A biogenomic masses for Embryophyta 263
4.50 ELF047A biogenomic masses for Metazoa 263
4.51 ELF049 biogenomic masses for Embryophyta 267
4.52 ELF049 biogenomic masses for Metazoa 269
4.53 ELF050 biogenomic masses for Embryophyta 275
4.54 ELF050 biogenomic masses for Metazoa 275
4.55 ELF053 biogenomic masses for Embryophyta 281
4.56 ELF053 biogenomic masses for Metazoa 281
4.57 Summary of taxonomic results by core 299

5.1 The ten most frequent non-European Embryophyta taxa 305
5.2 MetaDamage results by sample for Embryophyta 310
5.3 Comparing stratification results with other proxies 346
5.4 Mean moisture (%) in LOI samples from five cores 349
5.5 Summary of stratification analysis by core 350

6.1 The closest ten correlates of Panicoideae 360

7.1 Date estimates for Juglandaceae reads 382
7.2 Date estimates for Tilioideae reads 383

B.1 DNA yields for each sample per replicate 411

C.1 Read counts 414

D.1 Taxa found in the Viridiplantae data 418

E.1 Taxa found in the Metazoa (excluding Primates) data 440

Acknowledgements

To Robin Allaby, my supervisor, thank you for the many, many hours spent reading things I've written. This thesis is a beast, but nearly every page has benefited from your thoughtful criticism. More than that, you shaped the direction of this project when I was too wrapped up in details to notice where it was going. I greatly appreciated your insights and considerable expertise.

Rosie Ware, thank you for being so supportive with basically all of my other problems. It's your training that turned me into a competent ancient DNA researcher when I barely remembered how PCR worked. Your knowledge and skill were invaluable. When I use bleach, I will think of you.

Rosie Everett, I was so fortunate to join the group with you. We faced many of the same challenges and worked through them together. I found your great archaeological experience extremely helpful. You also helped me realise how interesting and important archaeology is, despite the lack of dinosaurs. I learned a lot from you and I think we made a great team.

Other group members, especially Lei Liu, Ryan Morrison, and Hannah Macdonald, thank you for your support and advice on all sorts of things.

Lesley Ward and Jeanette Selby of the Genomics Facility, thank you for going out of your way to train me on some very expensive equipment when other people could just pay you to do it properly. I really appreciated your time and patience.

To the C30/C46 cake club, I deeply regret that COVID-19 means we couldn't spend the last year together. This proud institution was instrumental in connecting me with other students and researchers and provided a great deal of advice and cake.

Finally, to Paul Forey, who has been the most supportive partner. You cooked, cleaned, and payed rent during the final months of this thesis, offering timely encouragement and reminders not to burn out in equal measure. You even reviewed the first part of the introduction for me. You don't have to read any more of this if you don't want to. Just stop here.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work was performed at the University of Warwick in the School of Life sciences. It was supervised by Professor Robin G. Allaby during the period of October 2017 to July 2021.

Chapter 3 (PIA) was previously published by the author in a peer-reviewed scientific journal as Cribdon *et al.* (2020). Preliminary results were also presented in poster form at the European Association of Archaeologists annual meeting (September 2018), 8th International Symposium on Biomolecular Archaeology (September 2018), and the annual conference for the Society of Molecular Biology and Evolution (July 2019).

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

- Figures 1.2 and 1.4 in Chapter 1 (Introduction) are reproduced from Walker *et al.* (2020) and Robb (2013) with permission. Figure 1.3 is reproduced in Chapter 1 and Chapter 7 (Discussion) from Platt *et al.* (2017) under a CC BY 4.0 license.
- Sediment cores were collected by Europe's Lost Frontiers colleagues.
- Sediment sampling for DNA analysis was shared with Roselyn Ware and Rosie Everett under the direction of Martin Bates and Simon Fitch.
- DNA extraction and library preparation for samples from cores ELF001-20 was performed by Roselyn Ware. Data from these samples is included in Chapter 6 (Mesophilic taxa and human disturbance indicators).
- The PIA program (Chapter 3) was initially developed by Robin Allaby, Oliver Smith, and Roselyn Ware and published in Smith *et al.* (2015).
- The biogenomic mass and stratification analysis were initially developed by Robin Allaby and Roselyn Ware and published in Gaffney *et al.* (2020).
- Core reports in Chapter 5 (Authentication results) were produced by Martin Bates, Richard Bates, and John Whittaker.
- Loss-on-ignition data in Chapter 5 was produced by Martin Bates.
- The MetaDamage analysis (Chapters 5 and 6) was initially developed by Rosie Everett and Logan Kistler and is in preparation for publication.
- Dates presented in Chapter 7 were produced by Martin Bates and Derek Hamilton (radiocarbon) and Tim Kinnaird (OSL).

Summary

Until $\sim$7000-8000 years ago, much of the southern North Sea was a coastal plain now named Doggerland. Doggerland was inhabited by Mesolithic people and is expected to be a rich source of archaeological information. This thesis reconstructs environmental changes through time in a region of Doggerland using sedimentary ancient DNA (sedaDNA) from marine cores.

Plant and animal DNA profiles are produced for individual samples using a novel, highly accurate sequence assignment method. These are used to infer environmental conditions and changes through each core. Patterns across cores are then examined. The validity of these results is evaluated by searching for age-associated DNA damage patterns and evidence that DNA has not moved vertically between sediment strata. Certain unexpected taxa that would be particularly informative for climate or human activity are investigated in greater depth where data allows. Finally, the authenticated results are drawn together and re-examined in the context of preliminary dating information from Europe's Lost Frontiers colleagues.

Overall, the cores appear to have captured environments through the inundation process. There are woodland and freshwater profiles; some are grass dominated and may represent reed beds; others combine seagrasses and small terrestrial signals to imply a brackish coastal environment. The project presents a detailed insight into an ancient submerged landscape, demonstrates the capability of emerging sedaDNA analysis methods, and provides a platform for further investigation of the relationship of sedaDNA to other environmental proxies.

Chapter 1

Introduction

At the end of the last Ice Age, the volume of water locked in ice sheets caused sea levels around the world to be dramatically lower than today. The resulting additional land area presented new opportunities to ancient peoples, not only as connections but as viable landscapes in their own right, until subsequent warming rose sea levels once more and submerged these areas along with evidence of their inhabitants. Once such landscape occupied the present southern North Sea (Walker *et al.* 2020) and is today known as Doggerland.

It has long been suspected that the North Sea was once inhabited. Bones from terrestrial animals, peat (which only forms in terrestrial settings), and occasional artefacts have been dredged up by fishing vessels for over a century (Peeters and Momber 2014). In 1913, Clement Reid published an early reconstruction of Doggerland based on dredged finds and submerged forests along present-day coasts (figure 1.1), noting that the high Dogger Bank in particular may have been attractive to prehistoric humans. Clark (1936) furthered this idea and presented Doggerland as a heartland of human occupation. Specifically, he argued that the distribution of sites for a major contemporaneous culture, the Maglemosian, and their preference for lowland wetlands suggested that much of their archaeological record would now be under the North Sea.

However, direct archaeological investigation was not feasible at the time, and interest in Doggerland waned over the 20th century. Its general perception simplified to little more than a land bridge, a place of migration but no particular settlement. It was only in 1998 that a key paper by Bryony Coles reinstated its importance as a rich and habitable landscape, nestled between the cold uplands above current sea level (Coles 1998). In recognition of Reid's work on the Dogger Bank and to signify its human importance, Coles coined the name Doggerland, suggestive of a lost country.

It is now widely accepted that Doggerland was a major region of occupation during the Holocene (Peeters and Momber 2014). In human terms, the beginning of the Holocene took place during the Mesolithic period, in which European people lived as hunter-fisher-gatherers, but the Holocene also saw the emergence of the Neolithic and its crucial component, agriculture (Robb 2013). Neolithic elements include domesticated plants and animals, ceramics, and more permanent construction, although these were not necessarily adopted simultaneously (Tresset and Vigne 2011). The submergence of Doggerland broadly coincides with the Neolithisation of Northwest Europe, suggesting that the landscape may also hold evidence on this important transition (Gaffney *et al.* 2017).

Doggerland is now the main region under investigation by Europe's Lost Frontiers, a project which aims to reconstruct submerged landscapes around the UK (Gaffney *et al.* 2017) and through which this thesis was produced. Building upon previous research but on a much greater scale, Europe's Lost Frontiers is using new technologies to

Figure 1.1: Fig. 4 from Reid (1913), showing a coastline reconstruction based on the lowest submerged forest in that study.

see past the marine environment, which although a barrier to traditional archaeology, has protected underlying land surfaces from much of the disturbance expected in a terrestrial setting. The North Sea therefore holds a particularly well-preserved record of recent millennia (Fitch 2018). Ancient land surfaces across much of the North Sea basin have been mapped using seismic surveys, allowing reconstruction of the underlying topography. Targeted sediment coring provided material for multi-proxy environmental reconstruction (Gaffney *et al.* 2020). All of this data is informing new models of sea level, ecology, and human activity (Walker *et al.* 2020). We are just beginning to reveal this environmental record and understand what Doggerland would have been like, how it changed as the climate warmed, and how this may have affected its people.

1.1 Current knowledge of the Doggerland environment

A primary factor in the reconstruction of the changing environments across Doggerland is our knowledge of sea levels over time. The timing and extent of inundation broadly determined when different regions were transformed from terrestrial to coastal and then fully marine environments. What was the overall pattern of inundation and when did regions become submerged? The most complete sea level model recently published can be found in Sturt *et al.* (2013; figure 6). Broadly, at the start of the Holocene, Doggerland probably reached as far north as the elevated Dogger Bank and formed a lowland plain across the North Sea basin. Valleys running south of the Bank and out of what would become the English Channel were among the first areas to submerge, fragmenting the landscape. The final remnants comprised Dogger Bank itself, a peninsula stretching east from Lincolnsire and Norfolk, and a strip along the Netherlands.

Europe's Lost Frontiers is in the process of producing updated sea level reconstructions incorporating new data, including seismic mapping of much of the North Sea basin. Coastline reconstructions at four key time points were published in Walker *et al.* (2020) and are reproduced with permission in figure 1.2. Walker *et al.* tell a similar story to Sturt *et al.* but propose two new terms for the region as inundation progresses, allowing for more intuitive interpretation as Doggerland became sea. Around 10,000 BP (years before present), they acknowledge a Doggerland proper (figure 1.2a). However, between 9,000-7,000 BP (figure 1.2b-c), Walker *et al.* refer to the northern uplands as Dogger Island and the remaining, highly fragmented landscape as the Doggerland Archipelago. By 7,000 BP (figure 1.2d), they recognise the loss of almost all islands but retention of a substantial coastal margin by referring to Dogger Littoral. For much of our period of interest, the Doggerland region was probably more marine than terrestrial.

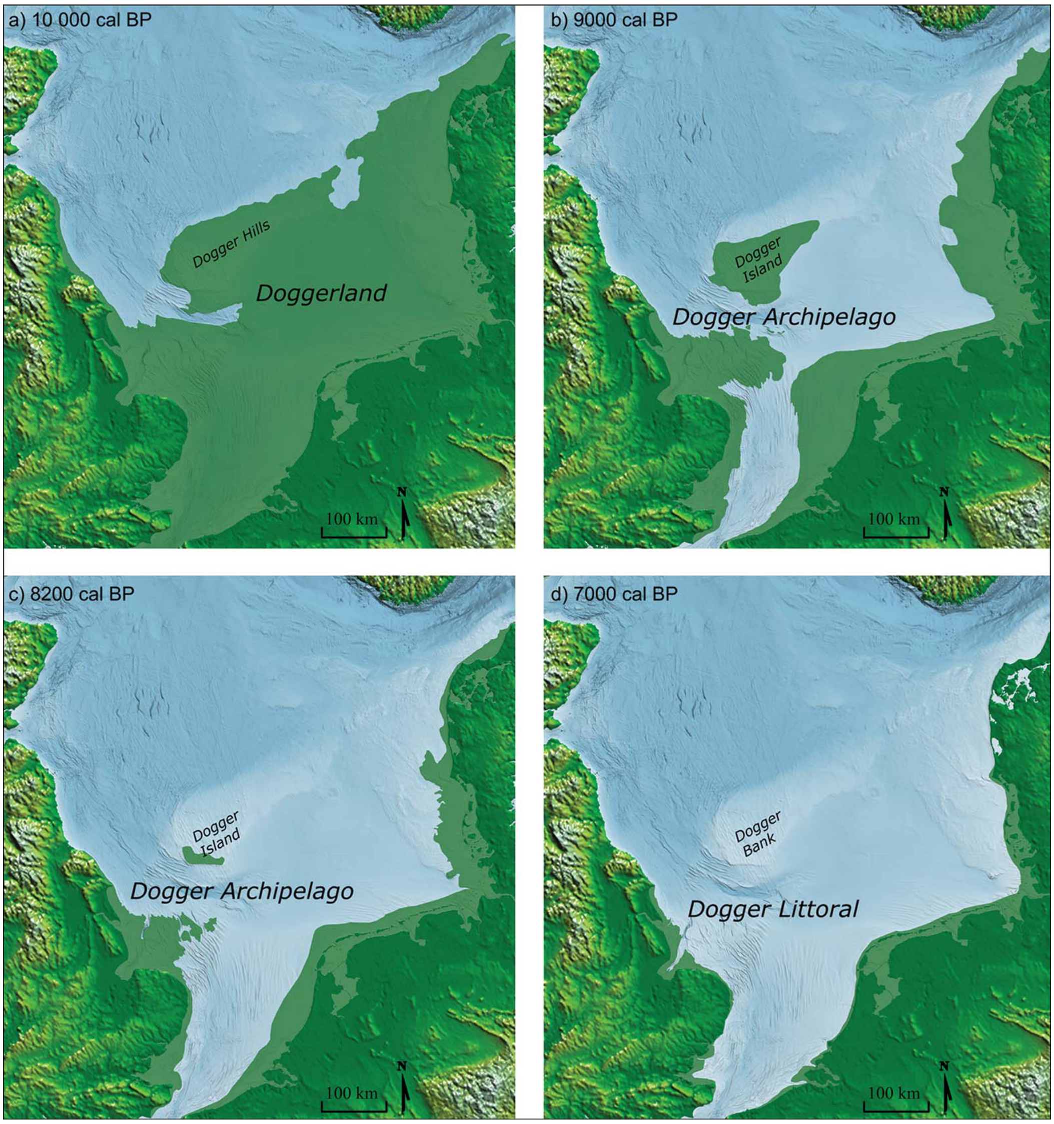


Figure 1.2: Walker *et al.* (2020) Fig. 2: North Sea coastline reconstructions. a) Doggerland c. 10,000 cal BP; b) Dogger Archipelago c. 9,000 cal BP; c) Dogger Archipelago c. 8,200 cal BP; d) Dogger Littoral c. 7,000 cal BP. Created by Merle Muru and used with permission.

This is the landscape-level story of inundation as currently known. However, recent modelling studies including Sturt *et al.* (2013) and Walker *et al.* (2020) have stressed the importance of local variation. Dramatic rises in sea level may have occurred, but impacts would likely have varied by region. Both Sturt *et al.* and Walker *et al.* argue this in particular for the Storegga slide tsunami. Around 8,150 BP, an underwater landslide off the coast of Norway generated an enormous tsunami that left extensive deposits on North Atlantic coastlines (Bondevik *et al.* 2012). Weninger *et al.* (2008) used these deposits and bathymetric modelling to argue that the tsunami caused the final flooding of Doggerland, with devastating effects for any remaining inhabitants. However, neither the Sturt *et al.* or Walker *et al.* models link the tsunami to a long-term, landscape-level rise in sea level. Walker *et al.* also note new deposits from the southern North Sea (Gaffney *et al.* 2020), a region previously without direct evidence, that demonstrate a variable impact dependant on local topography. Therefore, while sea level models can suggest patterns of inundation across a landscape, their ability to predict the effects of even dramatic events at smaller scales is limited.

For detailed snapshots of local environments, we instead turn to sediment cores. Cores can provide a time-series of multiple environmental proxies, including the sediment itself, pollen, diatoms, ostracods, plant and animal macrofossils, chemical biomarkers, and DNA. Unlike chance finds recovered from the seabed, cores also come with valuable contextual information: an exact location and depth, and potentially dates through either radiocarbon or optically-stimulated luminescence (OSL; calculating the time since certain mineral grains were last exposed to sunlight) dating.

The relatively few studies of Doggerland sediment cores have generally reported results typical of contemporary northwest Europe, with the vegetation influenced by the changing climate as well as local sea level. Figure 1.3, reproduced from Platt *et al.* (2017), illustrates the broad temperature trends from the Late Pleistocene to the present day via reconstructed Greenland air temperatures. The end of the Pleistocene saw dramatic temperature changes: the cool temperatures of the Last Glacial Maximum gave way to rapid warming approximately 14,700 BP at the beginning of the Bølling interstadial, followed by the short Older Dryas cool period around 14,100 BP, subsequent warming into the Allerod interstadial, and cooling again from 12,900 BP into the Younger Dryas. The Younger Dryas ended with rapid warming. This marked the beginning of the Holocene at approximately 11,700 BP. A period of more gradual warming followed, interrupted by the short cold event around 8,200 BP. Though difficult to see in figure 1.3, this warming culminated in the Holocene Thermal Maximum (Martin *et al.* 2020), which peaked at somewhat higher temperatures than the present day at approximately 8,000 BP. Finally, a gradual decline led to the current temperatures by around 5,000 BP.

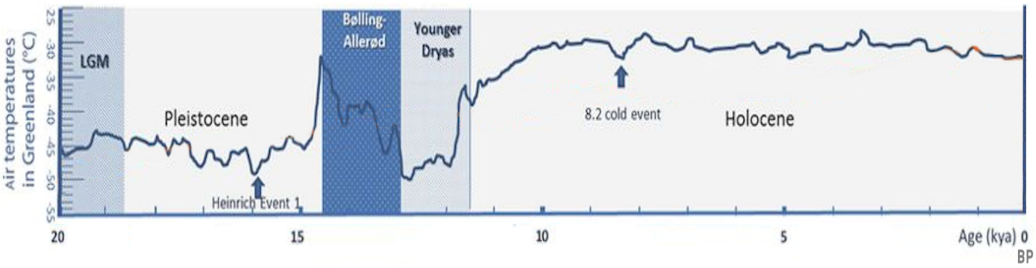


Figure 1.3: "Evolution of temperature in the Post-Glacial period according to Groeanland ice cores" by Platt *et al.* (2017), reproduced without alteration under a Creative Commons 4.0 license (https://creativecommons.org/licenses/by/4.0/). Note that the Holocene began around 11.7 ka, at the end of the Younger Dryas stadial, but this boundary is marked around 11.5 ka in this figure.

The existing environmental record from Doggerland cores broadly follows these climatic trends, beginning with grass-dominated tundra vegetation followed by mostly birch or pine woodland in the Younger Dryas (Krüger *et al.* 2017). As Holocene warming proceeded, woodland diversified and eventually became dominated by hazel (Wolters *et al.* 2010, Krüger *et al.* 2017, Brown *et al.* 2018). Many of the sediment cores in these studies also record inundation: that in Wolters *et al.* (2010) moves from woodland through carr to a brackish reedbed, and cores in Smith *et al.* (2015), Geary *et al.* (2017), Brown *et al.* (2018), and Gaffney *et al.* (2020) have freshwater signals replaced by brackish or estuarine. The overall picture in the Holocene is of vegetation not dissimilar to today with an increasing proportion of freshwater and eventually marine habitats as the water table rose and the sea came in.

1.2 Human effects of inundation

The effects of this environmental change on the people of Doggerland have been thoroughly debated. As Sturt *et al.* (2013) and Walker *et al.* (2020) note, change would have been locally variable, and effects would also have depended on a population's lifestyle and adaptability. How did these people live? Artefacts have been recovered from the North Sea for decades that demonstrate the presence of Mesolithic people, but these typically lack archaeological context, limiting the information available (Peeters and Momber 2014, Amkreutz and Spithoven 2019).

One approach is to extrapolate from more accessible archaeological sites around the North Sea basin. Fischer *et al.* (2007) examined the stable isotope ratios of Mesolithic humans and dogs from Denmark. Isotope ratios can indicate the proportion of freshwater, marine, and terrestrial foods consumed. Assuming that dogs eat similar foods to their humans, they see good evidence for a marine-based diet. Bonsall *et al.* (2009) review similar results for Mesolithic humans of Britain and other sites from southern Scandinavia. Dupont *et al.* (2009) combined isotope analysis with remains of food waste in shell middens to show further dependence on marine resources in Late Mesolithic Brittany. We could use these surrounding sites to infer a similar lifestyle in Doggerland.

However, a more recent study by Van der Plicht *et al.* 2016 examined the stable isotope ratios of Mesolithic human remains from the southern North Sea itself. In contrast to results from surrounding sites, freshwater resources appeared most important to these people, although both freshwater and marine increased in importance over time at the expense of terrestrial. Van der Plicht *et al.* go on to argue that this move away from terrestrial foods shows successful adaptation to rising sea levels. The rising water table could have maintained a sizeable wetland zone ahead of the advancing coastline, while dry land would only decrease. Mesolithic people developed their already productive exploitation of the rich wetland environments instead of the declining terrestrial resources.

Other researchers place greater emphasis on the negative effects of land loss. A recent study by Crombé (2019) correlates radical technological changes in stone weapons from northern France, Belgium, the southern Netherlands, and western Germany with short but abrupt cooling events. He argues that the technological changes were not to do with function, but with social cohesion in the face of competition. Fewer resources led to increased stress and territoriality between groups. Crombé then extends this argument from climate to sea level change, arguing that the lack of exploitation of marine resources found by Van der Plicht *et al.* 2016 actually demonstrates a lack of successful adaptation. While the rising water table may generally have maintained wetlands ahead of the advancing sea, the most persistent habitat would be the coast itself. An extended coastline could have maintained substantial resource-rich territory (Coles 2000), but if people did not utilise it, it would not go far to mitigate the loss of total land area.

In fact, Dupont *et al.* (2009) argue that even successful exploitation of marine resources, as they found in Brittany, was not necessarily helpful during inundation. They cite research on the geomorphological effects of sea level rise that shows important intertidal habitats such as saltmarsh and mudflats can disappear if the relief of the new coastline is no longer suitable (e.g. Galbraith *et al.* 2002, Hughes 2004). Even if a substantial coastline was maintained during inundation, the amount of *productive* territory may have declined on the coast at a similar rate to inland. No matter their subsistence, the people of Doggerland would eventually have felt the squeeze.

Nevertheless, resource base is not the only factor that determines how people cope with change. Leary (2009) favours a more nuanced approach, emphasising that the impact of natural hazards on people depends very much on social factors. He argues that what appear to be disasters with hindsight may not necessarily have been stressful if changes were gradual enough to be perceived as normal. Communities can be more resilient or vulnerable depending not just on resources, but adaptability of behaviour and links with neighbours, which are more difficult to infer from the limited evidence. Just like the process of inundation itself, we should expect there to have been significant local variation in responses.

1.3 Doggerland and Neolithisation

This sets a complex stage for a second key change in the history of Doggerland: the arrival of the Neolithic. The farming system that would dominate Europe originated in the Near East and spread out towards the northwest in fits and starts (Fort 2015, Silva and Vander Linden 2017). Figure 1.4, reproduced from Robb (2013), plots the spread of Neolithic activity across western Europe. Note the decrease in pace as the Neolithic moves further north. Robb collects existing estimates to suggest that Neolithic activities had become established in the British Isles some time between 6,450 and 5,950 BP. This is supported by Bayliss *et al.* (2011), a Bayesian modelling study also based on existing dates, that suggested an initial arrival date in the latter half of this range, between ~6,050 and 5,950 BP.

The key point is that, according to current estimates, by the time the Neolithic arrived in Britain, Doggerland was essentially gone. Recall that the Walker *et al.* (2020) model shows the submergence of the Dogger Archipelago occurring not long after 8,200 BP. This includes the region off Norfolk where most of the sediment cores from this project were taken. By 7,000 BP, additional land was more or less restricted to an extended coastline. The early British farmers at ~6,500-6,000 BP would have had a similar level of access to Doggerland as modern archaeologists do.

However, there is evidence of a Neolithic presence in other parts of the North Sea basin at least several centuries before Britain. The final push across the Channel seems to have followed a period of stasis (Robb 2013, Fort 2015, Betti *et al.* 2020). There are relatively well-supported dates for Neolithic activity around 7,450 BP in western France (Marchand 2007), 7,050 BP in the north, 7,250 BP in the Netherlands (Tresset and Vigne 2011), and 6,650 BP in Belgium (Meylemans *et al.* 2018). Tresset and Vigne suggest that part of the reason for the slowing of the spread at the northwestern fringes of Europe, including around the North Sea, was that the Neolithic influence had become diluted. After several thousand years and many hundreds of kilometres, it may involved fewer migrants and their objects and practices would have appeared less foreign. Mesolithic cultures do seem to have persisted for longer in these regions (Robb

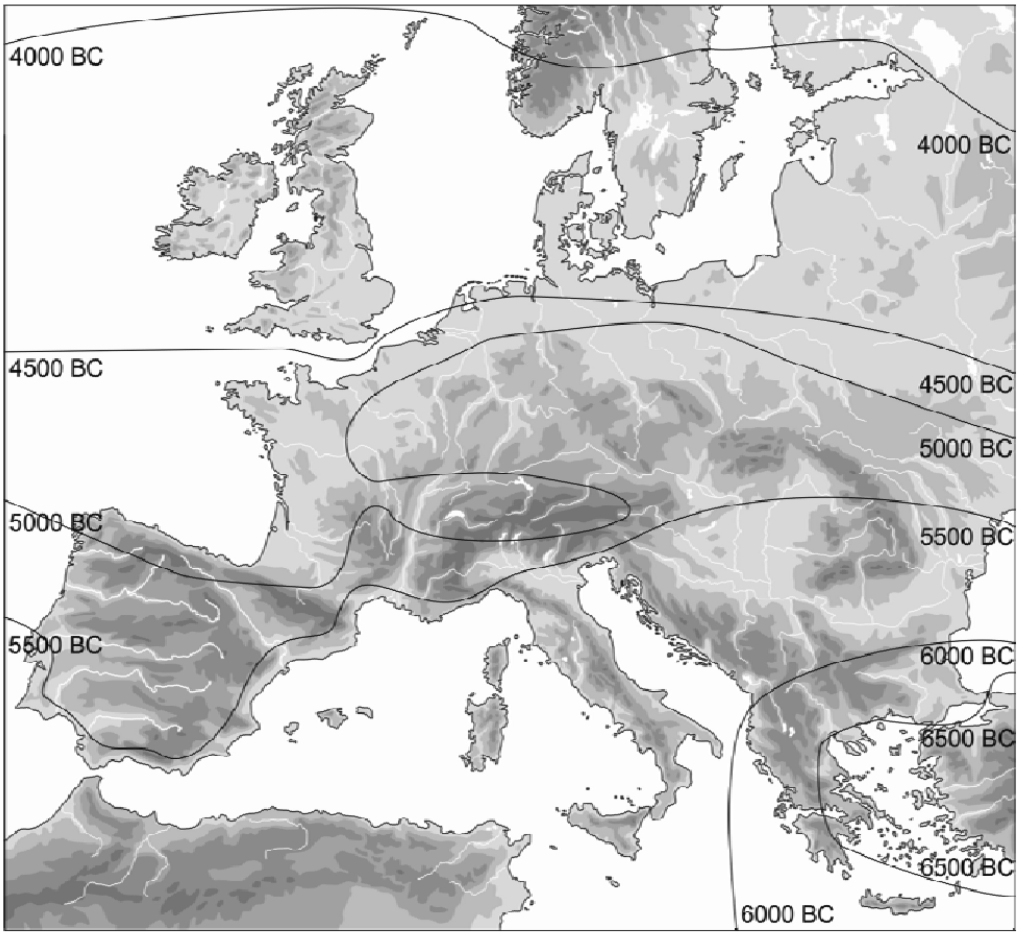


Figure 1.4: Robb (2013) Fig. 1: The spread of the Neolithic in Europe, dates in cal. BC. Reproduced with permission from University of Chicago Press.

2013, Bollongino *et al.* 2013, Rowley-Conwy 2014). Indeed, other authors argue that complex, well-adapted Mesolithic societies inhabiting coastal territories - environments inferior or at least unfamiliar to farmers - would themselves contribute to the slowdown by having less reason to change their way of life. This is despite exchange of goods and people with Neolithic neighbours (Coles 1998, Svizzero 2015). Furthermore, the cooler, wetter climate seems to have presented a challenge to crops, which had little time to adapt after being carried out of the Near East (Colledge *et al.* 2005). Betti *et al.* (2020) argue that a lower-quality growing season explains regions showing a marked slowdown in Neolithisation, including the Atlantic coast. These regions also showed significantly greater admixture between incoming Neolithic farmers and indigenous hunter-gatherers than elsewhere. Betti *et al.* suggest that struggling farmers were particularly dependent on hunting and gathering themselves, resulting in greater contact with the local population. This increased admixture, and potentially dependence on wild foods, continued even after the Neolithic expansion resumed. Although farming spread to regions around the southern North Sea relatively quickly, it was confronted with difficult growing conditions and well-adapted, potentially more resistant Mesolithic societies. The final push to Britain did not happen for quite some time.

Nevertheless, there is evidence of contact between continental farmers and Mesolithic societies in Britain and Ireland during this hiatus. Garrow and Sturt (2011) use evidence from islands around Britain to argue that the North Sea was not as much of a barrier as once thought (e.g. Coles 1998), and that regular boat travel may have occurred in both directions. Tantalisingly, Smith *et al.* (2015) present DNA evidence for wheat in a submerged site on the south coast of England dated to 8,000 BP, which the authors interpret as a sign of trade with farmers on the Continent. At this time, there was still a substantial Dogger Littoral and potentially islands that may hold similar evidence of Neolithic influence before it was fully adopted. Therefore, despite the Neolithic transition not occurring in Britain and Ireland until around 6,000 BP, there may be evidence of Neolithic products or activity beyond the present coastline on either side of the North Sea. It is certainly something to consider when studying Doggerland material.

1.4 SedaDNA and its application in the North Sea

An important emerging technology used in Europe's Lost Frontiers has been sedimentary ancient DNA (sedaDNA) analysis: the focus of this project. Ancient DNA (aDNA) is that which persists after the death of the organism. After death, the genome is no longer protected from biochemical or physical factors that threaten its stability, resulting in characteristic damage patterns (Dabney *et al.* 2013a): low quantity, short fragment lengths (often <50 bp; Dabney *et al.* 2013b), and miscoding lesions. These miscoding lesions primarily involve deamination of cytosine bases to uracil, which manifests as C-to-T substitutions in sequence data (Fulton and Shapiro 2019). Typical aDNA studies attempt to extract and analyse this DNA from physical remains of a particular organism. SedaDNA instead involves DNA from any of the organisms that contributed material to a sediment sample. The resulting metagenome (comprising genomic material from multiple organisms) can be used as a proxy for the local biological community.

Sediment has several advantages over physical remains as a source of aDNA. It is abundant and relatively easy to obtain, making destructive sampling more justifiable (Thomsen *et al.* 2009). Sources of sedaDNA are thought to include transient products such as urine, faeces, or shed integument in animals, and root cap cells or leaf litter in plants (Pedersen *et al.* 2015). Organisms typically contribute much more of these materials to the sedimentary record over their lifetime than physical remains after death, making detection of organisms through DNA more likely than through body remains, especially for rare organisms (Haile *et al.* 2009). DNA preservation can also be enhanced in sediment by its tendency to adsorb onto minerals, particularly clays (Coolen and Overmann 2007). This is thought to provide protection from nucleases, reducing DNA damage, although the mechanisms are not fully understood (Torti *et al.* 2015).

However, sedaDNA also brings additional challenges. With identifiable physical remains, it is possible to infer which sequences are derived from the target organism simply by aligning the data to a relevant genome (e.g. Van der Valk 2021). However, in sedaDNA, there is often not a single known target organism to compare sequences to. A huge range of organisms can contribute material to sediment. To match sequences to taxa, each sequence is typically compared to a database of known sequences, making identification strongly dependent on the completeness of those databases. Unfortunately, a complete database would require the full genome of every organism extant during the relevant period, which is unlikely to be achieved in the foreseeable future. Chapter 3 (PIA) discusses this in detail, but in short, the incompleteness of current databases leads to sequences being left unassigned or, worse, being assigned incorrectly. The problem has previously been addressed by curating relatively complete but limited databases. These databases can be restricted to organisms with very small genome size, typically bacteria (Velsko *et al.* 2018), or to short, taxonomically-informative "barcode" regions of genomes (Zinger *et al.* 2019). While a comparison algorithm can then assign sequences to these database more accurately, DNA that is either nonmicrobial or outside of barcode regions is discarded. This project involved developing an alternative method that can potentially assign any sequence with high accuracy while using incomplete databases.

Another challenge of metagenomic data is identifying contamination. Modern contamination has always been a problem in ancient DNA, but robust protocols have been developed to minimise its impact (Gilbert *et al.* 2005, Shapiro *et al.* 2019, Orlando *et al.* 2021). In particular, sterile conditions reduce contamination, and negative controls characterise what is left. After sequencing, ancient sequences can be identified using their characteristic damage patterns (Jónsson *et al.* 2013, Skoglund *et al.* 2014). However, existing age-authentication tools rely on comparing sequences to a single reference genome. Therefore, much like with assigning sequences to organisms, this project involved developing a new tool to analyse the damage patterns of metagenomic data effectively (Chapter 2: Main materials and methods, subsection 2.3.1).

SedaDNA also presents an additional contamination issue in the vertical movement of DNA between strata ("leaching"). This dissociates the DNA from its context, potentially mixing ancient DNA of different ages and making interpretation difficult. DNA movement is thought to primarily occur via water movement (Giguet-Covex *et al.* 2014), but the taphonomy of DNA in sediment is not well characterised (Birks and Birks 2016). A third new method developed through this project involves searching for evidence of stratification in DNA profiles between samples along a sediment core, ruling out leaching if profiles are sufficiently different (Chapter 2: Main materials and methods, subsection 2.3.2).

Finally, unlike typical physical remains, sediments can harbour live, metabolically active bacteria (Rivkina *et al.* 2000, Johnson *et al.* 2007). Any DNA recovered from these cells will be biologically modern, so will not show age-related damage patterns, but the bacteria may represent an in-situ ancient community (Inagaki *et al.* 2015). This project focused on plants and animals because of their greater use for environmental reconstruction, but also to avoid this authentication problem. Furthermore, living bacteria can release exonucleases that degrade DNA in their environment, although the adsorption of DNA to minerals provides some protection (Torti *et al.* 2015). This and the fact that DNA from living cells is less degraded can result in their DNA being present in much greater quantities than other endogenous DNA, wasting sequencing effort. Indeed, a large proportion of the sedaDNA data from this project was broadly assigned to bacteria.

The overall preservation potential for sedaDNA is expected to vary with sediment type. Frozen sediments were among the first to be targeted because of their low, stable temperatures (Willerslev *et al.* 2003), which appear to be the most important factors for DNA preservation (Lindahl and Nyberg 1972, Kistler *et al.* 2017). Ice and permafrost studies have since produced many successful results (e.g. Willerslev *et al.* 2007, Haile *et al.* 2009, Willerslev *et al.* 2014), and although via physical remains, permafrost also holds the current record for the oldest confirmed aDNA at over one million years (Van der Valk *et al.* 2021). The lack of liquid water also appears to protect frozen sediment from DNA leaching (Hansen *et al.* 2006), although there is still evidence of metabolically active bacteria (Rivkina *et al.* 2000). The main limitation of frozen sediment is simply its availability: its restriction to high latitudes or altitudes means that the variety of environments represented in frozen sedaDNA will also be limited. Terrestrial sedaDNA studies from other locations have mostly focused on caves (Rawlence *et al.* 2014), where temperatures can also be cool and stable in deeper regions. The sediment may also be protected from water, potentially reducing hydrolytic damage (Willerslev and Cooper 2005) and making leaching unlikely (Gilbert *et al.* 2008). However, Haile *et al.* (2007) report dry cave sediment where otherwise intact stratification was broken by its use as a sheep latrine. Finally, although again geographically limited, caves include archaeologically important sites of human habitation (Slon *et al.* 2017, Ardelean *et al.* 2020).

For a more general environmental picture, researchers have turned to lake sediments. There is already a long history of using material in lake sediment as proxies for environmental reconstruction, including pollen, other fossils (Pedersen *et al.* 2013), and algal pigments (Stivrins *et al.* 2018). Now many studies also involve sedaDNA, and it has generally been shown to be a useful complementary method with a comparatively local signal (Capo *et al.* 2021). Lake sediments are relatively promising sources of sedaDNA. The accumulation and decomposition of organic matter can create an anoxic layer at the bottom (Sobek *et al.* 2009), promoting DNA preservation (Corinaldesi *et al.* 2008) and minimising sediment disturbance through bioturbation (Pansu *et al.* 2015). Capo *et al.* (2021) suggest that fast sedimentation rates result in DNA entering the anoxic zone more quickly, which may explain the surprisingly good preservation in some temperate and tropical lakes. However, temperature is still a key factor. The deepest parts of a lake can enjoy cool and stable temperatures: water is most dense at 4°C, so forms a layer on the bottom unless there is significant mixing (Parducci *et al.* 2017). However, sediment from smaller bodies of water (including rivers) is likely to be close to ambient temperature (Haile *et al.* 2009), and latitude and altitude remain important (Fang and Stefan 1998).

Marine sediment combines the thermal stability of very large water bodies with further biochemical benefits for DNA preservation. Sufficiently deep sea sediments worldwide should maintain a stable 4°C layer because of cold water circulating from the poles. At least the top layer of marine sediment would be saturated with seawater, the salt in which further reduces hydrolytic damage by occupying the polar ends of water molecules, leading to fewer reactions with DNA (Lindahl and Nyberg 1972). Marine sediment is ubiquitous and is expected to contain vast amounts of DNA: over ten times as much as living biomass according to Dell'Anno and Danovaro (2005). Accordingly, there have been many successful sedaDNA studies, mostly sharing with lake sediment a focus on plankton and microbes (Armbrecht *et al.* 2019). Raniello and Procaccini (2002) also successfully recovered aDNA from seagrass material in marine sediment, although they did not target the sediment itself.

This project involves samples from marine cores. Many samples consist of sediment deposited in a marine context, but most cores also extend into deeper freshwater or terrestrial strata. The DNA profiles (Chapter 4: Taxonomic results) support this variety of sources through the diversity of taxa that sequences have been assigned to. These results are based on a final innovation developed through this project that offers a more nuanced means of quantifying DNA data than simply counting reads. First published in Gaffney *et al.* (2020), ecological reconstruction in this project is based not only on raw read counts, but counts adjusted for genome size, and therefore more representative of biomass. We term this measure biogenomic mass (Chapter 2: Main materials and methods).

1.5 Scope and aims of thesis

This thesis uses sedaDNA from North Sea cores to reconstruct ancient environments, focusing on a submerged river valley off the present east coast of Britain. It charts changes in the plant and animal DNA profiles over time, inferring ecological transitions that include marine inundation. It also further investigates key taxa, particularly those not native to the region today, that suggest particular climatic conditions or human activity. All interpretations are supported by a rigorous system of assigning reads to taxa, excluding contaminants, and authenticating ancient sequences. This project has involved the development of new methods that push the boundaries of ancient metagenomics. It aims to demonstrate the potential of sedaDNA for understanding submerged landscapes.

Chapter 2

Main materials and methods

2.1 Data collection

2.1.1 Sediment core collection

The sediment cores used in this project are a subset of those analysed for sedaDNA for Europe's Lost Frontiers. A map of all North Sea cores is shown in figure 2.1. The first round of cores, numbered ELF001-20 ("ELF" for "Europe's Lost Frontiers"), plotted a transect across what would have been the Dogger Archipelago (Walker *et al.* 2020). SedaDNA from these cores was primarily analysed by Roselyn Ware, although results from a subset of samples are included in Chapter 6 (Mesophilic taxa and human disturbance indicators) and Chapter 7 (Discussion).

The 27 cores that form the main focus of this project were taken in the second round, numbered ELF021-60. These are detailed in figure 2.2. Two cores, ELF022 and ELF027, were taken towards the north of the coring region: ELF022 on Dogger Bank, and ELF027 at the edge of the Outer Silver Pit. The remainder were taken in a northwest-southeast transect along a palaeochannel approximately 50 km north of the present Norfolk coast, referred to in Europe's Lost Frontiers publications as the Southern River system (Gaffney *et al.* 2020). Also part of the Dogger Archipelago, this region was probably among the last to be fully submerged, remaining above sea level until at least $\sim$8,200 BP (Walker *et al.* 2020). It was hoped that material and water collecting in the river would inform environmental proxies, allowing reconstruction of the surrounding catchment area. The system is also of high anthropological potential: the river, estuary, and coastline would have offered a variety of resources to Mesolithic inhabitants, who may have been particularly dependent on freshwater (Van der Plicht *et al.* 2016) and/or marine (Bonsall *et al.* 2009) foods. The flat lowlands preceding the coast may also have been attractive for early agricultural activities, if such ideas had arrived before the sea. It is certainly a promising region to search for indicators of environmental transition and human activity. The Southern River cores are detailed in figure 2.3.

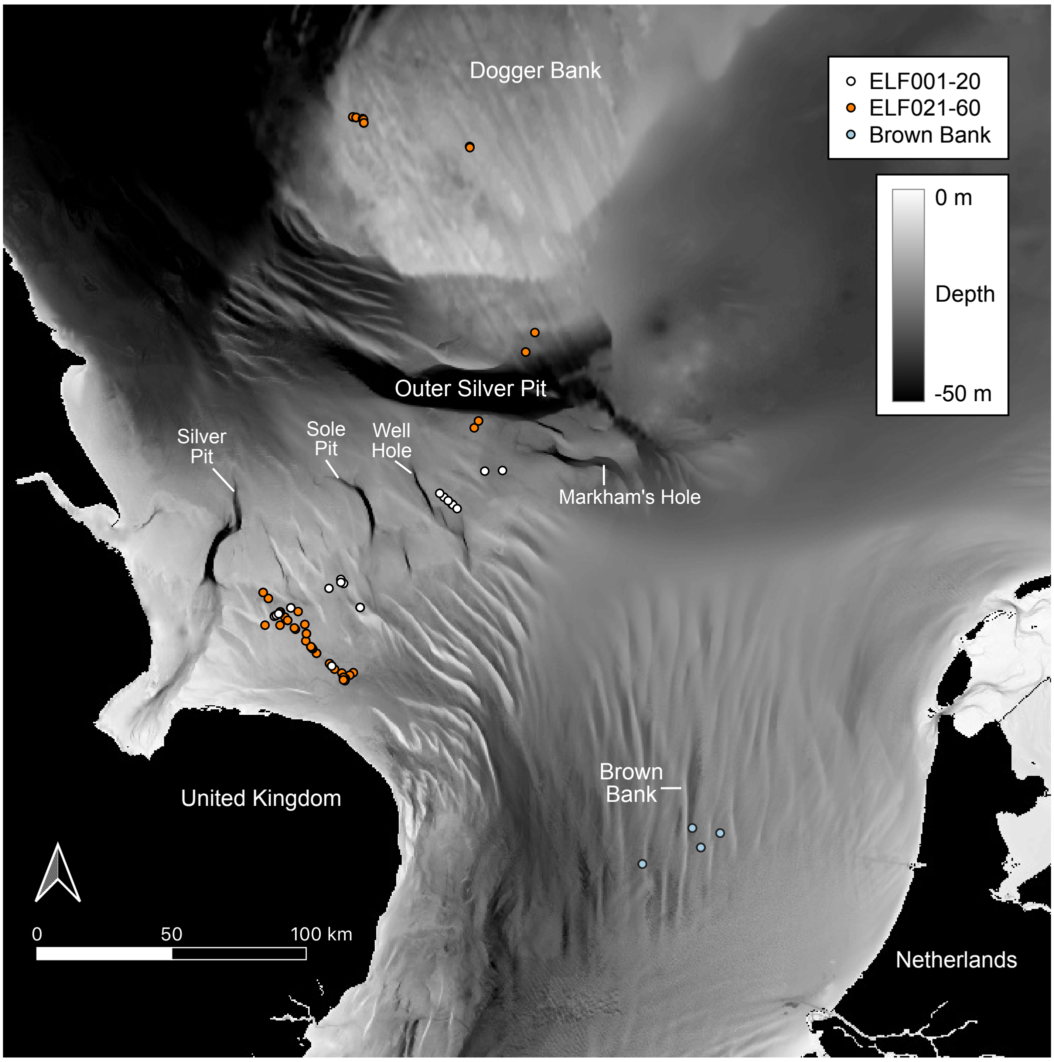


Figure 2.1: Bathymetry map showing core locations for the three rounds of sampling in the North Sea as part of the Europe's Lost Frontiers project. The first round took cores ELF001-20, the second ELF021-60, and the third took cores around the Brown Bank formation (not studied here). Modern land surfaces are shown in black.

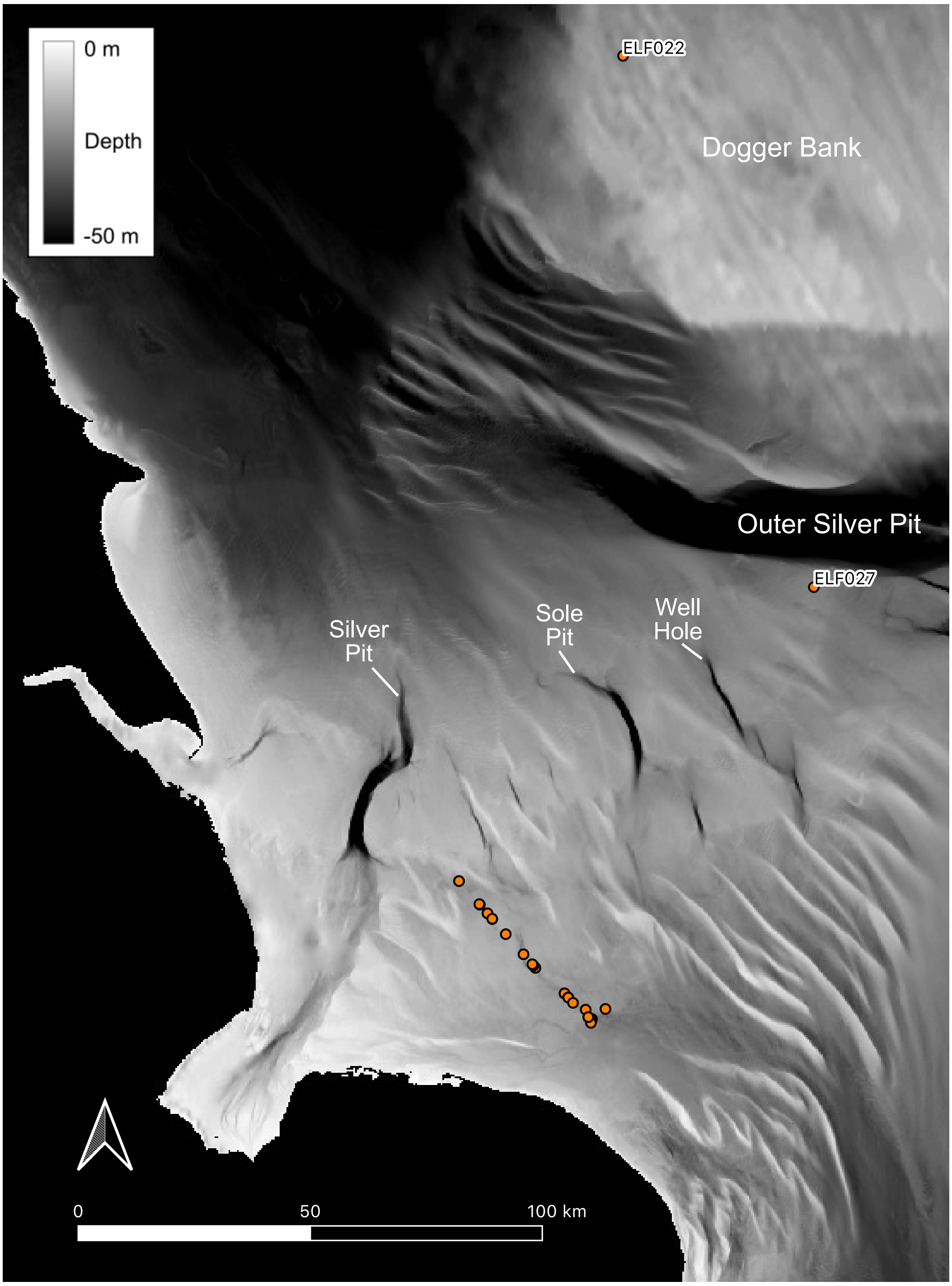


Figure 2.2: Bathymetry map of the North Sea showing the 27 sediment cores used in this project. Modern land surfaces are shown in black. The two cores outside of the palaeochannel transect, ELF022 and ELF027, are highlighted. See figure 2.3 for details of the palaeochannel.

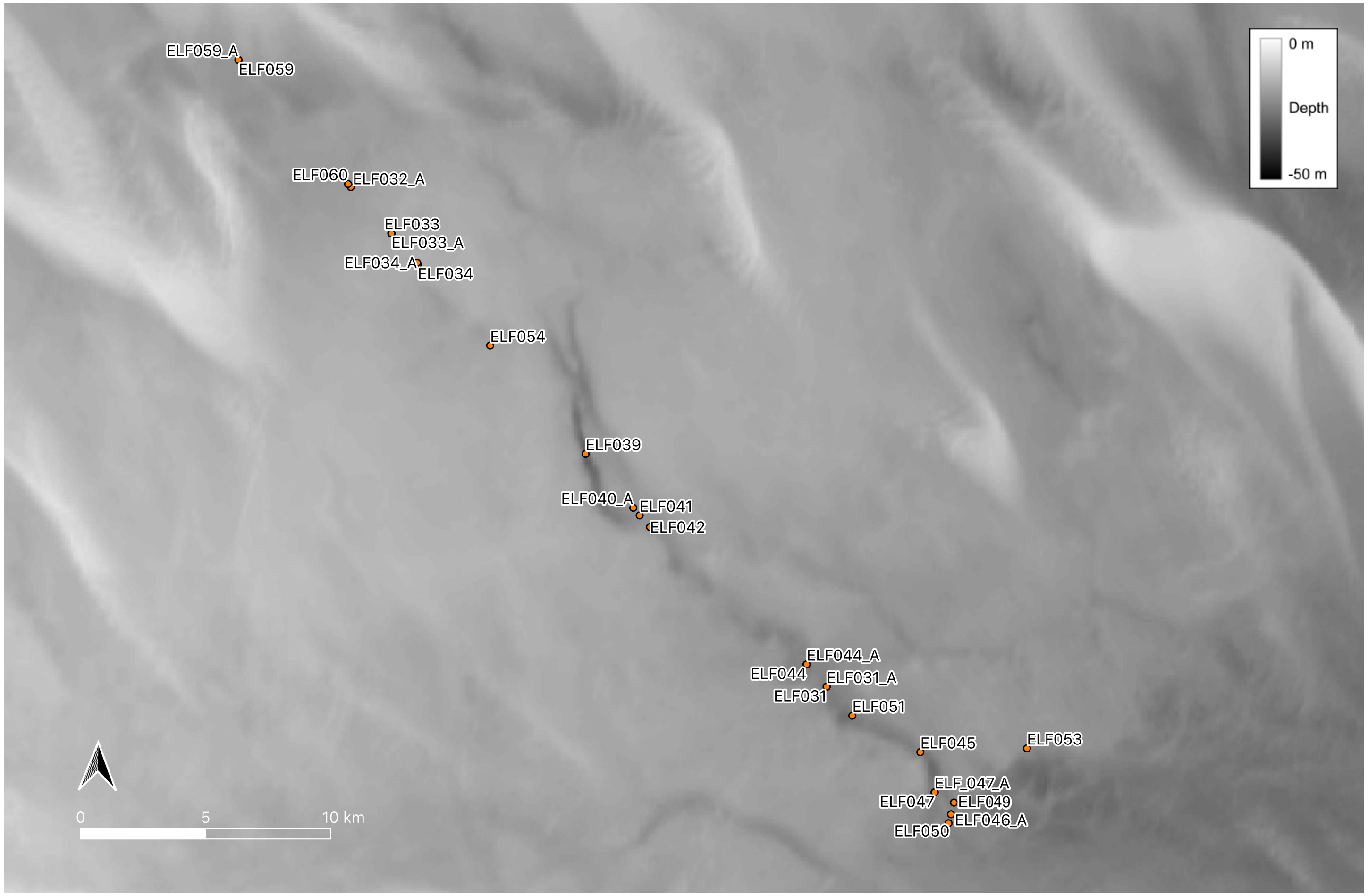


Figure 2.3: Bathymetry map of the palaeochannel transect containing all cores used in this project apart from ELF022 and ELF027 (see figure 2.2). The channel itself is visible from near core ELF039 to ELF047. Cores ELF049, ELF046A, ELF050, and ELF053 may extend beyond the mouth of the channel.

2.1.2 Sampling

The sediment cores were stored at 4°C before analysis. Cores were split and sampled at the University of Warwick ancient DNA facility following standard protocols (Gilbert *et al.* 2005). Strata were chosen for sampling by geoarchaeologists. Sampling took place inside a category two biosafety cabinet using sterile equipment. The cut surface of the core was discarded and the 1 cm of sediment next to the casing was avoided. Samples weighed 20g and were frozen within a few hours. 142 were taken for this project.

Seven additional samples were taken at a later date after the realisation that some cores may have contained deposits from the Storegga slide tsunami. These samples included ELF031A_58 (core ELF031A, depth 58 cm), ELF039_341, and ELF059A_355 from cores ELF021-60. As the cores had already been exposed to non-sterile conditions, the samples were taken in an archaeology facility using tools rinsed in bleach and 70% ethanol.

Core sampling was shared between Roselyn Ware, Rosie Everett, and myself. I performed all subsequent work for cores ELF021-60 independently.

2.1.3 DNA extraction

All steps up to and including sequencing were performed twice in replicate, resulting in two sets of data per sample. Samples were processed in batches of up to seven plus one negative control (reagents only). Negative controls were regularly quantified to check for excess DNA, but were also sequenced at the same depth as the samples, providing an estimate of the background contamination (e.g. the "kit-ome"; Leonard *et al.* 2007, Salter *et al.* 2014) with which to filter sample data later on.

The extraction protocol is based on that of Smith *et al.* (2015) and was previously published in Gaffney *et al.* (2020). 2 g (±0.05 g) of sediment was mixed with 5 ml CTAB buffer (2% w/v CTAB, 1% w/v PVP, 0.1 M Tris pH 8.0, 20 mM EDTA, 1.4 M NaCl) and incubated at 37°C with agitation for 7 days. The samples were centrifuged at 20,000 xg for 10 minutes. The supernatant was manually shaken with 4 ml chloroform:isoamyl alcohol (24:1) for 5 minutes. The aqueous phase was combined with 20 ml Buffer AW1 (Qiagen, cat#19081) and incubated at room temperature for 1 hour. This was then applied to EZ-10 silica-based spin columns (BioBasic, cat# SD5005) using a vacuum manifold.

The columns were then washed, first with 500 μl Buffer AW2 (Qiagen, cat #19082), and then 300 μl acetone to remove buffer residue. Both wash steps were followed by centrifugation at 6,000 xg for 1 minute. The columns were then removed from their collection tubes and air dried for 5 minutes. Finally, DNA was eluted in 65 or 75 μl Buffer EB (Qiagen, cat#19086, 10 mM Tris-Cl pH 8.5). The columns were incubated at 37°C for 10 minutes and centrifuged at 15,000 xg for 2 minutes. The eluted DNA was quantified using a high-sensitivity Qubit assay (Invitrogen, cat#Q32854; DNA yields listed in appendix B).

2.1.4 Library preparation and evaluation

Double-indexed libraries were produced using a protocol based on Meyer and Kircher (2010) and incorporating modifications from Kircher *et al.* (2012). The initial fragmentation step was omitted as ancient DNA is expected to already be <400 bp in length. This method was also previously published in Gaffney *et al.* (2020). Indexing primers are as in the Meyer and Kircher (2010) supplemental material (Indexing_Oligo_Sequences.doc), but with custom index barcodes designed following the three criteria in Meyer and Kircher (2010) Table 1.

1. **Blunt-end repair**: 20 μl of each DNA extraction in the batch (including the negative control) was combined with a master mix comprising 12.27 μl H_2O, 7 μl Buffer Tango (10X; VWR, cat#BY5), 0.53 μl dNTPs (25 mM each; Invitrogen, cat#10297018), 0.4 μl ATP (100 mM; Thermo Scientific, cat#R0441), 2 μl T4 polynucleotide kinase (10 U/μl; Thermo Scientific, cat#EK0032), and 0.8 μl T4 DNA polymerase (5 U/μl; Thermo Scientific, cat#EP0062) per sample. This produced a final reaction volume of 40 μl. The reactions were incubated in a thermal cycler for 15 minutes at 25°C and 5 minutes at 12°C.

2. **Spin column purification**: as per Kircher *et al.* (2012), purification was performed using spin columns instead of SPRI beads; I used the MinElute PCR purification kit (Qiagen, cat#28004). Reactions were combined with 200 μl Buffer PB, incubated at room temperature for 2 minutes, and transferred to MinElute columns. The columns were centrifuged at 6,000 xg for 1 minute and the flow-through discarded. 750 μl Buffer PE was added to wash, followed by two centrifugation steps at 6,000 xg for 1 minute, discarding the flow-through each time. The columns were then moved to clean 1.5 ml microfuge tubes and 20 μl Buffer EB was added. The columns were incubated at room temperature for 2 minutes and centrifuged at 15,000 xg for 2 minutes to elute the DNA.

3. **Adapter ligation:** the purified DNA extracts were combined with a master mix comprising 10.9 μl H_2O, 4 μl PEG-4000 (50%; supplied with ligase), 4 μl T4 DNA ligase buffer (10X; supplied with ligase), and 0.1 μl of P5+P7 adapter mix) as per Kircher *et al.* 2012) per sample. 1 μl of T4 DNA ligase (5 U/μl; Thermo Scientific, cat#EL0011) was then added separately to each reaction to make a final volume of 40 μl. Reactions were incubated at 22°C for 30 minutes.

4. **Spin column purification**: as in Step 2 above.

5. **Adapter fill-in**: the purified DNA extracts were combined with a master mix comprising 14.1 μl H_2O, 4 μl ThermoPol buffer (10X; supplied with *Bst* polymerase), 0.4 μl dNTPs (25 mM each; Invitrogen, cat#10297018), and 1.5 *Bst* polymerase (large fragment; 8 U/μl; NEB, cat#M0275S) per sample. This produced a reaction volume of 40 μl. Reactions were incubated at 37°C for 20 minutes, followed by 80°C for 20 minutes. This heat inactivation replaced further purification (Kircher *et al.* 2012).

6. **Indexing PCR**: 20 μl of each processed DNA template was combined with a master mix comprising 15.1 μl H_2O, 0.4 μl dNTPs (25 mM each; Invitrogen, cat#10297018), 10 μl SuperFi buffer (5X; supplied with Platinum SuperFi), 2 μl $MgSO_4$ (100 mM; NEB, cat#B1003), and 0.5 μl Platinum SuperFi polymerase (2 U/μl; Invitrogen, cat#12351010) per sample. 1 μl each of a forward and reverse indexing primer (10 μM) was added to each sample, with every sample in a sequencing run indexed with a unique pair of primers. This produced a reaction volume of 50 μl. Reactions were subject to 18 PCR cycles according to the following temperature profile:
    - Initial denaturation: 98°C for 30 seconds
    - Denaturation per cycle: 98°C for 10 seconds
    - Annealing per cycle: 60°C for 20 seconds
    - Elongation per cycle: 72°C for 20 seconds
    - Final extension: 72°C for 10 minutes

Libraries were visualised on a 2% agarose gel to check library formation. Primer and adapter dimers were reduced using SPRI bead purification following the reaction cleanup protocol in Rohland and Reich (2012). The library solution was combined with 45 μl SPRI bead solution (approximately 1 volume only to preserve shorter fragments) and allowed to settle on a magnetic rack for 5 minutes. The clear solution was removed and the beads were washed twice with 70% ethanol, incubating for 30 seconds each time. Remaining liquid was removed and the beads were air-dried for 5 minutes. The purified DNA was eluted in 20 μl TET buffer (1 mM EDTA pH 8.0, 10 mM Tris-Cl pH 8.0, 0.05% w/v Tween 20).

The libraries were then quantified using a broad-range Qubit assay (Invitrogen, cat#Q32853; DNA yields listed in appendix B) and a fragment size profile was produced using a 2100 Bioanalyzer with a high-sensitivity DNA chip (Agilent Technologies, cat#5067-4626). This allowed calculation of their molarities.

2.1.5 Sequencing

For each sequencing run, relevant libraries were normalised and pooled according to the NextSeq 500/550 Denature and Dilute Libraries Guide. First, 2 μl of each library was normalised to a given molarity (between 1 and 4 nM) by combining with an appropriate volume of molecular-grade water. Any libraries already below the target molarity (e.g. most negative controls) were not diluted further. 4 μl of each normalised library was then added to a single tube to make the raw pool. The molarity of the raw pool was calculated from its DNA concentration according to a broad-range Qubit assay (Invitrogen, cat#Q32853) and the mean fragment length of pooled samples. The raw pool was then diluted to one of the required molarities according to the NextSeq 500/550 Denature and Dilute Libraries Guide. This was the final pool taken forward for sequencing.

The final pools were sequenced on the Illumina NextSeq platform using a high-output, 150-cycle kit (75x75 paired end; cat#20024907). Kits were v2 for the initial sequencing runs but updated to v2.5 for the later deep sequencing.

The 149 samples were initially sequenced across three NextSeq runs. For the majority, the first DNA extractions were sequenced on the first run and the replicate extraction on the second. However, the negative control associated with the initial sequencing of samples ELF034_202 and ELF034_219 contained an unexpectedly high level of contamination, so the two samples were re-extracted and sequenced in replicate on the second run. Any negative effect on sequencing depth in the second run from increasing the number of libraries from 163 to 165 should have been negligible. The seven additional samples from potential Storegga deposits were sequenced in replicate on a third run. ELF059A_355 from the original samples was also suspected to be from a Storegga deposit, so was re-extracted and sequenced two more times on this third sequencing run. The third run only contained 20 libraries including negative controls, so these samples were sequenced to much greater depth.

Finally, 107 samples from cores ELF001-60 associated with interesting results from sedaDNA and/or other proxies were deep-sequenced in much greater depth. 39 were new samples taken in an archaeological facility as described above (subsection 2.1.2). Roselyn Ware sampled, extracted, and built libraries for this new material, while I performed the sequencing and data analysis. Analysis is ongoing at the time of writing, but initial results are included in Chapter 6 (Mesophilic taxa and human disturbance indicators) and Chapter 7 (Discussion).

2.2 Initial data analysis

2.2.1 Read processing and taxonomic assignment

Read counts per replicate at three key stages of initial data analysis are listed in appendix C. This analysis pipeline produced a large number of output files organised into the following directory tree:

```
analysis_output/
    MEGAN/
    blast/
    trimmed_fasta/
    trimmed_fastq/
```

First, raw FASTQ files were adapter-trimmed and collapsed in AdapterRemoval 2.2.2 (Schubert *et al.* 2016):

```
AdapterRemoval --file1 $filevar --file2 $filevar2 --basename
analysis_output/trimmed_fastq/$filevar --minlength 30 --trimns
--trimqualities --minquality 30 --collapse --gzip
```

`$filevar` represents the file name of the R1 FASTQ file, the base for all output file names, and `$filevar2` the R2. Sequences shorter than 30 bp were discarded, trailing Ns or low-quality bases were trimmed off, and the R1 and R2 sequencing files were merged where possible.

Although they passed the quality filters, the .truncated.gz output files were discarded because the additional trimming is likely to have removed any age-associated damage signal, which would cause problems during authentication (see subsection 2.3.1). Therefore, only the .collapsed.gz FASTQs were taken forward and converted to FASTAs:

```
zcat analysis_output/trimmed_fastq/$filevar.collapsed.gz |
awk 'NR%4 !=0' | awk 'NR%3 !=0' | sed 's/@/>/g' >
analysis_output/trimmed_fasta/$filevar.collapsed.wd.fasta
```

These .wd.fasta files are "with duplicates"; they contain copies of the same sequences which are assumed to be PCR duplicates. The next stage removed duplicates using fastx_collapser from the FASTX-Toolkit 0.0.13 (Gordon and Hannon 2010):

```
fastx_collapser -i
analysis_output/trimmed_fasta/$filevar.collapsed.wd.fasta -o
analysis_output/trimmed_fasta/$filevar.collapsed.fasta
```

This produced a quality-controlled base FASTA for each sample. The next step was to assign the sequences to broad taxonomic groups. An initial BLAST was performed against the full nucleotide GenBank database (downloaded on 05-09-2019) using blastn 2.6.0 (Zhang *et al.* 2000):

```
blastn -db BLAST_nt_2019-09-05/nt -num_threads $threads -query
analysis_output/trimmed_fasta/$filevar.collapsed.fasta -out
analysis_output/blast/$filevar.collapsed.txt  -num_alignments 10
-outfmt "6 std staxids";
```

A BLAST search looks through the reference database and outputs up to a given number of reference sequences that match the query sequence sufficiently closely (referred to as hits). Output format 6 is tabular, reducing file size, and taxonomic IDs of hits (`staxids`) were required for the next step. In format 6, `-num_alignments` states the maximum number of hits collected per query. 10 was sufficient for this stage.

This initial BLAST output and the base FASTA were then fed into the command line version of MEGAN5 (Huson *et al.* 2016), a taxonomic assignment program, with default settings. This produced a base RMA file of assigned reads for each sample:

```
echo "load taxGIFile='gi_taxid-March2015X.bin';
set loadAllReadsIntoMemory=false ;
import blastFile='analysis_output/blast/$filevar.collapsed.txt'
fastaFile='analysis_output/trimmed_fasta/$filevar.collapsed.fasta'
meganFile='analysis_output/MEGAN/$filevar.collapsed.rma'
minScore=30.0 minSupport=5 minComplexity=0.0
useMinimalCoverageHeuristic=false useSeed=false useCOG=false
useKegg=false paired=false useIdentityFilter=false
textStoragePolicy=Embed blastFormat=BlastTAB
mapping='Taxonomy:BUILT_IN=true,Taxonomy:GI_MAP=true';
quit;" >> command.txt

xvfb-run --auto-servernum --server-num=1 MEGAN -g -S -L
MEGAN5-academic-license.txt -c command.txt

rm command.txt
```

The two main taxonomic groups of interest were Viridiplantae and Metazoa (excluding Primates). Primates were excluded because non-human primates are not expected in Holocene northern Europe, and the inevitably large amount of contaminant human DNA would have significantly increased processing time and file sizes. Again using MEGAN5, the two groups were extracted from the base RMAs as two separate RMAs and were analysed in parallel from this point. Viridiplantae and Metazoa RMAs were extracted as follows, with taxon specified by the $taxa variable below:

```
echo "open file='analysis_output/MEGAN/$filevar.collapsed.rma';
extract what=document file='
analysis_output/MEGAN/$filevar.$taxa.collapsed.rma' names=$taxa
allBelow=true;
open file='analysis_output/MEGAN/$filevar.$taxa.collapsed.rma';
export what= reads file='
analysis_output/trimmed_fasta/$filevar.$taxa.collapsed.fasta';
quit;" >> command.txt

xvfb-run --auto-servernum --server-num=1 MEGAN -g -S -L
MEGAN5-academic-license.txt -c command.txt

rm command.txt
```

The Viridiplantae FASTA was complete. However, the Metazoa RMA was opened manually in the MEGAN6 GUI and Primates excluded to export the Metazoa (no Primates) FASTA.

Once extracted, each new FASTA was then BLASTed more thoroughly in preparation for more stringent taxonomic assignment, using `-max_target_seqs` 500 instead of `-num_alignments` 10 to give up to approximately 500 hits per read:

```
blastn -db BLAST_nt_2019-09-05/nt -num_threads $threads -query
analysis_output/trimmed_fasta/$filevar.$taxon.collapsed.fasta -out
```

```
analysis_output/blast/$filevar.$taxon.collapsed.txt -max_target_seqs
500 -outfmt "6 std staxids";
```

Assigning metagenomic reads to taxa is complicated by the uneven representation of taxa in reference databases. Phylogenetic Intersection Analysis (PIA) is a program developed by the Allaby group to address the main issues (Chapter 3). PIA 5.3 was used to assign reads in the Viridiplantae and Metazoa (excluding Primates) FASTAs using the new BLAST results:

```
perl PIA.pl -f $filevar.$taxa.collapsed.fasta -b
../blast/$filevar.$taxa.collapsed.txt -t $threads
```

This concluded the initial analysis pipeline. However, MEGAN5 was also used to assign the Viridiplantae FASTAs via their thorough BLAST output for comparison against PIA (Chapter 4: Taxonomic results). The settings were otherwise identical to the initial base assignment above. A taxonomic summary of each sample was then manually extracted from each new Viridiplantae RMA in the MEGAN6 GUI as a tab-separated list in the format "taxonID_to_count". This was comparable to PIA output.

2.2.2 Filtering by negative controls

Once reads were assigned, taxa in samples could be filtered by taxa in negative controls. Negative controls for each assignment method were merged by sequencing run. This gave a total control file against which all sample files from that sequencing run could be filtered. On the replicates sequencing run, the control Blank_B9 was mixed with ELF053_275 due to an indexing error, but as the amount of post-PIA data was very small, the combined file was simply treated as Blank_B9. More of a problem was control Blank_B21R, which had an unexpectedly large amount of data, including many taxa common in sample files. This contrasted with the apparently empty Qubit and Bioanalyzer results, suggesting another indexing error. That control and its related sample files (replicates for ELF054_291, 315, 330, and 356; ELF046A_270; and ELF044_90 and 137) were discarded.

The remaining sample files were then filtered according to their total control file using a custom Perl script. This script excludes a taxon from the sample file if the number of reads assigned to the taxon in the control is $\geq$5% the number in the sample. The strictest approach is to exclude all taxa from the sample that are present in the negative control, but this more nuanced method allows the taxa to pass if the signal in the sample is much stronger than in the blank.

2.2.3 Merging replicates

All samples but those associated with Blank_B21R (subsection 2.2.2) had data in replicate. Replicates from the initial ELF021-60 sequencing data were initially analysed separately, but were then compared to assess whether they should be merged (merging was irrelevant for the few analyses involving cores ELF001-20 or deep-sequencing data in this thesis). This involved running a modified form of the stratification analysis (subsection 2.3.2) on the Embryophyta (land plants; Viridiplantae excluding algae) data, which had returned far more reads than Metazoa (excluding Primates).

The stratification analysis compares pairs of input files taxon by taxon. There were 137 sample pairs in the initial ELF021-60 sequencing data. 57% had at least one taxon significantly different ($p < 0.05$) between replicates, which arguably indicates different taxonomic profiles. The mean percentage of significantly different taxa per sample was 33%, although this varied greatly (standard deviation = 28). The percentage of different taxa was somewhat correlated with overall read count (Spearman's coefficient = 0.6114357): the more data, the more different replicates appeared. While we may expect pairs of replicates with disparate read counts to appear more different because of differences in sample size, there was no correlation between read counts of replicates within a sample. Pairs of replicates were just as likely to have similar or dissimilar amounts of data. Instead, the increasingly different taxa with increasing read count may be explained by limited sampling overall. Samples with more data contain a greater number of rare taxa, which are less likely to appear in both replicates, making the replicates appear less similar.

I therefore decided to merge replicates: little information would be lost by merging similar replicates, and merging dissimilar replicates would result in a more representative sample. This resulted in a single post-PIA list of assigned reads for each sample, filtered by negative controls, for Viridiplantae and Metazoa (excluding Primates).

2.2.4 Genome size adjustment

Palaeoenvironmental reconstruction would be most straightforward if we could estimate biomasses of taxa. Biomass should influence read count: the greater the biomass of a taxon, the more DNA available for deposition in the sediment. Because this study used non-selective shotgun sequencing, we would expect more sedaDNA to result in more reads, assuming that taphonomy and DNA recovery within a sample is constant. However, read count is also affected by the database representation and genome sizes of organisms. The former is largely out of our control, but we can attempt to correct for genome size, bringing read count a step closer to biomass.

Building on methods used by Roselyn Ware in Gaffney *et al.* (2020), I collected C-value (pg DNA per haploid genome) estimates for Viridiplantae and Metazoa from the Kew (Leitch *et al.* 2019) and Gregory lab (Gregory 2018) databases respectively. The mean C-value for each taxon in the data was calculated from relevant measurements in the databases. These are listed in appendices D (Viridiplantae) and E (Metazoa).

Viridiplantae C-values were restricted to the single curated "prime" measurement for each taxon where possible. There was no such option for Metazoa, so those estimates are more likely to have been influenced by unreliable measurements. Furthermore, approximately 25% and 20% of taxa in the Viridiplantae and Metazoa databases respectively could not be automatically identified using the NCBI taxonomy database, and manual identification was not feasible on the timescale of this project. As the taxonomies used in the genome size databases do not contain every rank, and sometimes use outdated classification, mean C-values for some taxa may have been calculated without relevant measurements or with those from taxa now considered unrelated.

In an attempt to quantify some reliability, appendices D and E also list the coefficients of variation (CV; standard deviation divided by mean) alongside each mean C-value, where $CV \leq 1$ suggests a relatively precise estimate and $CV \geq 1$ imprecise. Many of the most imprecise estimates are for genome sizes of higher taxa, which is to be expected. Fortunately, these taxa are rarely useful for ecological reconstruction so are usually ignored. Taxa that are used for reconstruction despite imprecise genome size estimates will be noted individually.

To correct for genome size, read counts were divided by estimated C-value, making count proportional to the amount of DNA per cell, and therefore proportional to biomass. These genome-size-adjusted counts are termed biogenomic mass (Gaffney *et al.* 2020).

2.2.5 Adding ecological information

Reconstructing an ecosystem from a list of taxa requires ecological information about the taxa involved. I maintain a database of all sedaDNA taxa in the Allaby group containing their taxonomic information, estimated genome size, European and Great British native status, and what we term "ecological category" (appendices D and E list taxa relevant to this thesis). There are twelve ecological categories for Viridiplantae (table 2.1) and five for Metazoa (table 2.2). Taxa are assigned to the first relevant category down the list. For example, a freshwater fern would be assigned to "freshwater aquatics", and a terrestrial fern to "ferns". The order of categories loosely follows pollen diagrams produced by colleagues in Europe's Lost Frontiers (Gaffney *et al.* 2020).

Table 2.1: Ecological categories for taxa in Viridiplantae.

Ecological category		Description
Habitat	Freshwater aquatics	
	Salt/brackish aquatics	
	Mixed aquatics	Fresh and salt/brackish water
	Halophytes	Salt-tolerant, so often coastal
	Xerophytes	Drought-tolerant
Form	Trees/shrubs	Woody seed plants
	Herbs	Non-woody seed plants excluding Poales
	Trees/shrubs and herbs	
	Grasses and relatives	Poales and child taxa
	Ferns	
	Bryophytes	Liverworts, mosses, and hornworts
	Mixed	Containing taxa from multiple form categories

Table 2.2: Ecological categories for taxa in Metazoa.

Ecological category	Description
Freshwater aquatics	
Salt/brackish aquatics	
Mixed aquatics	Fresh and salt/brackish water
Terrestrial	
Mixed	Terrestrial and aquatic

Genome-size-adjusted samples were collated by core alongside taxonomic and ecological information for all relevant taxa. This produced the main results tables.

2.2.6 Separating irrelevant and problematic taxa

The main results tables were further divided by taxon. First, taxa outside the scope of this study were removed. Viridiplantae was restricted to Embryophyta (land plants), therefore excluding algae. Although reads assigned to Primates by MEGAN at an early stage of analysis had been removed, a small number of additional Primate reads were later identified by PIA. These were also excluded. I refer to the remaining taxa as the "ingroup". The total biogenomic mass for each sample refers to the ingroup.

The final filtering stage separated taxa not native to Europe. The sediment samples originate from the southern North Sea and are expected to be on the order of several thousand years old, so we would expect the taxa to be comparable to those in present-day Europe. Negative control filtering should have removed most reads assigned to common contaminants. Any remaining non-European taxa are assumed to be false positives. These are examined separately in Chapter 5 (Authentication results).

However, biogenomic masses from non-European taxa still contribute to the totals for each sample because some analyses use cumulative counting. In this system, biogenomic masses for a taxon also count towards its parent taxa. For example, if a sample had a biogenomic mass of five for *Betula* (birches) and ten for Betulaceae, then under cumulative counting, the total for *Betula* would still be five, but for Betulaceae it would be fifteen. This makes parent taxa more comparable, as those sequences assigned to *Betula* are also assigned to Betulaceae, if indirectly. If *Betula* was excluded as a non-European native but Betulaceae remained, the total biogenomic mass for Betulaceae would still be fifteen. Because even excluded ingroup taxa can potentially be counted via their parents, the grand total for each sample also counts all ingroup taxa. Otherwise, it would not be possible to convert read counts into proportions. Cumulative counting is used for the Pianka similarity index and the stratification analysis.

At the end of the main analysis pipeline, there are post-PIA, negative control filtered, biogenomic masses for each sample, arranged by core, for Embryophyta and Metazoa (excluding Primates). These taxonomic results are discussed in Chapter 4.

2.3 Authentication

A significant part of this project involved improving sequence assignment. With PIA, we can be much more confident that a read originated from the taxon it has been assigned to. However, correctly identifying a read is a separate issue to authenticating its age. Filtering by negative controls should exclude frequent modern contaminants, but it is not foolproof, and we must also consider the possibility of sedimentary DNA moving between strata (leaching). Exploration of these two elements, DNA age and vertical stability, are detailed in this section.

2.3.1 Age authentication using DNA damage

DNA sustains damage after cell death (Dabney *et al.* 2013a, Kistler *et al.* 2017). The two key patterns are fragmentation and C-to-T (or complementary G-to-A) substitutions. These substitutions are more common towards the ends of molecules because any bases left in a single-stranded overhang are more exposed to attack. Therefore, ancient sequences are characteristically short with increased C-to-T substitutions at the 5' end and G-to-A at the 3'.

The leading tool for authentication via damage patterns, mapDamage2.0 (Jónsson *et al.* 2013), analyses the size distribution of reads and identifies C-to-T and G-to-A substitutions through alignment with a reference genome. Input reads must be from organisms closely related enough to the reference for a meaningful alignment, and there must be enough reads to distinguish between base mismatches caused by damage and by other factors. This limits the application of mapDamage2.0 in metagenomic studies like this, with a large number of taxa mostly represented by few reads each.

Therefore, Rosie Everett, Logan Kistler, and myself developed the MetaDamage analysis tool that also calculates C-to-T and G-to-A mismatches between input sequences and references, but provides each with an individual reference from GenBank instead of using a single global reference genome. This allows MetaDamage to compare diverse taxa in a sample. A minimum number of sequences is still recommended, but as the sequences can be from any taxon, it is much more applicable to metagenomics. MetaDamage was presented at the 2021 Society of Antiquaries of London meeting (Allaby *et al.* 2021) and is in preparation for publication.

In brief, MetaDamage runs an input FASTA through a simple BLAST search to identify a corresponding list of reference sequences. Each reference is then re-aligned to its query. Mismatches between the reference and query of any kind are counted along the first 25 bases from each end. The counts for each sequence are averaged to produce an overall rate for every type of mismatch on both ends. These are plotted in a similar way to the fragment misincorporation plots of mapDamage2.0, allowing comparison between the two age-associated mismatch types (C-to-T on the 5' end and G-to-A on the 3') and the others (a baseline). Excess age-associated mismatches at the very ends of the molecules, particularly position zero, indicate age-associated damage, and are a sign that the input sequences are on average ancient.

MetaDamage also provides a measure of variance by calculating a 95% confidence interval for the two most important output values: the position-zero C-to-T mismatch rate on the 5' end and the G-to-A rate on the 3'. Chapter 5 (Authentication results) shows how the confidence interval reveals a lack of reliability in results from samples with too few reads.

MetaDamage 2.0 was used to analyse the post-PIA, control-filtered Embryophyta and Metazoa (excluding Primates) data per sample. This generated a damage profile for each of the taxonomic profiles of Chapter 4. These results form section 5.1 of Chapter 5 (Authentication results). MetaDamage was also used to authenticate reads from specific taxa pooled from multiple samples, especially unexpected taxa that may be particularly likely to represent modern contamination. These results are the basis of Chapter 6 (Mesophilic taxa and human disturbance indicators).

2.3.2 Stratification analysis

MetaDamage can flag samples that appear to have a significant modern component. However, as most damage is expected to occur relatively soon after the death of an organism (Kistler *et al.* 2017), it is unlikely to distinguish between ancient DNA of different ages. The vertical movement of DNA between strata, which would result in DNA profiles of mixed ages, must therefore be considered separately. Vertical DNA movement would dissociate DNA from its stratigraphic context and make interpretation challenging. Previous attempts to exclude DNA movement have relied on the presence or absence of key taxa, such as the confinement of moa DNA to pre-European strata in New Zealand sediments studied by Haile *et al.* (2007). Haile *et al.* were subsequently able to detect downwards movement of DNA because some pre-European strata also contained DNA from sheep. This method relies on prior knowledge of temporal ranges for key taxa, which is not possible for the relatively unknown environments of this study. Instead, Robin Allaby, Roselyn Ware, and myself developed a program that searches for evidence of stratification based solely on patterns in the DNA profiles, first published in Gaffney *et al.* (2020). The taphonomy of DNA in sediment is still not well understood, but we assume that the movement observed in studies such as Haile *et al.* (2007) involves free DNA diffusing or being carried (such as by water) through undisturbed sediment; displacement through processes such as bioturbation is another matter. This vertical DNA movement would lead to the DNA profiles of adjacent samples in a core becoming more similar, either upwards or downwards or in both directions. Therefore, a lack of DNA movement can be inferred from significantly different DNA profiles in adjacent samples.

For each sample, the stratification analysis takes the read count for each taxon as a proportion of the total reads. This attempts to control for data yield. The proportion is treated as the probability of observing that taxon following a binomial distribution. Each proportion is used to build a beta distribution, where parameters α and β are reads assigned to that taxon and to other taxa respectively. The analysis then compares pairs of adjacent samples up the core. It asks whether the read count for that taxon could have been drawn from the same distribution in both samples: a p-value is calculated from the overlap between the two beta distributions. Similar DNA profiles should contain approximately the same proportions of taxa.

The p-value shows the statistical significance of any change. To show direction and magnitude, the analysis also calculates an index of change for each comparison. A negative index shows a proportional decrease for that taxon and a positive index shows an increase. It is calculated as the difference in proportion across the samples as a percentage of the smaller proportion:

$$Change\ Index = \frac{(b-a)}{a} * 100\ where\ b > a$$

$$Change\ Index = \frac{(a-b)}{b} * 100\ where\ a > b$$

Where *b* is the proportion in the lower sample ("below") and *a* is the proportion in the upper sample ("above").

Finally, the completeness of signal is calculated for each core. A statistically significant change in a taxon between samples can demonstrate stratification: if the taxon is present above but absent or significantly reduced below, it suggests a lack of downwards DNA movement, and vice versa for upwards. Therefore, each pair of adjacent samples can have two possible instances of stratification: upwards and downwards. These are determined by the presence of at least one statistically significant ($p<0.01$) increase and decrease respectively. The percentage of instances that show stratification, of the 2*(number of samples), is the completeness of signal for the core.

I performed the stratification analysis on negative control filtered, European Embryophyta read counts for each core, counted cumulatively. Under cumulative counting, very high taxa (e.g. kingdom, class) usually made samples appear disproportionately similar because they included almost all reads. The few significant differences among these higher taxa occurred in Liliopsida and Mesangiospermae between very high-yield samples: those that returned very large read counts. These changes could be explained by the artefact of assignment where the proportion of reads assigned to higher taxa in *Zostera*-dominated samples increases with data yield (Chapter 2: Taxonomic results, subsection 4.30.2). In short, very high taxa did not give valid signals. The data was therefore re-analysed with the following taxa excluded:

- Embryophyta
- Tracheophyta
- Euphyllophyta
- Spermatophyta
- Magnoliopsida
- Mesangiospermae
- eudicotyledons
- Gunneridae
- Pentapetalae
- Liliopsida

Also excluded were taxon comparisons involving fewer than 14 reads. 14 is the minimum required for a p-value of <0.01, below which there is a lack of statistical power. The minimum read count also limits excessive appearance/disappearance effects of rare taxa. The stratification analysis weights appearances or disappearances strongly, assigning indices of change of "Inf" (infinity) and "-Inf" (negative infinity) respectively. This could result in exaggerated differences between profiles if, for example, a rare taxon appeared stochastically along a core. The minimum read limit aims to exclude these consistently rare taxa.

The p-value and index of change can indicate a change in DNA signal for a given taxon across samples. However, a lack of change does not necessarily suggest that DNA movement occurred: similar DNA profiles may also arise from a similar ecology. Therefore, the stratification analysis is used to *exclude* DNA movement where possible and flag other samples for further investigation, particularly using different environmental proxies. The results are discussed in Chapter 5 (Authentication results), section 5.2.

It is also important to remember that, although a DNA profile is influenced by biomass from the ancient environment, it is not a simple proxy. Using proportions instead of absolute read counts attempts to account for library performance, so taxa in samples that only differ in data yield (number of reads returned) should appear similar. However, unlike in Chapter 4, there is no attempt to account for genome size, and a significant proportion of taxa are likely to be excluded due to a lack of data. The stratification analysis evaluates DNA profiles but is not well suited to inferring changes in biomass of particular taxa. It will not be used to infer ecological change, for example.

Finally, I adapted the stratification analysis to compare not adjacent samples, but pairs of replicates for the same samples (subsection 2.2.3). This informed whether replicates would be merged before further analysis.

The next chapter looks back at the read assignment program PIA to explain the algorithm and its development tests in detail.

Chapter 3

PIA: more accurate taxonomic assignment of metagenomic data demonstrated on sedaDNA from the North Sea

Excluding very minor changes, this chapter was previously published as Cribdon *et al.* (2020). A PDF is included in appendix A, and the paper and supplementary material can also be accessed free of charge at https://doi.org/10.3389/fevo.2020.00084. Robin Allaby, Oliver Smith, Roselyn Ware, and myself wrote and designed PIA. I performed benchmarking and accuracy testing with some input from Roselyn Ware. I was the primary author of the manuscript with review and editing by Robin Allaby and Roselyn Ware. Vincent Gaffney was the Principal Investigator of Europe's Lost Frontiers, the project through which the sedaDNA dataset was obtained.

Abstract

Assigning metagenomic reads to taxa presents significant challenges. Existing approaches address some issues, but are mostly limited to metabarcoding or optimised for microbial data. We present PIA (Phylogenetic Intersection Analysis): a taxonomic binner that works from standard BLAST output while mitigating key effects of incomplete databases. Benchmarking against MEGAN using sedaDNA suggests that, while PIA is less sensitive, it can be more accurate. We use known sequences to estimate the accuracy of PIA at up to 96% when the real organism is not represented in the database. For ancient DNA, where taxa of interest are frequently over-represented domesticates or absent, poorly-known organisms, more accurate assignment is critical, even at the expense of sensitivity. PIA offers an approach to objectively filter out false positive hits without the need to manually remove taxa and so make presuppositions about past environments and their palaeoecologies.

3.1 Introduction

Next-generation sequencing allows detailed metagenomic analysis of a wide range of ancient samples. Studies have attempted to recreate biological communities from material including coprolites (Appelt *et al.* 2014, Bon *et al.* 2012), dental calculus (Warinner *et al.* 2015, Weyrich *et al.* 2017), ice cores (Willerslev *et al.* 2007), sediment (Birks and Birks 2016, Smith *et al.* 2015), stalagmites (Stahlschmidt *et al.* 2019), rodent middens (Kuch *et al.* 2002), and mollusc shells (Der Sarkissian *et al.* 2016). Our understanding of contamination and best laboratory practice has made good progress (Gilbert *et al.* 2005, Shapiro *et al.* 2019) and methods for authenticating ancient DNA sequences are developing (Key *et al.* 2017, Renaud *et al.* 2019). However, identifying ancient metagenomic sequences is still a challenge, particularly for shotgun data.

Shotgun sequencing has three key advantages over metabarcoding for ancient metagenomics. First, it can capture information from anywhere in the genome, greatly increasing sensitivity. Every DNA molecule extracted from a sample has the potential to be identified, provided that reference databases are adequate. Second, read count and genome size could be used to calculate biogenomic mass: a proxy of biomass (Gaffney *et al.* 2020). Third, metabarcoding is far less likely to record DNA damage signals. Damage accumulates in DNA over time (Kistler *et al.* 2017), so is important for authentication of ancient reads, and occurs most rapidly on the single-stranded overhangs at the ends of molecules. A characteristic damage signal is C-to-T deamination; changes to the base sequence make it less likely that metabarcoding primers will anneal, so damaged molecules are less likely to be sequenced. Furthermore, primer regions are typically removed during analysis, so even if the very ends of molecules are amplified, they will not be considered. Shotgun sequencing can potentially sequence whole molecules, especially when fragments are short, as is the case for ancient DNA. This preserves any damage signal intact. Overall, shotgun data has the potential to supply highly sensitive and informative metagenomic data.

However, because sequences can come from anywhere in the genome, accurately assigning shotgun reads to taxa requires a much larger reference database than for metabarcoding. The GenBank database is the most comprehensive (Benson *et al.* 2016), but even this is highly incomplete. Only a tiny fraction of organisms have had their full genomes sequenced and most are not represented at all. Reads from unrepresented organisms may go unassigned. Worse, the uneven representation of taxa that are in a database can create two additional problems that may lead to incorrect assignments.

The first problem is the over-representation of some taxa. This was recently identified as an issue for BLAST (Zhang *et al.* 2000), the “gold standard” of taxonomic binning (Herbig *et al.* 2016), by Shah *et al.* (2018). When BLAST searches against a database, it starts at the top and returns the first *n* hits that pass a quality filter, not the best *n* hits. If an over-represented taxon is a reasonable match, BLAST could return *n* hits and finish before it has a chance to identify closer but less represented taxa further down the database. Better matches may be missing from the list of hits. Even if BLAST does check the whole database, the list of hits may be disproportionately full of over-represented taxa. Taxonomic assignment methods that consider this list may then assign with too much weight to these taxa.

The second problem with an uneven database is “oasis” taxa in “sparse” areas. Consider a sparsely-populated area of the database with just one or a few taxa represented, not including the real taxon (figure 3.1, panel B). A specific sequence is unlikely to hit anything and will probably be left unassigned. But a conservative sequence may hit that one or few taxa, not necessarily because they are a good match, but because there is nothing else closer. The list of BLAST hits for that read will not be empty, but will have very low diversity. This can give the illusion of a confident match. Taxonomic binners that use a phylogenetic intersection or “lowest common ancestor” approach, robust to conservative sequences, can produce false positives because of oasis taxa.

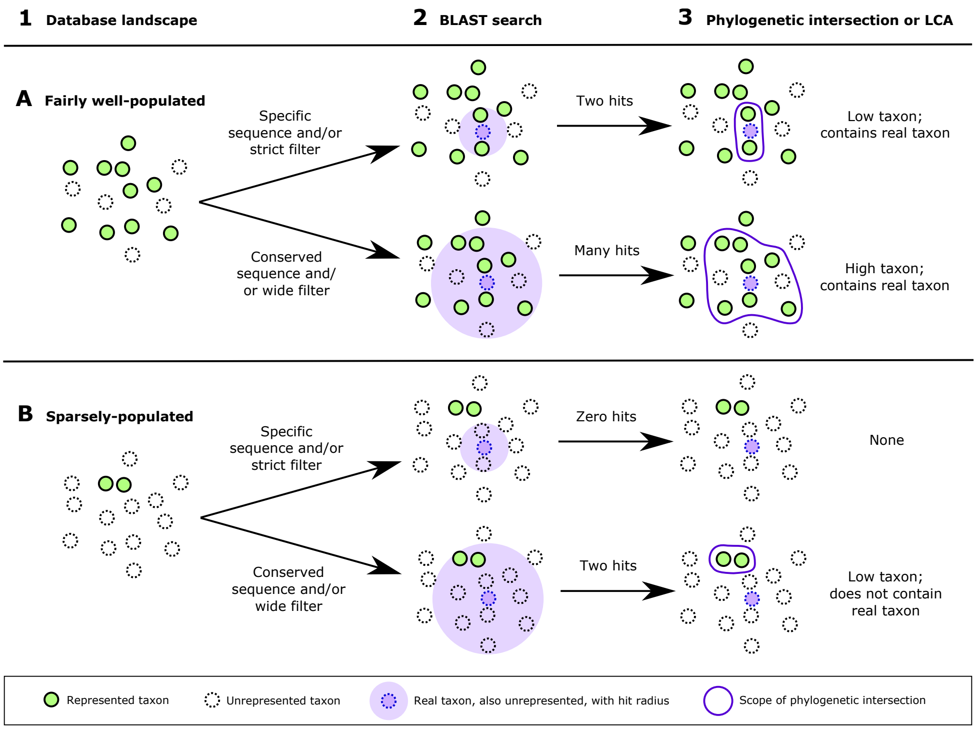


Figure 3.1: Database landscape diagrams of an (A) well-populated and (B) sparsely-populated region. Column (1) is the state of the database region. (2) shows a BLAST search for a sequence from an unrepresented taxon; for a reference sequence to be a hit, it must be inside the hit radius. Column (3) is what phylogenetic intersection or LCA (“lowest common ancestor”) those BLAST hits would produce: what the read would be assigned to. In a well-populated region (A), the more conservative the sequence (or the wider the quality filter), the higher the taxon the read would be assigned to. The intersection should contain diverse relatives of the real taxon, so is likely to contain the real taxon. In a sparsely-populated region (B), a specific sequence (or strict filter) is unlikely to be assigned. But a more conservative sequence (or wider filter) may hit the one or few represented taxa in the region: oases. These are probably not close relatives of the real taxon. However, the resulting low-diversity list of hits gives the intersection an illusion of confidence.

BLAST and BLAST-like algorithms have a minimum quality filter that affects how similar a reference sequence must be to count as a hit and how much empty space there must be around a read for it to go unassigned (figure 3.1, “hit radius”). But as with many aspects of taxonomic assignment, this filter has a trade-off between accuracy and sensitivity. A very strict filter would increase the resistance of reads to not-very-similar oases, but make them less attracted to more similar sequences that could be informative. This is especially an issue for aDNA, where even a read from an organism that is in the database may not share an identical sequence because of DNA damage or mutations over time. The minimum quality filter cannot protect from oasis taxa alone.

One of the main arguments in favour of metabarcoding is its use of confined, curated databases that aim to be functionally complete for the study taxon in the study area, such as the Arctic flora database in Sønstebø *et al.* (2010). Uneven representation is limited if all taxa are represented to some degree. It is currently realistic to sequence a barcode region of several hundred species for a study, as in Sønstebø *et al.* But because shotgun sequencing can access the whole genome, a complete shotgun database must have the full genome of all organisms, which will not happen in the foreseeable future. Metabarcoding databases are typically far more "complete" in that more of the study taxa are represented. However, this still assumes that an environment can even be well-studied enough for a complete list of taxa. This is debatable, especially for ancient ecosystems. Despite metabarcoding databases being easier to fill, arguably neither can ever be truly complete. Metabarcoding does not fully address uneven representation in databases. Both metabarcoding and shotgun approaches would benefit from an alternative solution.

A method that accepts shotgun data while also improving the database is SPARSE (Zhou *et al.* 2018). It rebuilds a given database as hierarchical clusters of similar sequences. If a taxon is represented by several very similar genomes, these genomes will be combined into a single cluster. The final SPARSE database has every present taxon represented by one genome, addressing the problem of over-represented taxa. However, SPARSE is designed for microbial data in relatively well-studied systems, where the database is both relatively well-populated and small enough to be rebuilt on a typical lab server. It does not address the problem of oasis taxa in sparse areas, nor would it be easily applicable to studies of organisms with larger genomes.

A popular standard tool for metagenomic studies not limited to microbes is MEGAN (Huson *et al.* 2007, 2016). This analyses output from various reference-matching programs, including BLAST. Its sister program, MALT (Herbig *et al.* 2016), aims to generate comparable output to BLAST at greatly increased speed before assigning taxonomy in the same way as MEGAN. This shared method is the LCA (Lowest Common Ancestor) algorithm (Huson *et al.* 2016). The default naive LCA is best suited to taxonomic binning. For each read, hits are first quality-filtered against multiple criteria. Good hits are assumed to belong not to the single organism they were sequenced from, but the "lowest common ancestor" (ancestral node) of all associated taxa. Being associated with multiple taxa suggests that the hit sequence is conservative, so should be assigned to a higher taxon. The more conserved the sequence, the more diverse the associated taxa, so the higher the taxon to which the hit is assigned. Following the same logic, the read is then assigned to the lowest common ancestor of its list of processed hits.

The LCA is robust to overrepresented taxa in the list of hits. The lowest common ancestor is calculated on presence/absence, not number of occurrences. However, accurate assignment still depends on the list containing accurate hits to begin with, which overrepresented taxa can prevent (Shah *et al.* 2018).

The LCA also addresses unrepresented taxa: even if the real taxon is not in the database, the list of hits should include relatives, so the read should be assigned to an "ancestor" that encompasses the real taxon. The more sparse the database, the more diverse the list of hits, so the higher the taxon the read is assigned to. In very sparse regions, this means that reads are likely to be under- or unassigned but not incorrectly over-assigned (Huson *et al.* 2007). However, we argue that the LCA approach may incorrectly assign these reads if they are influenced by oasis taxa. If, for instance, a sparse region were occupied by clumps of taxa rather than an even spread of relatives around the unrepresented taxon (figure 3.1, panel B), the list of hits may be dominated by one of those taxon clumps, resulting in a relatively specific "ancestor" close to the oasis but not necessarily the real taxon.

MEGAN does have a further check against false positives: the min-support filter (Huson *et al.* 2007, Huson 2019). Once all reads have been assigned, resulting taxa are only reported if they contain a minimum number of reads. If a read was assigned to a taxon that does not meet this threshold, it is pushed up the taxonomy until it reaches a taxon that does. This excludes very rare taxa, which Huson *et al.* argue are more likely to be false positives. However, we argue that oasis taxa could escape this check. Being the only represented taxon in that database region, an oasis could potentially pull in reads that would otherwise be assigned to multiple local taxa. The fewer other taxa around, the stronger the oasis effect, and the greater the number of reads incorrectly assigned to that taxon. Oasis taxa can systematically generate false positives that are not necessarily rare.

In this paper, we present PIA as a taxonomic binner which, like MEGAN, works from gold-standard BLAST output and is not designed specifically for microbial data, yet goes further to address the shortcomings of BLAST and databases. It also filters BLAST hits by a strict quality threshold. It also accounts for over-represented taxa by only counting each hit taxon once. It also avoids over-assigning conservative hits and sequences by finding a lowest common ancestor, here called a phylogenetic intersection to avoid ambiguity when dealing with ancient sequences that may genuinely be ancestral. However, there are two key differences between MEGAN and PIA. First is a difference with finding the intersection. MEGAN accepts an LCA calculated from just one taxon (i.e., that taxon itself), but if PIA does not have at least two taxa, it discards the read. It assumes that the real taxon is not in the database, so will not assign directly to a taxon in the database. It only assigns to a higher taxon, assuming that the real taxon lies within that phylogenetic range. This avoids over-assigning unrepresented reads to close relatives. Second is a diversity check that measures the extent of population in the region of the database. Reads assigned in sparse regions, vulnerable to the influence of oasis taxa, are discarded. PIA discards the majority of reads, but those that remain are robustly assigned. The resulting assignments are reliable despite low read counts.

This study evaluates PIA by benchmarking its performance against MEGAN with empirical and simulated data. The empirical data was generated as part of the Europe's Lost Frontiers project. This aims to reconstruct submerged palaeolandscapes around the UK, particularly Doggerland, which now lies under the North Sea. One arm of the project is multi-proxy analysis of sediment cores. This study uses our sedaDNA data from core ELF039, chosen because most samples had a relatively high data yield and the geological context suggested a potentially interesting story. For more information, see Gaffney *et al.* (2020).

3.2 Algorithm

A very early version of PIA was originally presented in Smith *et al.* (2015). Although the central approach has not changed, it has been substantially rewritten and refined. Scripts are available from https://github.com/Allaby-lab/PIA.

3.2.1 The input BLAST file

The two inputs for PIA are a FASTA of query sequences and a corresponding BLAST file. The BLAST file must be in format 6 (tabular) with all standard columns followed by an additional column containing taxonomic IDs associated with the reference sequence hit. This column is how PIA assigns hits to taxa. We also use the `-max_target_seqs` parameter to limit the number of hits returned per query sequence, recognising that the hits returned will be the first *n* to meet a quality threshold (Shah *et al.* 2018). Although PIA aims to reduce the impact of overrepresented taxa in databases once the BLAST is complete, it is important that this BLAST takes enough hits to reach underrepresented taxa. `-max_target_seqs` should be as high as practical. We suggest 500 as a default. Finally, note that BLAST can be run with *x* number of threads. Many of our larger samples took days to BLAST despite using several threads. This is by far the most computationally expensive part of the pipeline.

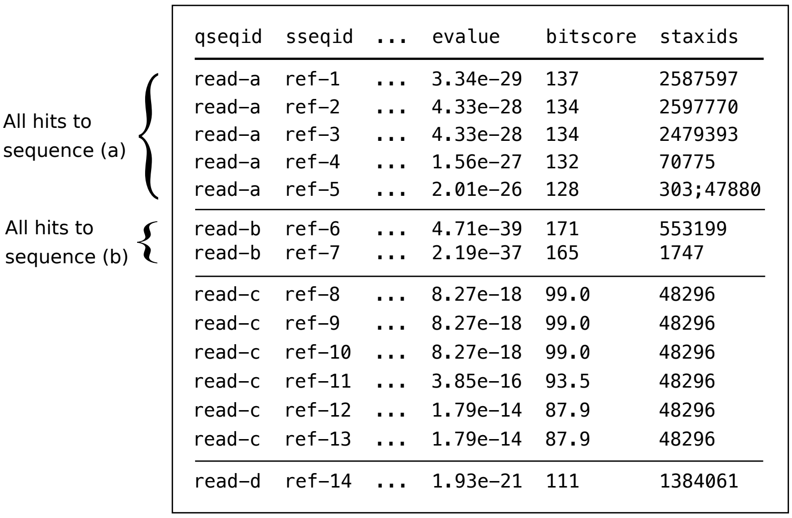
A typical pre-PIA BLAST command:

```
blastn -db [nucleotide database] -num_threads [x] -query [input
FASTA] -out [output] -max_target_seqs 500 -outfmt "6 std staxids"
```

The resulting BLAST file (figure 3.2) lists hits first by query sequence, so all hits to a query are together, and then by descending Expect value (E), so better matches are generally further up the list. However, within E value, the order is simply the order in which the hits occur in the database.

3.2.2 PIA

PIA is computationally light enough to be run on a laptop with small sample files (FASTA approximately <3 MB). The index-building step required before first use should



qseqid	sseqid	...	evalue	bitscore	staxids
read-a	ref-1	...	3.34e-29	137	2587597
read-a	ref-2	...	4.33e-28	134	2597770
read-a	ref-3	...	4.33e-28	134	2479393
read-a	ref-4	...	1.56e-27	132	70775
read-a	ref-5	...	2.01e-26	128	303;47880
read-b	ref-6	...	4.71e-39	171	553199
read-b	ref-7	...	2.19e-37	165	1747
read-c	ref-8	...	8.27e-18	99.0	48296
read-c	ref-9	...	8.27e-18	99.0	48296
read-c	ref-10	...	8.27e-18	99.0	48296
read-c	ref-11	...	3.85e-16	93.5	48296
read-c	ref-12	...	1.79e-14	87.9	48296
read-c	ref-13	...	1.79e-14	87.9	48296
read-d	ref-14	...	1.93e-21	111	1384061

Figure 3.2: Example partial BLAST output structure in format "6 std staxids". The standard (std) fields are the first columns, starting with query sequence (qseqid) and ending with Expect (E) value (evalue) and score (bitscore). Additional fields, here the taxonomic IDs (staxids) associated with the reference, are at the end. Each row is a hit between the query sequence and a reference sequence from the database. Hits are ordered first by query sequence, then by E value from lowest to highest.

take no more than a few minutes. Time to analyse the seven samples used in this study on one thread ranged from approximately ten seconds to ten minutes. PIA can also be multi-threaded for larger samples, for which we recommend a server.

Figure 3.3 illustrates the PIA algorithm. PIA considers one read at a time. Reads with no BLAST hits are discarded. For reads with hits, PIA first calculates the coverage of the top hit:

$$\%\, coverage = (matchlength)/(readlength) * 100$$

If the coverage does not meet a threshold (default 95%), the read is discarded. The taxonomic assignment of the read is strongly influenced by the top hit, so it only accepts a very close match.

PIA then considers each hit in order of the BLAST file. First, the hit is assigned to a taxon. If a hit is associated with multiple taxa, PIA assumes that this indicates a conservative sequence and assigns the hit to the phylogenetic intersection of those taxa. The assigned taxon is then evaluated. If there has already been a hit to the taxon, the hit is discarded. Because hits are listed in order of E value, this means that only the best hit for each taxon is retained. This taxon check aims to mitigate the problem of overrepresented taxa. Provided that the BLAST found enough hits to reach underrepresented taxa in the database at all, this check gives them equal weight to overrepresented taxa. Every taxon is reduced to a single hit.

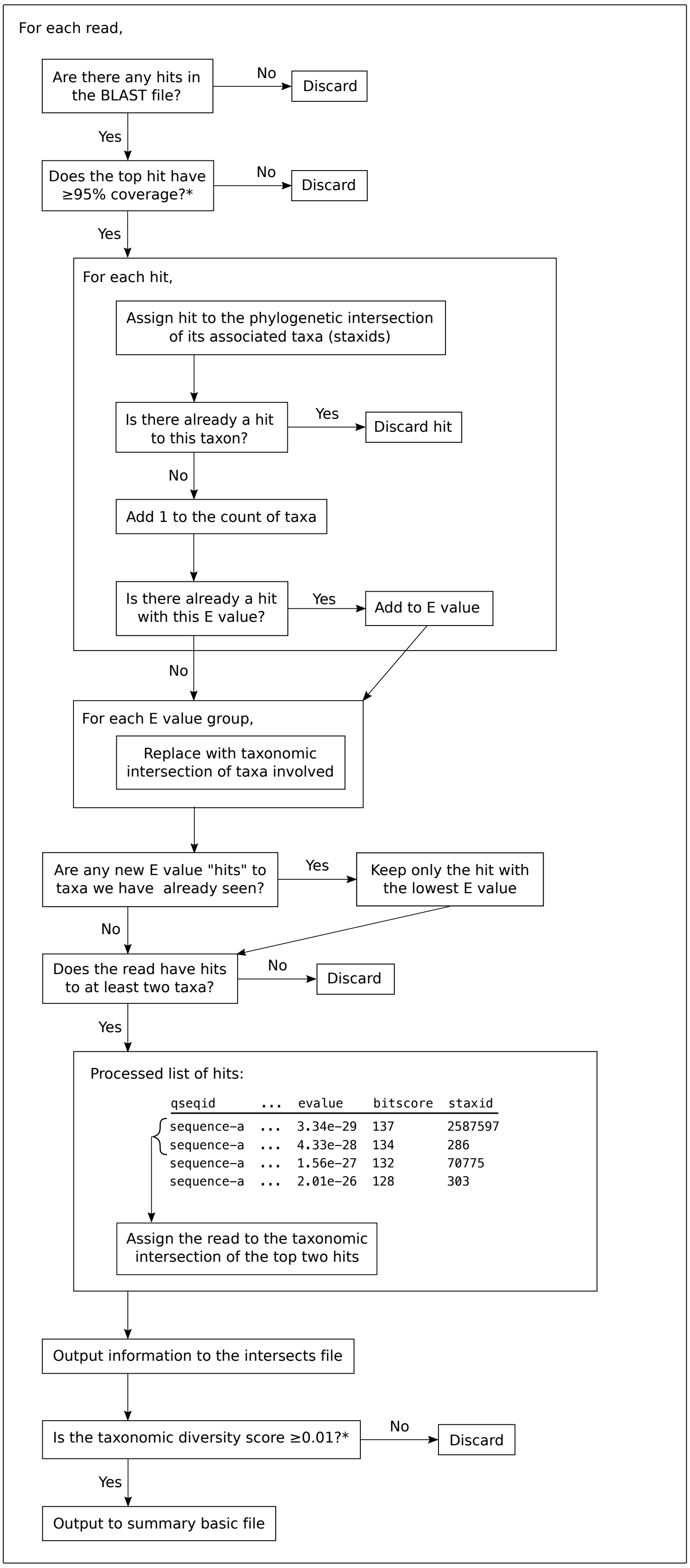
The second check performed on each hit is the E value. If there has already been a hit that passed the taxon check with this E value, those hits are grouped together. Once all hits for this read have been taxon-checked and grouped by E value, the E value groups are collapsed to a single "hit" per E value. This "hit" is the phylogenetic intersection of the group members. If a read is found to be equally similar to sequences from several different organisms, PIA again assumes that this indicates a conservative sequence. Finally, if these new “hits” are to previously seen taxa, then as before, only the hit with the best E value is retained.

Once the list of BLAST hits for the read has been reduced to one (best) hit per taxon, PIA assigns the read to the phylogenetic intersection of the top and second-top hits. If only one hit remains, there cannot be an intersection, so the read is discarded. Finding the intersection firstly avoids over-assigning conservative sequences. Secondly, it avoids over-assigning reads from unrepresented taxa to close represented relatives. PIA assumes that the real taxon is not in the database, so it will not assign directly to any organism in the database. The intersection is only taken between the top two hits because, after the taxon check and grouping by E value, those two hits may already represent distantly-related and/or higher taxa.

The final step is the diversity check, which filters reads by taxonomic diversity score:

$$Taxonomic\ diversity\ score = (t - 1)/c$$

Where t is the number of different taxa in the original list of BLAST hits and c is a predefined cap on the number of hit taxa to consider. The score measures how populated this area of the database is. A well-populated region will have more hits. If the region is sparsely-populated, there may be a disproportionately high number of hits to oasis taxa. Reads which seem to match an organism in a too sparsely-populated area are discarded.



qseqid	...	evalue	bitscore	staxid
sequence-a	...	3.34e-29	137	2587597
sequence-a	...	4.33e-28	134	286
sequence-a	...	1.56e-27	132	70775
sequence-a	...	2.01e-26	128	303

*Adjustable parameter

Figure 3.3: Flowchart illustrating the PIA algorithm. There are three key checks that may result in a read being discarded: sufficient coverage of the top BLAST hit, at least two hits remaining after processing, and a high enough taxonomic diversity score. Reads that pass are assigned to the intersection of the top two remaining BLAST hits.

3.3 Methods

3.3.1 Analysis of empirical sedaDNA data

PIA and MEGAN were compared in a parallel analysis of seven samples from the Europe's Lost Frontiers project (Gaffney *et al.* 2020). These samples are from sediment core ELF039 which was taken from a palaeochannel approximately 50 km north of the present Norfolk coast. No dates were available for that core at the time of writing, but the channel is interpreted as a river valley that underwent marine inundation during the early Holocene. The samples were shotgun sequenced on a NextSeq 550 as part of our work using sedaDNA for palaeoenvironmental reconstruction. We typically focus on plants because of their high biomass in most environments, increasing the chance of DNA deposition, and the abundance of ecological and distribution information available. Accordingly, this study made use of reads from Viridiplantae.

Chapter 2 (Main materials and methods) of this thesis describes data collection and analysis. This study used data from the initial DNA extractions of ELF039_250, 321, 355, 384, 415, 460, and 485. ELF039_145 was excluded due to low data yield. Replicates were not sequenced at this time. The seven sample files had reads assigned by MEGAN and PIA and were filtered by negative controls. The results showed little variation between samples, so the samples were concatenated together before visualisation with Krona (Ondov *et al.* 2011; see online supplementary material at https://doi.org/10.3389/fevo.2020.00084).

3.3.2 Accuracy testing with simulated data

Benchmarking against MEGAN suggested that PIA may successfully increase the accuracy of taxonomic assignments at the cost of sensitivity. To test the accuracy more objectively, we ran both MEGAN and PIA on two test datasets of known GenBank sequences. For each dataset, the control condition used the original BLAST database from the benchmarking analysis (downloaded on 05-09-2019). An "exclusion" condition excluded all taxa in the test dataset from the BLAST database. This aimed to simulate the unrepresented taxa, common in metagenomic data, that PIA is designed to analyse. In each condition, we tracked the assignments of individual sequences and compared them to the actual source organisms. Most stages involved custom scripts available from https://github.com/Allaby-lab/PIA-accessories and detailed in the online supplementary material (https://doi.org/10.3389/fevo.2020.00084).

Each test dataset comprised 250 GenBank sequences downloaded through the NCBI website. For the first dataset, sequences were first filtered to Embryophyta and to a length of 30-150 bp to reflect typical aDNA. We then iterated through "All other taxa" from the "Results by taxon" option until taxa were represented by no more than 44 relevant sequences. Metagenomic data is likely to contain poorly-represented sequences. Single sequences from 245 taxa were downloaded as a FASTA with GIs included. An additional five 30-150 bp sequences were added from well-represented domesticates: *Hordeum vulgare*, *Musa acuminata*, *Triticum dicoccon*, *Triticum aestivum* and *Zea* *mays*. These were run through BLASTn to check that they did match their taxa labels, as model organism sequences are frequently assigned to incorrect taxa. The second dataset was built in the same way but was first filtered to Mammalia instead of Embryophyta. The low-frequency taxa were represented by up to 47 relevant sequences and the five high-frequency taxa were *Camelus bactrianus*, *Camelus dromedarius*, *Balaenoptera bonaerensis*, *Chlorocebus aethiops* and *Papio anubis*. Finally, each FASTA file was re-formatted to single-line using fasta_formatter from the FAST-X toolkit 0.0.13 (Gordon and Hannon 2010).

The FASTAs were run through BLAST with the same settings as in benchmarking. The exclusion condition only differed in the reduced database. For every taxon, a list of GIs for all sequences from that taxon was downloaded from GenBank. These lists were concatenated into a master GI list. The BLAST option `-negative_gilist` was used to exclude this list from the database. For each BLAST file, the MEGAN and PIA analyses were performed with the same settings as in benchmarking. See the online supplementary material for details (https://doi.org/10.3389/fevo.2020.00084).

It became apparent after analysis that two Mammalia sequences may be affected by human contamination: GI 2198752 (accession no. U84666.1, *Cavia porcellus* Y5 scRNA gene, partial sequence) and GI 13508496 (accession no. AY028924.1, *Mammut americanum* 16S ribosomal RNA gene, partial sequence; mitochondrial gene for mitochondrial product). We ran BLAST on both sequences to check, changing `-max_target_seqs 500` to `-num_alignments 1` to produce easily readable output with the default limit of 500 hits. Other settings were the same as in benchmarking.

Finally, a small separate test of GenBank data was used to evaluate the performance of PIA on highly divergent taxa. Because of the diversity check, we expect PIA to unnecessarily discard reads assigned to taxa with few living relatives because their region of the database will always appear incomplete. We ran BLAST and PIA on the available GenBank sequences from two monotypic orders: Ginkgoales (containing the gymnosperm Ginkgo; 22,600 sequences) and Microbiotheria (containing the marsupial Dromiciops gliroides; 417 sequences). This used the same settings as in benchmarking.

3.4 Results

3.4.1 Analysis of empirical sedaDNA data

Taxonomic assignments of early Holocene sedaDNA from a submerged palaeochannel in the North Sea by MEGAN and PIA are compared in figures 3.4 and 3.5. The most frequent taxa are labelled in full. Of these, taxa not native to Europe are highlighted in bold.

The taxonomic profiles of the MEGAN and PIA outputs are broadly similar (figures 3.4 & 3.5). Figure 3.4 begins at Mesangiospermae, to which the vast majority of reads are assigned by both methods. Most reads are assigned to *Zostera marina* (eelgrass), related taxa in Potamogetonaceae or to its parent order Alismatales, suggesting a wetland or fully aquatic environment with at least some saltwater influence. There is also a sizeable signal from grasses (Poaceae). In the largest remaining segment, Pentapetalae (figure 3.5), both profiles show a diverse range of taxa found in northwest Europe today. This includes Rosaceae (strawberry, bramble, apple, drupe trees), *Salix* (willow), *Populus* (poplar), and Fagales (birch, oak).

However, the numbers of reads making up these taxa differ significantly. Though proportionally similar, the MEGAN profile was built from 88,497 reads compared to just 27,547 accepted by the PIA. The MEGAN profile also has higher taxonomic richness, containing 374 taxa versus 210 (table 3.1). Those MEGAN taxa are also generally more specific. MEGAN assigned far more reads to genus or lower. Overall, the results are consistent with MEGAN placing more emphasis on sensitivity than PIA.

Because the samples originate from northwest Europe in the early Holocene, we would expect DNA sequences to be comparable to European taxa today. The samples have been filtered by negative controls which should have removed most assignments to common modern contaminant taxa present in reagents. We therefore assume any assignments to non-European taxa to be false positives.

Many of the most frequent non-European taxa assigned to by MEGAN are domesticated grasses such as *Oryza*, *Setaria italica* and *Sorghum bicolor* (figure 3.4). In Pentapetalae (figure 3.5), most of the terminal taxa in the MEGAN output – those genera and species that suggest a higher sensitivity than PIA – are non-European and therefore likely false positives. Table 3.1 quantifies all assignments: 40.11% of taxa in the MEGAN profile are suspect compared to 20.95% for PIA. In total, MEGAN assigned 12.78% of reads to non-European taxa and PIA assigned just 0.52%. The false positive taxa have lower counts on average, suggesting that the minimum support filter in MEGAN is a valid approach, but in this case PIA was more effective at removing this sort of false positive.

It appears that the lower sensitivity of PIA is associated with higher accuracy. To investigate this more objectively, we ran PIA on test sequences of known origin.

Table 3.1: Numbers of European and non-European taxa hit and the numbers of reads assigned to each category in the MEGAN and PIA benchmarking output. Reads assigned to non-European taxa are suspected to be false positives for this data.

	MEGAN	PIA
Total taxa	374	210
European	224 (59.89%)	166 (79.05%)
Non-European	150 (40.11%)	44 (20.95%)
Total reads	88497	27547
To European taxa	77,189 (87.22%)	27,405 (99.48%)
To non-European taxa	11,308 (12.78%)	142 (0.52%)

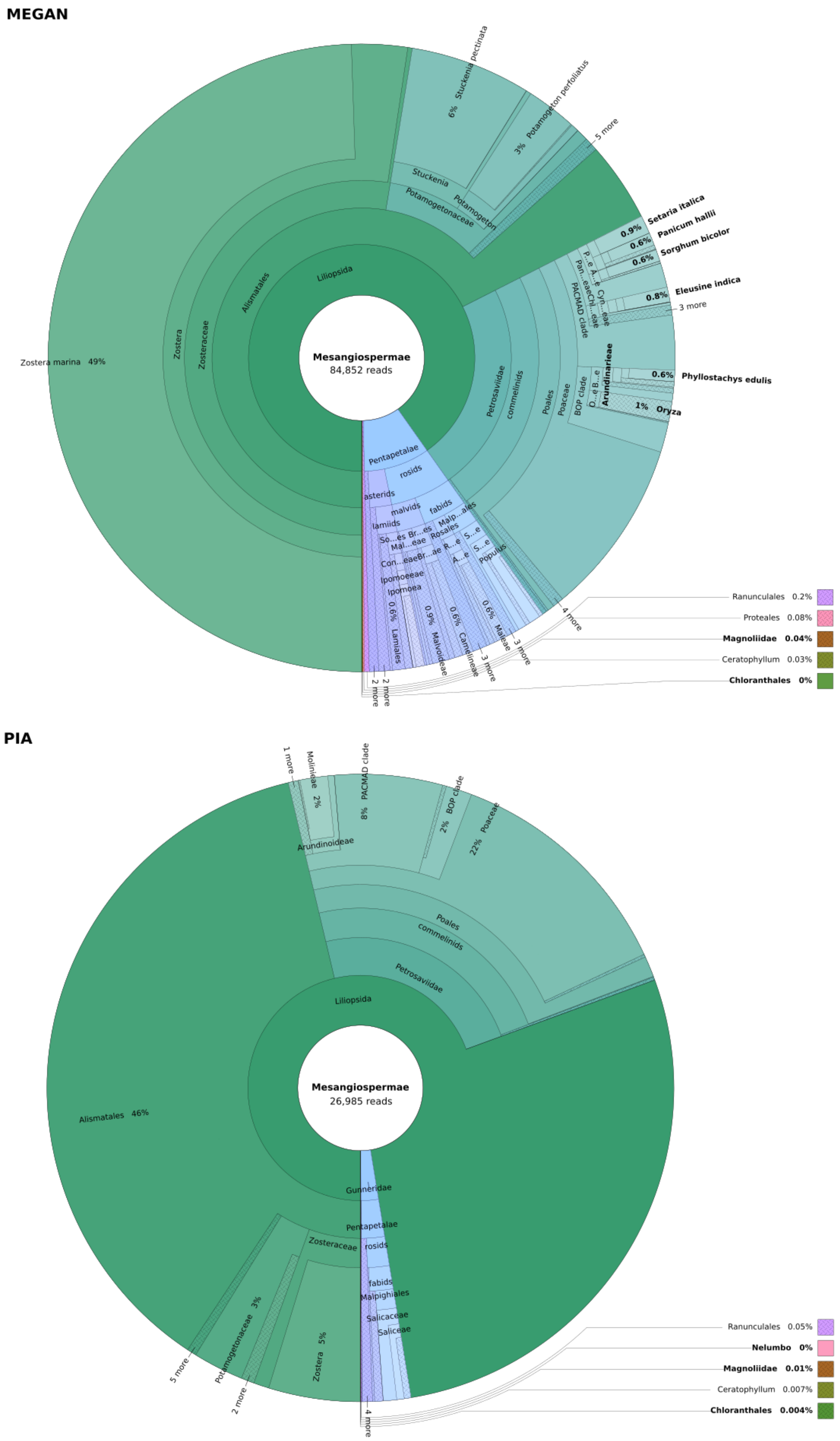
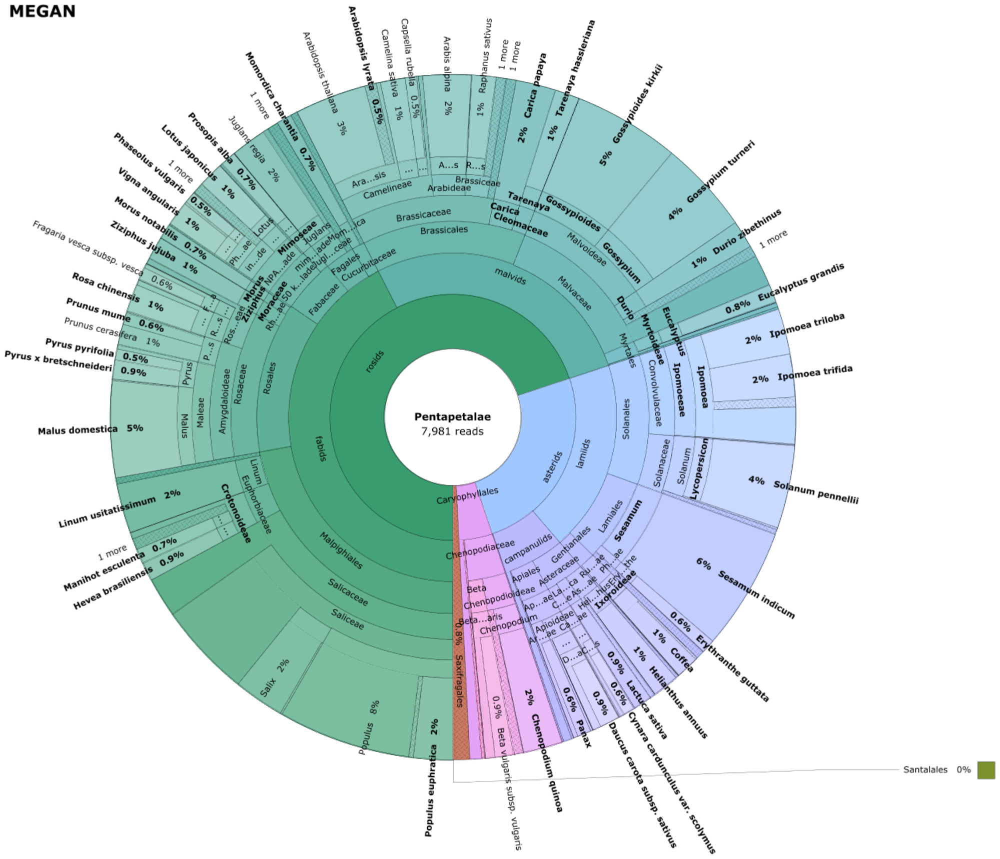


Figure 3.4: Taxonomic profiles of the combined MEGAN and PIA outputs for the seven sediment samples after filtering each by negative controls. The highest taxon shown is Mesangiospermae, which includes the vast majority of reads. Taxa not native to Europe, which are suspected to be false positive assignments for this data, are highlighted in bold. Colours indicate the proportion of reads assigned to that taxon.

MEGAN



PIA

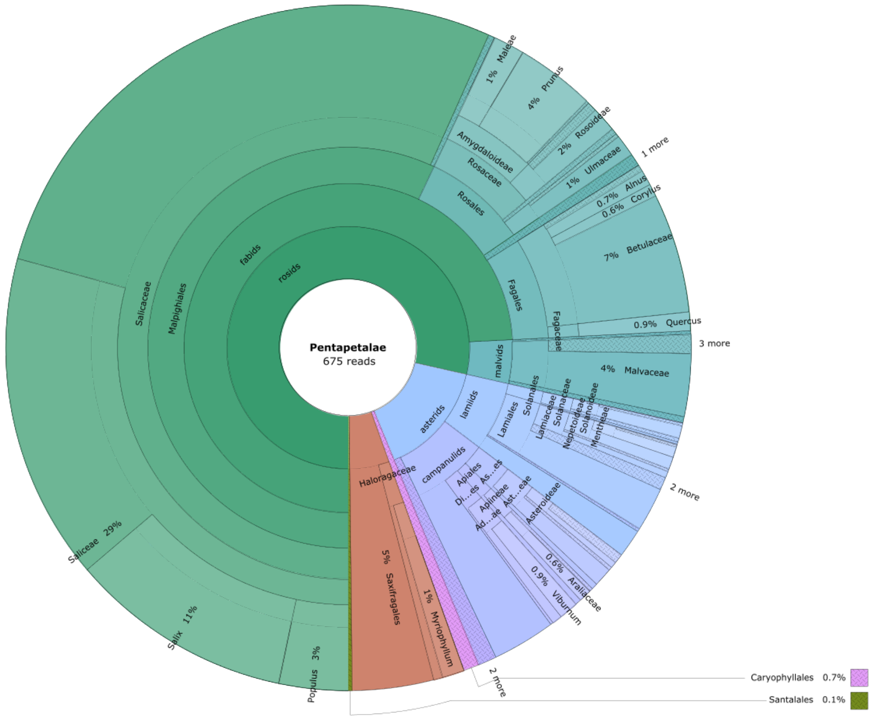


Figure 3.5: Taxonomic profiles of the combined MEGAN and PIA outputs for the seven sediment samples after filtering each by negative controls. The highest taxon shown is Pentapetalae, which is the largest wedge in figure 3.4 not easily seen. Taxa not native to Europe, which are suspected to be false positive assignments for this data, are highlighted in bold. Colours indicate the proportion of reads assigned to that taxon

3.4.2 Accuracy testing with simulated data

Embryophyta

Individual reads, their source organism and all four assignments are listed in the first worksheet of supplementary table S2 (https://doi.org/10.3389/fevo.2020.00084). Table 3.2 provides a summary. We considered an assignment correct if it was to the actual taxon or one of its parent taxa. For example, if PIA assigned a read from *Betula* to the family Betulaceae, it would be a correct assignment at family level. Family level is typically precise enough to be useful for environmental reconstruction in plants. An assignment to Viridiplantae would be correct at kingdom level. An assignment to Poaceae would be incorrect.

In the control condition, MEGAN assigned 91% of sequences and PIA 52%, mirroring the higher sensitivity of MEGAN observed in the analysis of real data. Both were highly accurate at 97% and 100% respectively. MEGAN was somewhat more precise, with 62% of assignments correct to family level or below, compared to 53.49% for PIA. Overall, MEGAN showed a much greater ability to assign sequences at the cost of a very small drop in accuracy compared to PIA.

The exclusion condition, where the source taxa had been removed from the database, shows a similar pattern of results with generally worse performance by both tools. However, MEGAN appears to suffer more. The “Change” columns in table 3.2 shows that MEGAN assigns proportionally fewer sequences at all, correctly, and with precision than PIA. Notably, accuracy of MEGAN falls to 80% but that of PIA remains at a healthy 96%.

Despite the exclusion database generally presenting more of a challenge, there were a small number of sequences that were assigned better than with the complete database. PIA did not assign the *Lapageria rosea* and *Lupinus luteus* sequences in the control condition but matched MEGAN’s broad Mesangiospermae assignment for the exclusion. Both MEGAN and PIA assigned the *Metasequoia glyptostroboides* and *Magnolia* x *soulangeana* sequences more precisely in the exclusion condition, although not particularly so. This unexpected behaviour may be due to peculiarities of the database around those sequences.

Mammalia

Full results are listed in the second worksheet of supplementary table S2 (https://doi.org/10.3389/fevo.2020.00084). Table 3.3 provides a summary. In the control condition, the Mammalia dataset showed a similar pattern to Embryophyta. MEGAN assigned more reads and with more precision; both programs were very accurate. The exclusion condition resulted in worse performance for both programs, again with a greater impact on MEGAN. However, the decrease in accuracy was even more pronounced than for Embryophyta. MEGAN only assigned 60% of sequences accurately. PIA assigned 83% accurately, which while better, is far from the 96% accuracy seen for Embryophyta.

Note that these accuracy results are likely a slight underestimate, as the two questionable sequences do appear to be mislabeled. Both BLAST outputs are dominated by *Homo sapiens* and other primates. MEGAN and PIA generally assigned them either to higher mammal taxa or close parent taxa of humans, both of which are reasonable if the sequences are actually human.

As with Embryophyta, a small number of sequences were assigned better with their taxa excluded from the database. MEGAN assigned the *Stenella attenuata* sequence incorrectly in the control but broadly correct after exclusion. PIA assigned the *Kogia sima* sequence more precisely after exclusion, though only by one level. Finally, the only time the *Halichoerus grypus* sequence was assigned was by PIA after exclusion, and it did so correctly to family.

Monotypic taxa

PIA assigned 5% of reads from Ginkgoales and with only 77% accuracy. For Microbiotheria, PIA assigned 37% of reads; 100% were accurate but the most precise was only to Metatheria. The proportion of reads assigned to each was considerably lower than the 50-60% from the mixed test datasets above.

Table 3.2: Percentages of the 250 sequences assigned by MEGAN and PIA in the Embryophyta accuracy test. The control condition BLASTed against the full GenBank nucleotide database (downloaded on 05-09-2019). The exclusion condition omitted the source taxa from the database. Of those reads assigned, percentages assigned incorrectly or correctly are given. The final two rows detail whether correctly-assigned reads were assigned to higher taxa or to at least family. These rows sum to the total percent correct.

	Control BLAST		Exclusion BLAST		Change	
Embryophyta	MEGAN	PIA	MEGAN	PIA	MEGAN	PIA
Assigned	91.20%	51.60%	76.00%	45.60%	-15.20%	-6.00%
Incorrect	3.07%	0.00%	20.00%	4.39%	16.93%	4.39%
Correct	96.93%	100.00%	80.00%	95.61%	-16.93%	-4.39%
Correct to above family	35.09%	46.51%	46.84%	60.53%	11.75%	14.02%
Correct to family or below	61.84%	53.49%	33.16%	35.09%	-28.68%	-18.40%

Table 3.3: Percentages of the 250 sequences assigned by MEGAN and PIA in the Mammalia accuracy test. Details are the same as for table 3.2.

	Control BLAST		Exclusion BLAST		Change	
Mammalia	MEGAN	PIA	MEGAN	PIA	MEGAN	PIA
Assigned	93.60%	57.60%	76.40%	52.40%	-17.20%	-5.20%
Incorrect	2.99%	0.00%	40.31%	16.79%	37.32%	16.79%
Correct	97.01%	100.00%	59.69%	83.21%	-37.32%	-16.79%
Correct to above family	28.21%	45.14%	41.36%	49.62%	13.36%	4.48%
Correct to family or below	68.80%	54.86%	18.32%	33.59%	-50.48%	-21.27%

3.5 Discussion

Ancient metagenomics has much potential, but taxonomic assignment of reads can be improved. Databases are highly uneven, resulting in the joint problems of over-represented taxa filling up hit lists at the expense of poorly-represented but closer matches, and oasis taxa in sparsely-populated areas drawing in reads and giving an illusion of confident assignment. There are methods that partly address these problems in some circumstances, but we demonstrate here that PIA performs strongly, providing an objective approach to remove false positives from data sets. Benchmarking on plant sedaDNA data against a standard tool, MEGAN, showed that PIA produces a comparatively low-resolution taxonomic profile. Far fewer reads are assigned and those that are rarely make it to genus. However, we argue that much of the sensitivity of MEGAN in this context is over-sensitivity. Both methods describe core ELF0039 as coming from a primarily wetland environment, with a clear signal from fresh and saltwater plants in Alismatales and the riparian *Salix*, along with some signal from grasses in Poaceae and woodland trees in Fagales. Yet the MEGAN profile assigned nearly 13% of reads to clearly questionable taxa, such as the tropical *Sorghum bicolor*, Australasian *Eucalyptus* and American *Carica papaya*, that if taken at face value would present a radical departure from the established palaeoecology of Europe. Once such taxa are removed as 'known' false positives, the MEGAN analysis only retrieves a little more than PIA (figure 3.5), which add little to the palaeoecological reconstruction and likely still contain false positives. One example is *Arabidopsis thaliana*, a known model organism not expected to feature greatly in the Mesolithic. In our context, the additional accuracy of PIA appears to outweigh the increased sensitivity of MEGAN.

The accuracy test on simulated data returned similar results. With a full BLAST database, MEGAN assigned nearly twice as many sequences with greater precision and only marginally lower accuracy than PIA. However, when the source taxa were excluded from the database, exacerbating the problems caused by incomplete databases and better representing real metagenomic data, the improvements of MEGAN over PIA diminished and the difference in accuracy became substantial. For Embryophyta sequences, PIA maintained a very high accuracy of 96%, whereas that of MEGAN fell to 80%.

Both programs performed less well with the Mammalia dataset, but PIA still returned 83% accuracy after exclusion of source taxa compared to 60% from MEGAN. We suspect that this difference may simply be due to the fact that there are far fewer species of mammal than embryophyte, so removing 250 mammal taxa will have removed proportionally more of the relevant database than removing the same number from Embryophyta. Both PIA and MEGAN performed very well in the control condition, so it is unlikely to be directly due to the mammal sequences themselves. Instead, we suggest that the exclusion condition simulated a more incomplete database for Mammalia than Embryophyta. PIA still outperformed MEGAN. However, it is clear that while PIA copes better with incomplete databases, it is not a perfect solution.

Additionally, two specific limitations of PIA are apparent from its algorithm. First, PIA cannot assign to leaf taxa. It can only assign to a species if there are subspecies in the database, for example. PIA does not fully take advantage of sequences with very high taxonomic resolution. If higher resolution is desired, it may be helpful to first identify reads to higher taxa more accurately using PIA, then further analyse any sequences assigned to taxa of interest using a different approach.

The second limitation is a result of the taxonomic diversity check. PIA discards assignments to taxa in sparse areas of the database because these areas are vulnerable to the influence of oasis taxa. However, this assumes that sparsity is due to incompleteness. There are divergent taxa with very few living relatives that will occupy a naturally sparse database region. PIA is less likely to accept assignments to these taxa. To demonstrate this, we ran PIA on the available GenBank sequences from Ginkgoales and Microbiotheria, which are orders containing a single species. PIA assigned fewer reads from these taxa than from the mixed Embryophyta or Mammalia datasets. Such divergent taxa are unusual, but are less likely to be recovered by PIA. Again, PIA shows a lack of sensitivity that may limit its application in some studies.

However, even with these caveats, we have demonstrated that the improved ability of PIA to address the challenges of an incomplete reference database can result in highly accurate taxonomic assignment of metagenomic shotgun data. PIA produced fewer false positives than the standard approach. The more likely false positives are to occur, the more necessary it becomes to manually sort taxa into plausible and implausible, which requires subjective presuppositions about the source of the data. This is particularly problematic for ancient metagenomics where little is known about the study environment. PIA offers an objective alternative with an estimated 96% accuracy for plants.

Chapter 4

Taxonomic results

4.1 Comparing taxonomic assignment by MEGAN and PIA

Assignment with MEGAN and PIA is compared in detail for a subset of data from core ELF039 in Chapter 3 (PIA). This short section reports a broader analysis of the Viridiplantae results for all ELF021-60 samples. Reads were assigned by MEGAN5 and PIA using the results of the thorough BLAST search (`-max_target_seqs 500`). Samples were filtered by negative controls and replicates were merged. The final read counts for MEGAN and PIA, and the MEGAN:PIA ratio, are listed in table 4.1.

Table 4.1: Read counts for all samples before processing and when assigned by MEGAN or PIA, and the MEGAN:PIA read count ratio. After MEGAN or PIA assignment, samples were filtered by negative controls and replicates merged. Columns are shaded with most saturation for the most extreme values. The ratio column has red for ratios below 1 (PIA assigned more reads) and blue for above 1 (MEGAN assigned more reads).

Sample	Raw count	MEGAN count	PIA count	MEGAN:PIA ratio
ELF022_140	266431	0	2	0.00
ELF027_160	8274674	4630	775	5.97
ELF027_174	36426308	200444	102955	1.95
ELF027_177	1184516	349	69	5.06
ELF031_033	906692	1	0	NA
ELF031_043	1042697	146	16	9.13
ELF031_056	5241159	665	50	13.30
ELF031A_058	105532766	41467	7771	5.34
ELF031A_072	3146129	22664	8205	2.76
ELF031A_088	1941179	2477	800	3.10
ELF031A_107	1059728	940	249	3.78
ELF031A_123	1200038	3491	775	4.50
ELF031A_152	1058369	3616	343	10.54
ELF031A_177	1167607	711	276	2.58
ELF031A_202	1258821	185	56	3.30
ELF031A_219	165408	14	0	NA
ELF031A_281	390393	6	6	1.00
ELF031A_310	150551	0	0	NA
ELF032A_047	12309084	15318	2838	5.40
ELF032A_074	40860623	27611	5438	5.08
ELF032A_095	11997084	20066	3681	5.45
ELF032A_117	29181519	10708	1404	7.63
ELF032A_153	16661747	16401	2384	6.88
ELF032A_177	16269458	20645	6410	3.22
ELF033_046	5852731	1034	32	32.31
ELF033_075	7278259	2255	29	77.76
ELF033_118	9817442	18912	3290	5.75
ELF033_155	12239164	1740	70	24.86
ELF033_176	6481542	3239	213	15.21
ELF033_183	1529949	335	60	5.58
ELF033_187	2320753	1514	270	5.61
ELF033_195	326132	384	54	7.11
ELF033_203	334004	262	48	5.46
ELF033A_050	2499171	279	8	34.88
ELF033A_070	4032588	703	26	27.04
ELF033A_098	2153881	666	49	13.59
ELF033A_111	4547485	941	69	13.64
ELF033A_126	1002832	111	7	15.86
ELF033A_158	1522189	2361	449	5.26
ELF034_061	1173521	53	1	53.00
ELF034_079	904449	14	0	NA

Continued on next page

Table 4.1 continued

Sample	Raw count	MEGAN count	PIA count	MEGAN:PIA ratio
ELF034_094	2332425	295	0	NA
ELF034_132	2388531	132	2	66.00
ELF034_157	259937	7	1	7.00
ELF034_177	3756487	517	4	129.25
ELF034_185	1530287	84	4	21.00
ELF034_202	3025498	264	32	8.25
ELF034_219	1094648	144	15	9.60
ELF034A_063	36584170	14067	1870	7.52
ELF034A_081	5208489	2359	528	4.47
ELF034A_126	16806593	5073	493	10.29
ELF034A_146	2184818	205	3	68.33
ELF034A_166	709251	102	12	8.50
ELF034A_172	5755371	543	5	108.60
ELF034A_183	2695543	118	7	16.86
ELF034A_195	2012268	159	2	79.50
ELF034A_225	1065852	30	0	NA
ELF034A_261	1410734	65	1	65.00
ELF034A_282	330456	7	3	2.33
ELF039_145	116883	0	0	NA
ELF039_250	24402887	293156	122474	2.39
ELF039_321	61257128	301336	157186	1.92
ELF039_341	53400769	3118	234	13.32
ELF039_355	11319691	14309	5347	2.68
ELF039_384	6388822	3748	369	10.16
ELF039_415	24128997	20146	5820	3.46
ELF039_460	21199771	10215	2834	3.60
ELF039_485	767755	134	18	7.44
ELF040A_095	14630337	14513	2242	6.47
ELF040A_112	5433987	1731	132	13.11
ELF040A_192	6161930	9474	971	9.76
ELF040A_208	3716264	732	44	16.64
ELF040A_298	564827	3875	1462	2.65
ELF040A_350	326888	3582	1449	2.47
ELF040A_487	129161	10	5	2.00
ELF041_087	1069047	909	157	5.79
ELF041_110	1440606	2840	624	4.55
ELF041_180	1187651	4399	1584	2.78
ELF041_295	1699659	1690	75	22.53
ELF042_065	5130157	26018	8926	2.91
ELF042_151	1084493	5564	2113	2.63
ELF042_250	2157404	4973	863	5.76
ELF042_350	1599061	9851	2686	3.67
ELF044_090	5918708	16105	7041	2.29
ELF044_137	2498990	2582	638	4.05
ELF044A_097	3851033	41076	18264	2.25
ELF045_090	1763219	3832	1765	2.17
ELF045_145	40272649	54742	29014	1.89
ELF045_252	3159654	9712	3810	2.55
ELF045_346	4130095	5844	1041	5.61
ELF045_450	2848986	9694	519	18.68
ELF045_522	10219099	7233	2066	3.50
ELF046A_270	834665	9	0	NA
ELF047_070	1656409	876	68	12.88
ELF047_150	4096257	13416	5529	2.43
ELF047_241	782240	6931	2703	2.56
ELF047_274	2264573	1588	407	3.90
ELF047_325	2645831	8411	4540	1.85
ELF047_386	3653092	4598	1889	2.43
ELF047A_050	5161831	5231	1969	2.66
ELF047A_150	617870	94	4	23.50
ELF047A_256	3325395	230	20	11.50
ELF047A_354	3504397	240	17	14.12
ELF049_295	8815997	6449	1250	5.16
ELF049_361	1565038	596	140	4.26
ELF050_250	393246	5	1	5.00
ELF050_354	496587	12	3	4.00
ELF050_450	344654	70	9	7.78
ELF050_595	260600	6	3	2.00
ELF051_096	1485178	413	67	6.16
ELF051_120	1658132	25	11	2.27
ELF051_151	1635114	559	24	23.29
ELF051_196	12342349	6991	523	13.37
ELF051_255	13852327	11151	2787	4.00
ELF051_292	11121017	21545	8514	2.53
ELF053_179	327994	6	0	NA
ELF053_214	246529	19	3	6.33
ELF053_275	2060870	46	1	46.00
ELF053_289	1009480	114	20	5.70
ELF053_336	1194948	38	4	9.50
ELF054_058	14363736	15349	2247	6.83
ELF054_140	3321726	528	87	6.07
ELF054_182	1191833	2388	467	5.11
ELF054_268	1713992	372	35	10.63
ELF054_291	4726612	398	25	15.92
ELF054_315	2729077	476	23	20.70
ELF054_330	2826912	313	12	26.08
ELF054_356	8347570	1019	2	509.50
ELF059_210	10679410	17135	6274	2.73
ELF059_230	16183635	4971	1415	3.51
ELF059_270	15313599	8756	2590	3.38
ELF059_280	6946403	1678	108	15.54
ELF059_337	5618718	3334	1222	2.73
ELF059_359	3362236	10555	4938	2.14
ELF059_378	4815136	2749	138	19.92
ELF059A_135	16494644	60480	27335	2.21
ELF059A_190	31963255	40562	18256	2.22
ELF059A_250	37218086	49563	21335	2.32

Continued on next page

Table 4.1 continued

Sample	Raw count	MEGAN count	PIA count	MEGAN:PIA ratio
ELF059A_320	2939068	512	108	4.74
ELF059A_355	102493315	135879	55422	2.45
ELF060_250	8970504	5498	2303	2.39
ELF060_350	5930745	13126	6266	2.09
ELF060_420	3180010	15314	6384	2.40
ELF060_465	6906126	4601	1900	2.42

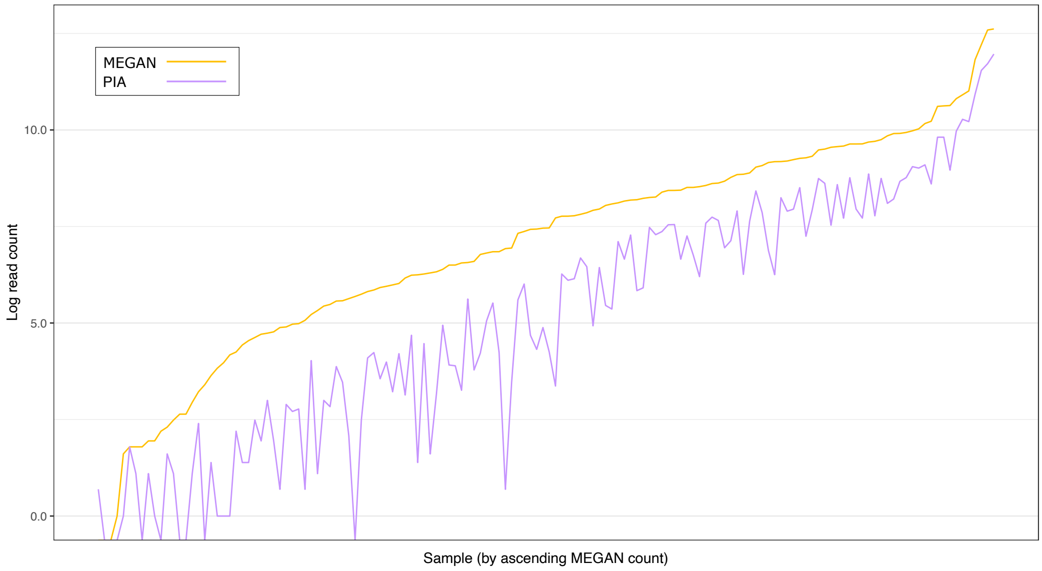


Figure 4.1: Read counts for Viridiplantae in all ELF samples when assigned with MEGAN or PIA, filtered by negative controls, and replicates merged. Read counts have been natural-log-transformed.

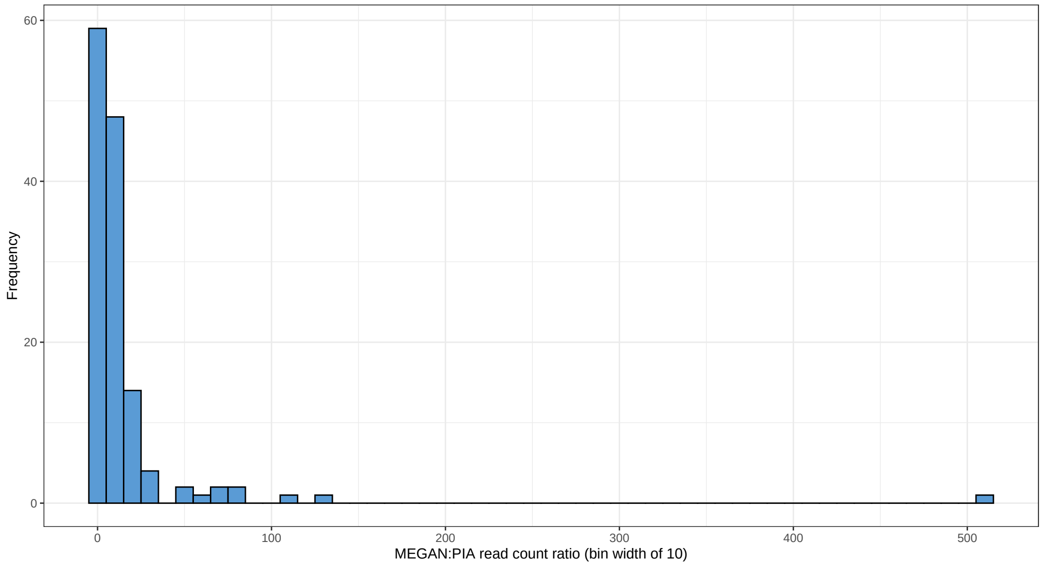


Figure 4.2: Histogram of ratios between MEGAN and PIA read counts for Viridiplantae, filtered by negative controls, in all samples. Ratios could not be calculated for the nine samples where PIA returned 0 reads.

The expected result after Chapter 3 (PIA) would be for MEGAN to show greater sensitivity but lower accuracy than PIA. Figure 4.1 addresses sensitivity by comparing read counts (note the log scale). The MEGAN count is greater than PIA for all but two samples: the PIA count was equal for ELF031A_281 and greater for ELF022_140 (table 4.1). However, those PIA counts are only 6 and 2 reads respectively. Both are too close to zero to be strong counterexamples. MEGAN does appear more sensitive than PIA.

Estimating accuracy requires detailed analysis and was not feasible for this number of samples. However, it was used to investigate the one sample with a particularly high MEGAN:PIA count ratio; a sample where, at first glance, MEGAN outperformed PIA beyond normal expectations. Figure 4.2 shows a histogram of count ratios. The majority (86%) of ratios are between 1 and 50, indicating that MEGAN returned up to 50 times more reads than PIA. However, there is a clear outlier with a ratio of 509.5. This sample, ELF054_356, was analysed at taxon level (table 4.2).

Table 4.2: MEGAN and PIA taxonomic profiles for Viridiplantae from ELF054_356, which had a particularly large MEGAN:PIA ratio. Data had been filtered by negative controls and replicates combined. Taxa are ordered by total read count and those not native to Europe are in bold.

Common name	Taxon	Total	MEGAN	PIA
	All	1021	1019	2
	Gossypium raimondii	**75**	**75**	**0**
	Gossypioides kirkii	**70**	**70**	**0**
	Solanum pennellii	**56**	**56**	**0**
Narrowleaf lupin	Lupinus angustifolius	53	53	0
Adzuki bean	**Vigna angularis var. angularis**	**52**	**52**	**0**
Pineapple	**Ananas comosus var. bracteatus**	**45**	**45**	**0**
Common flax	**Linum usitatissimum**	**34**	**34**	**0**
	Gossypium turneri	**31**	**31**	**0**
Cacao	**Theobroma cacao**	**30**	**30**	**0**
Greater duckweed	Spirodela polyrhiza	30	30	0
	Sesamum indicum	**20**	**20**	**0**
Alpine rock-cress	Arabis alpina	20	20	0
	Caulerpa verticillata	17	17	0
Morning glories	**Ipomoea**	**16**	**16**	**0**
	Lotus japonicus	**16**	**16**	**0**
	Brassica rapa	16	16	0
Cultivated carrot	Daucus carota subsp. sativus	14	14	0
Chinese rose	**Rosa chinensis**	**14**	**14**	**0**
	Micromonas commoda	13	13	0
Opium poppy	Papaver somniferum	13	13	0
	Chlorella variabilis	12	12	0
Threefork morning glory	**Ipomoea trifida**	**12**	**12**	**0**
Cultivated rice	**Oryza sativa Indica Group**	**12**	**12**	**0**
	Zygnematales	12	11	1
Aiea morning glory	**Ipomoea triloba**	**11**	**11**	**0**
Peanut	**Arachis hypogaea**	**11**	**11**	**0**
Pineapples	**Ananas**	**10**	**10**	**0**
Greater broomrape	Orobanche rapum-genistae	9	9	0
English walnut	Juglans regia	9	9	0
Physic nut	**Jatropha curcas**	**9**	**9**	**0**
Apple tribe	Maleae	9	9	0
	Characiochloris acuminata	**8**	**8**	**0**
	Micromonas pusilla CCMP1545	8	8	0
	Trebouxiophyceae sp. MX-AZ01	**8**	**8**	**0**
Seep monkeyflower	**Erythranthe guttata**	**8**	**8**	**0**
Spider flower	**Tarenaya hassleriana**	**8**	**8**	**0**
	Panicum hallii	**8**	**8**	**0**
	Auxenochlorella protothecoides	7	7	0
	Mesostigma viride	7	7	0
Sunflower	**Helianthus annuus**	**7**	**7**	**0**
Common bean	**Phaseolus vulgaris**	**7**	**7**	**0**
Rubber tree	**Hevea brasiliensis**	**7**	**7**	**0**
	Camelineae	7	7	0
	Eutrema	**7**	**7**	**0**
Date palm	**Phoenix dactylifera**	**7**	**7**	**0**
Cultivated rice	**Oryza sativa Japonica Group**	**7**	**7**	**0**
	Bathycoccaceae	6	6	0
Bracken	Pteridium aquilinum	6	6	0
Quinoa	**Chenopodium quinoa**	**6**	**6**	**0**
Cassava	**Manihot esculenta**	**6**	**6**	**0**
Willow tribe	Saliceae	6	6	0
Peach	**Prunus persica**	**6**	**6**	**0**
Thale cresses	Arabidopsis	6	6	0
Pineapple	**Ananas comosus**	**6**	**6**	**0**
	Chlamydomonadaceae	5	5	0
	Volvox carteri f. nagariensis	**5**	**5**	**0**
Spreading earth moss	Physcomitrella patens	5	5	0
Cultivated beet	Beta vulgaris subsp. vulgaris	5	5	0
White carob	**Prosopis alba**	**5**	**5**	**0**
Mung bean	**Vigna radiata var. radiata**	**5**	**5**	**0**
Wild strawberry	Fragaria vesca subsp. vesca	5	5	0
Papaya	**Carica papaya**	**5**	**5**	**0**
Durian	**Durio zibethinus**	**5**	**5**	**0**
Asparagus	Asparagus officinalis	5	5	0
Foxtail millet	**Setaria italica**	**5**	**5**	**0**
	Ulvophyceae	4	4	0
	Bryopsidales	4	4	0

Continued on next page

Table 4.2 continued

Common name	Taxon	Total	MEGAN	PIA
Polypod ferns	Polypodiales	4	4	0
Goosefoot subfamily	Chenopodioideae	4	4	0
Cacao subfamily	**Byttnerioideae**	**4**	**4**	**0**
	Chlamydomonadales	3	3	0
	Mamiellaceae	3	3	0
	Caryophyllales	3	3	0
Amaranth family	Amaranthaceae	3	3	0
Spurge family	Euphorbiaceae	3	3	0
Buttercup order	Ranunculales	3	3	0
	Zygnemophyceae	3	3	0
	Chlorellaceae	2	2	0
	Streptophytina	2	2	0
	Bryopsida	2	2	0
	Glycine	**2**	**2**	**0**
	Mamiellophyceae	1	1	0
Ferns	Polypodiopsida	1	1	0
Carrots	Daucus	1	1	0
	Crotonoideae	**1**	**1**	**0**
Rose subfamily	Rosoideae	1	1	0
Alismatids	Alismatales	1	1	0
Bluethreads	**Burmannia**	**1**	**0**	**1**

A significant proportion of taxa in table 4.2 are not native to Europe (bold), so are assumed to be false positives. This includes eight of the most frequent taxa. Table 4.3 summarises the frequencies of assignments to European and non-European taxa.

Table 4.3: Numbers of Viridiplantae reads assigned by MEGAN and PIA to European and non-European taxa in ELF054_356. Data had been filtered by negative controls and replicates combined.

	MEGAN	PIA
Total European	352	1
Total non-European	667	1
Total	1019	2
% European	34.54	50.00

In Chapter 3 (PIA), MEGAN assigned approximately 87% of reads from the test samples to European taxa. Performance in ELF054_356 was considerably worse at only 35%. This suggests that the reads in this sample were particularly difficult to assign. Notice that many taxa, including the ten highest-frequency, are cultivated plants or close wild relatives, which are highly likely to be over-represented in GenBank. PIA is designed to avoid assigning reads to over-represented taxa. For ELF054_356, PIA discarded most of the data, but MEGAN attempted to assign and did so poorly. This is expected behaviour; ELF054_356 is simply a particularly visible example of the differences between PIA and MEGAN. Overall, it appears that PIA had the expected effect of producing fewer but more reliable assignments.

4.2 Introduction to core summaries

The bulk of this chapter reports the taxonomic results for the 27 cores from ELF021-60 that were analysed for sedaDNA. These are biogenomic masses for Embryophyta and Metazoa (excluding Primates) after filtering by negative controls. Non-European taxa are assumed to be false positives and have mostly been excluded; later MetaDamage analysis in Chapter 5 (Authentication results) confirmed that these reads have most likely been mis-assigned. However, because some statistical measures count cumulatively, biogenomic masses for non-European taxa can count towards European parent taxa and the sample total. Therefore, the totals row in the data tables of this chapter includes any non-European taxa and may exceed the sum of visible biogenomic masses, although as non-European taxa are rare, the difference is typically very small.

The cores are discussed in turn. The first two cores, ELF022 and ELF027, were taken far north of the Southern River transect. The remaining cores then follow in order from the head of the river system to the mouth, starting with ELF059 in the northwest. If replicate cores were taken at the same location, the core labelled 'A' (such as ELF059A) is discussed after its sister core.

The first subsection for each core reports the total Embryophyta and Metazoa (excluding Primates) read counts after filtering by negative controls. These read counts are used to estimate data yield, which is not possible with biogenomic mass. A low read count suggests a less representative sample of the environment, so interpretation of such samples requires caution. Note that, as for the totals row of the data tables in this chapter, these plots may include reads assigned to non-European taxa. The effect is usually minimal, but a small number of samples are plotted with reads despite being later described as empty. The additional reads had been assigned to non-European taxa.

The second and largest subsection shows biogenomic mass by taxon. These are summarised first as Krona charts (Ondov *et al.* 2011), a form of nested pie chart showing only the most frequent taxa. Figure 4.3 shows an annotated example. Taxa are coloured by ecological category (see key in figure 4.4), and those native to Europe but not Great Britain are highlighted in bold. The outmost circle contains taxa to which reads were directly assigned. Parent taxa are nested inside, showing how the taxa in a sample are related. It can be helpful, for instance, to know when reads are assigned to both a genus and its family. This suggests that reads assigned to the family may have originated from the genus, so they could be considered together when interpreting. Note that redundant parent taxa, containing the same number of reads as the child, are usually collapsed to save space. This often includes the highest taxa. Many charts therefore do not start with Embryophyta or Metazoa, but a slightly lower taxon. Common starting taxa are listed here in descending order:

Plants:	Animals:
▪ Embryophyta	▪ Metazoa
▪ Tracheophyta	▪ Eumetazoa
▪ Euphyllophyta	▪ Bilateria

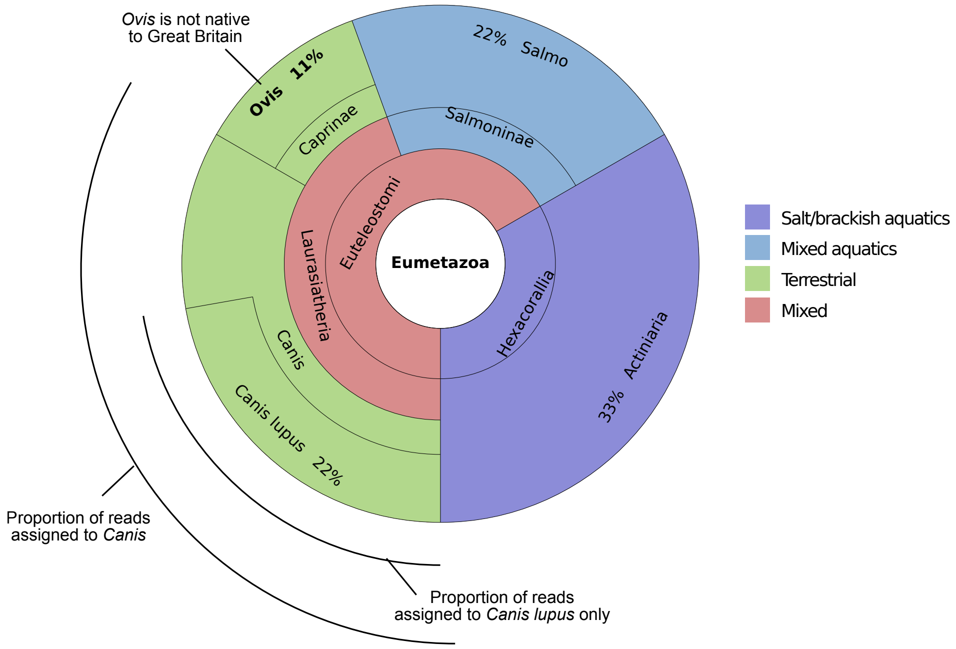


Figure 4.3: Example Krona chart built from the fictitious Metazoa data in table 4.4.

Viridiplantae		Metazoa
Freshwater aquatics	🟦	Freshwater aquatics
Salt/brackish aquatics	🟪	Salt/brackish aquatics
Mixed aquatics	🟦	Mixed aquatics
Halophytes	🟩	
Xerophytes	🟪	
Trees/shrubs	🟩	
Herbs	🟩	Terrestrial
Trees/shrubs and herbs	🟫	
Grasses and relatives	🟨	
Ferns	🟪	
Bryophytes	🟪	
Mixed	🟥	Mixed

Figure 4.4: Key to ecological categories.

Complementing the Krona charts are the tables from which they are derived. For example, table 4.4 corresponds to figure 4.3. These list all taxa in a core by ecological category and taxonomy (A-Z). The Metazoa tables have an additional column, "Group", containing broad taxonomic group, as the Metazoa ecological categories are not taxonomically informative. As in the Krona charts, taxa in bold are not native to Great Britain.

Note that interpretation of biogenomic masses is most valid when comparing the same taxon across samples. Biogenomic mass attempts to account for genome size,

Table 4.4: Example taxonomic results table used to generate figure 4.3. Taxa not native to Great Britain are in bold.

Ecological category	Group	Common name	Taxon	Test_sample
			Ingroup	9
Salt/brackish aquatics	Sea anemones	Sea anemones	Actiniaria	3
Mixed aquatics	Fishes	European salmon and trout	Salmo	2
Terrestrial	Mammals	Grey wolf	Canis lupus	2
Terrestrial	Mammals	Dogs, wolves and jackals	Canis	1
Terrestrial	**Mammals**	**Sheep genus**	**Ovis**	**1**

but it does not address the other major factor affecting data recovery besides biomass: database representation. Most sequences go un- or under-assigned because appropriate references are missing from the database. Species represented by a few genes will attract fewer matches than those represented by a full genome. This makes it difficult to compare biogenomic mass between taxa. Comparing key taxa across samples is more straightforward.

The penultimate section for each core is a pair of heatmaps that illustrate the Pianka similarity index (Pianka 1973) between adjacent samples for Embryophyta and Metazoa (excluding Primates). Originally developed for measuring niche overlap between individual organisms, the index compares the proportional use of a list of resources (here the biogenomic masses of taxa) between subjects (here samples). An index of 0 indicates no overlap, and 1 complete overlap. Samples with more overlap are inferred to have a more similar ecology. These Pianka scores are calculated from cumulative biogenomic masses, so that each taxon includes masses from itself and any child taxa in the data. This allows valid comparison of higher taxa, but very high taxa will proportionally count for almost all of both samples. This inflates Pianka scores, making samples appear more similar. Nevertheless, Pianka scores should be sufficient to highlight particularly different pairs of samples in this dataset.

Finally, a short summary section draws together the sedaDNA results for the core. After each core has been discussed, the final part of this chapter draws the individual results together to search for higher-level patterns.

4.3 ELF022

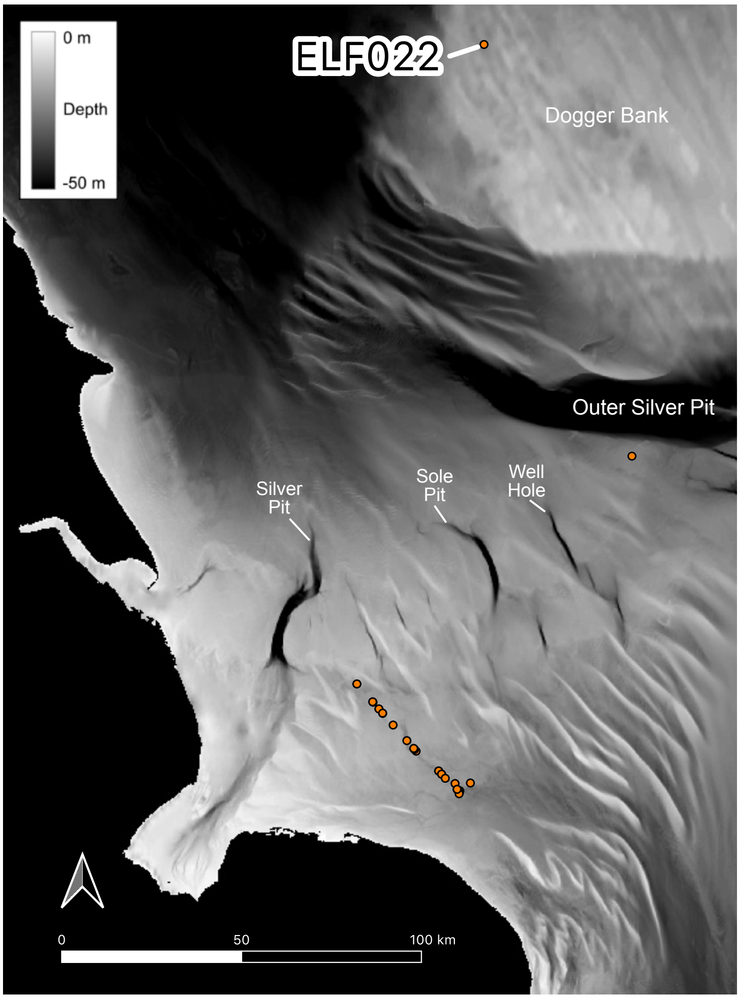


Figure 4.5: Partial bathymetry map of the North Sea highlighting ELF022. Darker grey indicates greater depth. Modern land surfaces are shown in black.

4.3.1 Read counts

Despite ELF022 being one of the longest cores, only a single sample was taken, and only five reads passed filtering. Results are therefore limited.

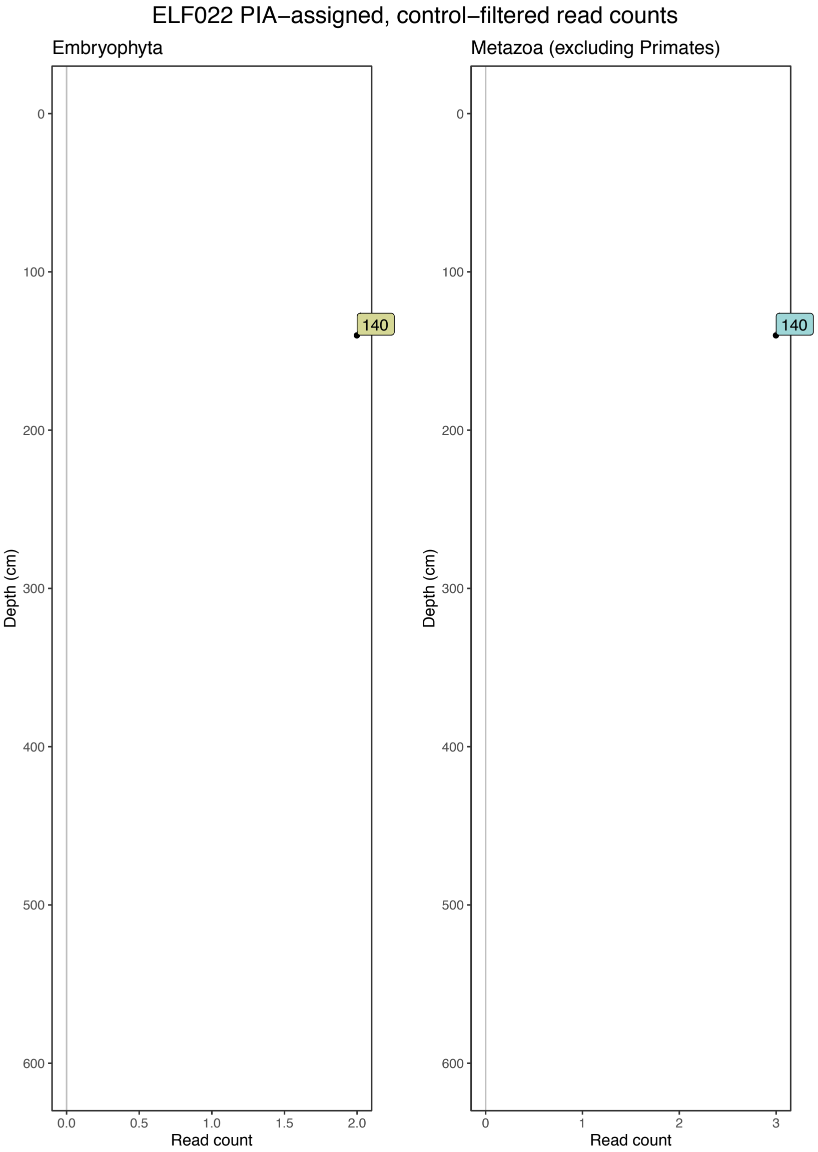


Figure 4.6: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF022. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.3.2 Taxonomic profiles

ELF022_140

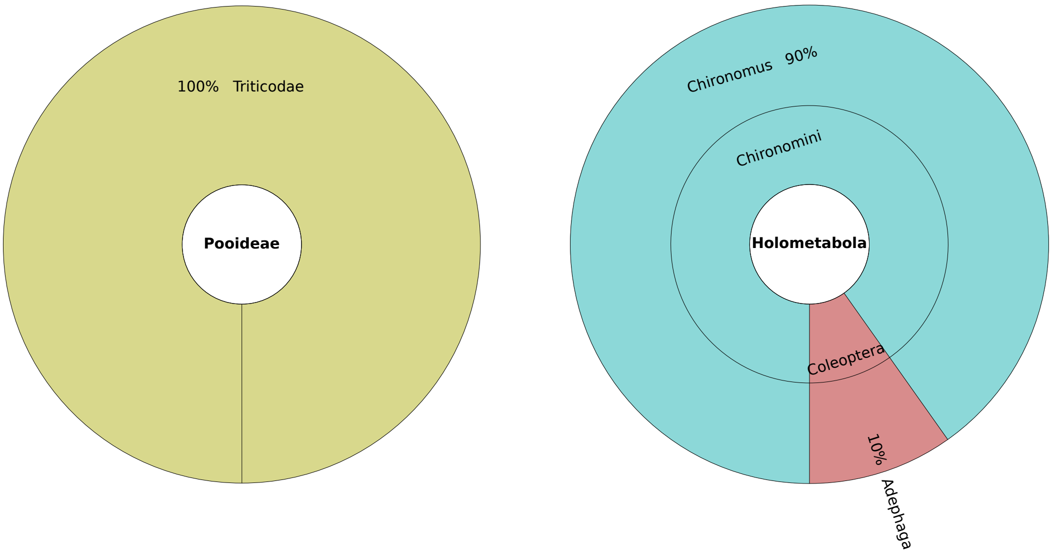


Figure 4.7: Krona charts showing the most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF022 by biogenomic mass. See key in figure 4.4.

Table 4.5: Biogenomic masses of European taxa in Embryophyta from ELF022. Samples are in cm.

Ecological category	Common name	Taxon	140
		Ingroup	0.38
Grasses and relatives		Triticodae	0.11

Table 4.6: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF022. Samples are in cm.

Ecological category	Group	Common name	Taxon	140
			Ingroup	12.26
Freshwater aquatics	Dipterans		Chironomus	11.06
Mixed	Coleopterans		Adephaga	1.20

The single sample, 140 cm, has only three ingroup taxa. The single Embryophyta taxon is Triticodae, a small group of terrestrial grasses which includes the heavily-studied wheat and barley, so should be treated with caution. The two Metazoa taxa are metamorphosing insects (Holometabola): most biogenomic mass is assigned to the midge *Chironomus*, which has freshwater larvae, whereas Adephaga contains freshwater and terrestrial taxa.

4.3.3 Pianka scores

Not applicable as ELF022 has only one sample.

4.3.4 Summary

ELF022 has very limited data. The only plant signal in the single sample is from grasses, whereas the animals are consistent with freshwater but could be terrestrial.

4.4 ELF027

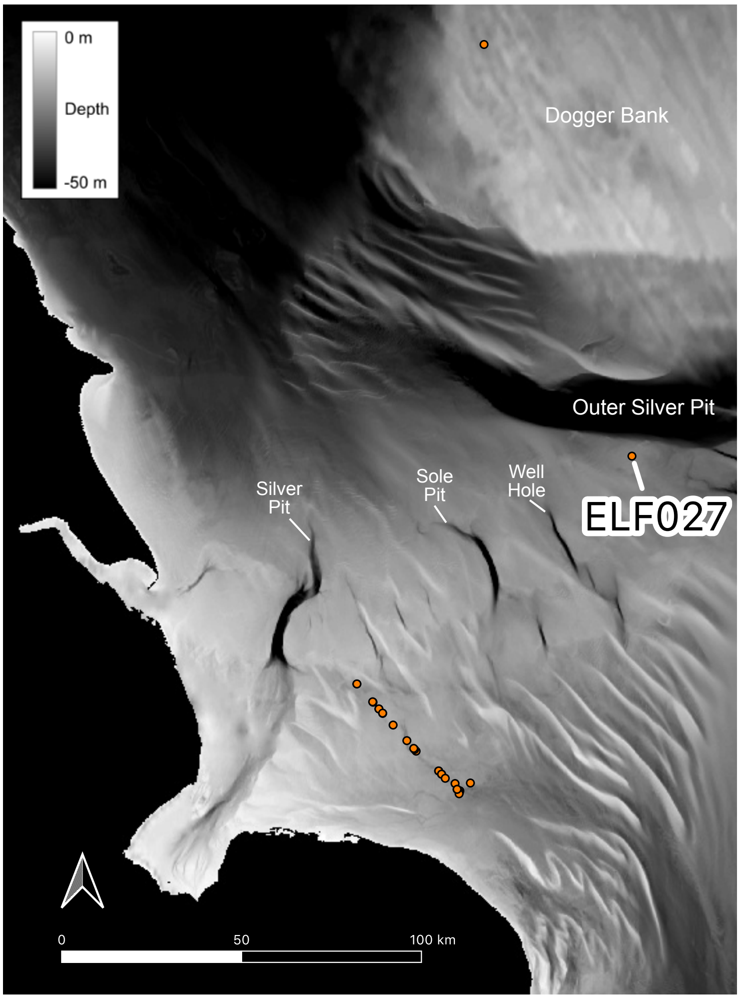


Figure 4.8: Partial bathymetry map of the North Sea highlighting ELF027. Darker grey indicates greater depth. Modern land surfaces are shown in black.

4.4.1 Read counts

Be mindful of the extreme variation in Embryophyta read count when reading figure 4.9: the middle sample, 174 cm, has one of the highest counts across all cores, yet 160 also has a reasonable count at 764. The deepest sample, 177, has a low count for Embryophyta at 64, but not quite zero. The Metazoa counts show a similar spike at much smaller scale: 177 is empty, 160 medium, and 174 particularly high, although with only 22 reads. Almost every sample across all cores has far more data for Embryophyta than Metazoa, perhaps because producers typically have a higher biomass than consumers.

Figure 4.9: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF027. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.4.2 Taxonomic profiles

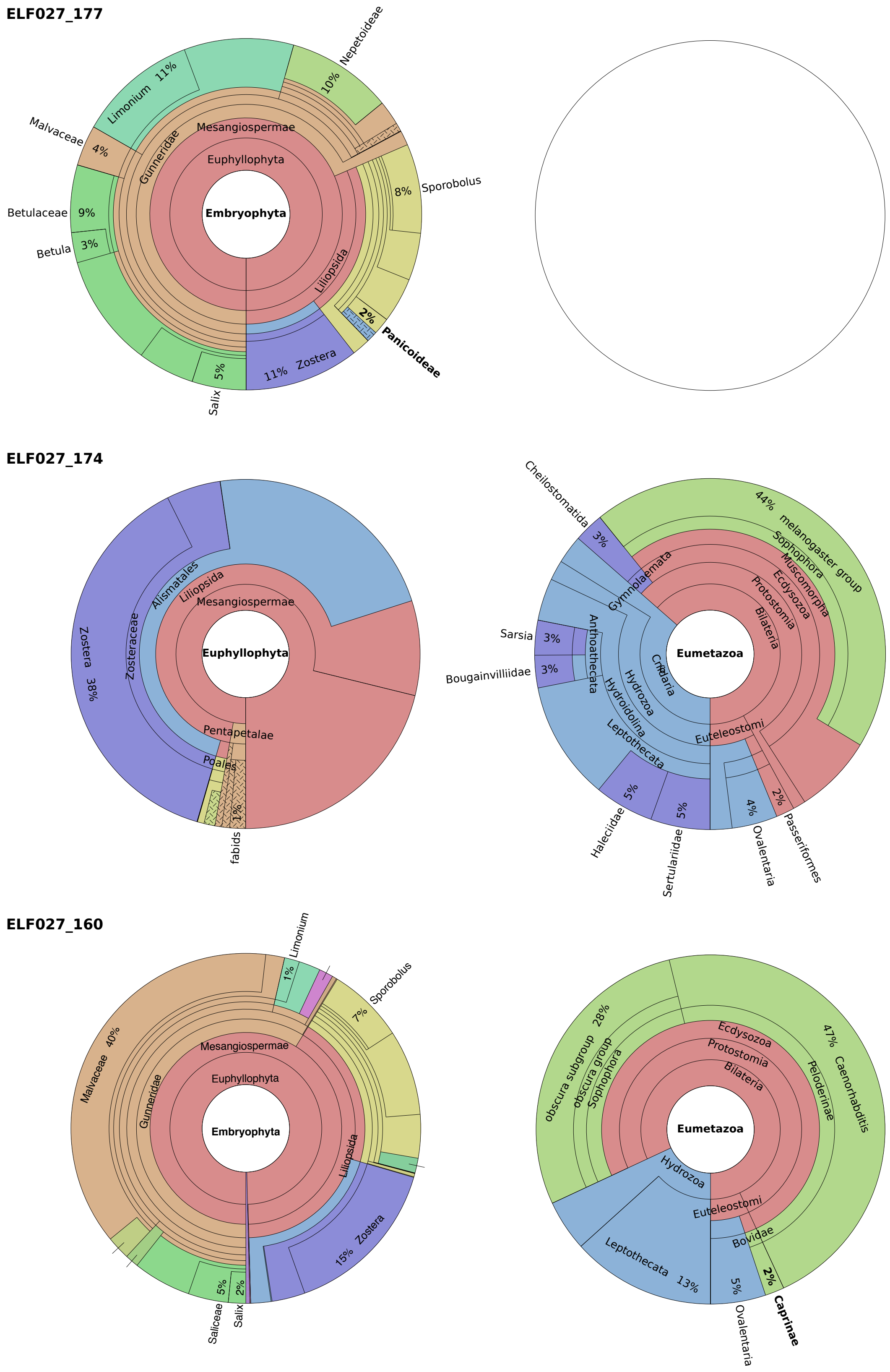


Figure 4.10: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF027 by biogenomic mass. Taxa in bold are not native to Great Britain. See key in figure 4.4.

Table 4.7: Biogenomic masses of European taxa in Embryophyta from ELF027. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	177	174	160
		Ingroup	29989.84	34.83	29383.84	571.17
Freshwater aquatics	Duckweed family	Lemnoideae	7.32	0.00	7.32	0.00
Freshwater aquatics	Pondweeds	Potamogeton	0.54	0.00	0.54	0.00
Salt/brackish aquatics	Eelgrass family	Zosteraceae	1486.24	0.00	1469.72	16.51
Salt/brackish aquatics	Eelgrasses	Zostera	11256.88	3.67	11172.48	80.73
Mixed aquatics	Alismatids	Alismatales	6570.24	0.00	6560.23	10.02
Mixed aquatics	Pondweed family	Potamogetonaceae	0.50	0.00	0.50	0.00
Mixed aquatics	Pondweeds	Stuckenia	0.50	0.00	0.00	0.50
Mixed aquatics	Reeds	Arundinoideae	3.89	0.00	3.54	0.35
Mixed aquatics	Common reed tribe	Molinieae	23.97	0.35	20.09	3.53
Halophytes	Orache	Atriplex	0.69	0.00	0.69	0.00
Halophytes	Goosefoots	Chenopodium	3.20	0.00	3.20	0.00
Halophytes	Glassworts	Salicornia	0.46	0.00	0.46	0.00
Halophytes		Salicornia subg. Salicornia	10.14	0.00	8.70	1.45
Halophytes	Leadwort family	Plumbaginaceae	34.50	3.51	20.25	10.74
Halophytes	Sea-lavenders	Limonium	29.84	3.84	18.54	7.46
Halophytes	Saltmarsh grass	Puccinellia	0.56	0.00	0.28	0.28
Trees/shrubs		Viburnum	1.05	0.00	1.05	0.00
Trees/shrubs	Birch family	Betulaceae	8.64	2.16	3.24	3.24
Trees/shrubs	Alders	Alnus	4.84	0.00	3.23	1.61
Trees/shrubs	Birches	Betula	3.88	0.97	2.91	0.00
Trees/shrubs	Oaks	Quercus	1.06	0.00	1.06	0.00
Trees/shrubs	Willow family	Salicaceae	231.05	3.61	198.56	28.88
Trees/shrubs	Willow tribe	Saliceae	114.37	1.82	92.59	19.97
Trees/shrubs	Poplars	Populus	33.63	0.00	33.63	0.00
Trees/shrubs	Willows	Salix	60.55	1.73	50.17	8.65
Trees/shrubs		Amygdaleae	2.21	0.00	2.21	0.00
Trees/shrubs	Stone fruit trees	Prunus	42.00	0.00	40.18	1.83
Trees/shrubs	Apples	Malus	0.99	0.00	0.99	0.00
Trees/shrubs	Lime subfamily	Tilioideae	0.80	0.00	0.80	0.00
Trees/shrubs	Limes	Tilia	12.80	0.00	7.20	5.60
Herbs	Chicory tribe	Cichorieae	0.28	0.00	0.28	0.00
Herbs	Gentian tribe	Gentianeae	0.28	0.00	0.28	0.00
Herbs	Dead-nettle subfamily	Lamioideae	0.63	0.00	0.63	0.00
Herbs	Catmint subfamily	Nepetoideae	4.52	3.39	1.13	0.00
Herbs	Mint tribe	Mentheae	2.46	0.00	1.23	1.23
Herbs	Plantains	Plantago	5.96	0.00	5.96	0.00
Herbs	Goosefoot family	Chenopodiaceae	31.84	0.00	26.27	5.57
Herbs	Spurge family	Euphorbiaceae	0.31	0.00	0.31	0.00
Herbs		Camelineae	3.14	0.00	3.14	0.00
Herbs	Asparagus family	Asparagaceae	0.10	0.00	0.10	0.00
Herbs	Lady's slipper orchids	Cypripedioideae	0.04	0.00	0.04	0.00
Herbs		Epidendroideae	0.32	0.00	0.32	0.00
Herbs		Malaxideae	0.53	0.00	0.53	0.00
Herbs	Lily family	Liliaceae	0.03	0.00	0.03	0.00
Herbs	Poppy subfamily	Papaveroideae	0.34	0.00	0.34	0.00
Trees/shrubs and herbs		Gunneridae	0.94	0.47	0.00	0.47
Trees/shrubs and herbs		Apiineae	0.93	0.00	0.93	0.00
Trees/shrubs and herbs	Ivy family	Araliaceae	0.50	0.00	0.50	0.00
Trees/shrubs and herbs		Asterales	0.57	0.00	0.57	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	2.17	0.27	1.90	0.00
Trees/shrubs and herbs		Asteroideae	2.76	0.00	1.75	1.00
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	0.55	0.00	0.55	0.00
Trees/shrubs and herbs	Aster tribe	Astereae	6.22	0.00	6.22	0.00
Trees/shrubs and herbs		Gentianales	0.90	0.00	0.90	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	0.88	0.00	0.88	0.00
Trees/shrubs and herbs	Dead-nettle family	Lamiaceae	0.78	0.78	0.00	0.00
Trees/shrubs and herbs		Stachydeae	0.72	0.00	0.72	0.00
Trees/shrubs and herbs	Plantain family	Plantaginaceae	1.84	0.00	1.84	0.00
Trees/shrubs and herbs		Solanales	0.38	0.00	0.38	0.00
Trees/shrubs and herbs		Caryophyllales	76.40	0.00	76.40	0.00
Trees/shrubs and herbs	Goosefoot subfamily	Chenopodioideae	2.45	0.00	2.45	0.00
Trees/shrubs and herbs	Knotweed family	Polygonaceae	0.55	0.00	0.55	0.00
Trees/shrubs and herbs	Mustard order	Brassicales	1.30	0.00	1.30	0.00
Trees/shrubs and herbs	Mallow order	Malvales	16.96	0.00	7.60	9.35
Trees/shrubs and herbs	Mallow family	Malvaceae	394.15	1.31	189.90	202.95
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	8.88	0.00	2.78	6.11
Trees/shrubs and herbs	Mistletoe family	Viscaceae	0.01	0.01	0.00	0.00
Grasses and relatives	Grass order	Poales	2.05	0.59	1.17	0.29
Grasses and relatives		Cyperoideae	5.06	0.00	3.37	1.69
Grasses and relatives	True sedges	Carex	4.56	0.00	4.56	0.00
Grasses and relatives	Rushes	Juncaceae	1.04	0.00	1.04	0.00
Grasses and relatives		Coleanthinae	0.42	0.00	0.42	0.00
Grasses and relatives		PACMAD clade	180.58	0.00	180.58	0.00
Grasses and relatives		Chloridoideae	70.16	1.44	47.09	21.62
Grasses and relatives	Zoysia tribe	Zoysieae	150.00	1.52	106.06	42.42
Grasses and relatives	Dropseed	Sporobolus	150.48	2.86	109.52	38.10
Grasses and relatives		**Panicoideae**	**1.62**	**0.54**	**0.54**	**0.54**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**9.58**	**0.00**	**7.99**	**1.60**
Grasses and relatives		**Paniceae**	**4.02**	**0.00**	**2.68**	**1.34**
Ferns	Ferns	Polypodiopsida	0.14	0.00	0.14	0.00
Ferns	Horsetails	Equisetum	0.18	0.00	0.18	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	0.49	0.00	0.33	0.16
Ferns	Polypod ferns	Polypodiales	0.24	0.00	0.16	0.08
Ferns	Spleenwort suborder	Aspleniineae	0.36	0.00	0.27	0.09
Ferns		Thelypteridaceae	0.10	0.00	0.10	0.00

Continued on next page

Table 4.7 continued

Ecological category	Common name	Taxon	Total	177	174	160
		Ingroup	29989.84	34.83	29383.84	571.17
Ferns	Buckler and male ferns	Dryopteris	0.15	0.00	0.00	0.15
Ferns		**Pteridoideae**	**0.12**	**0.00**	**0.12**	**0.00**
Bryophytes		Bryidae	1.93	0.00	0.00	1.93
Mixed		Tracheophyta	11.07	0.00	11.07	0.00
Mixed		Euphyllophyta	0.34	0.00	0.34	0.00
Mixed	Core angiosperms	Mesangiospermae	6199.00	0.00	6199.00	0.00
Mixed	Monocots	Liliopsida	2569.84	0.00	2569.84	0.00

Table 4.8: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF027. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	177	174	160
			Ingroup	63.10	0.00	44.80	18.30
Salt/brackish aquatics	Bryozoans		Cheilostomatida	1.22	0.00	1.22	0.00
Salt/brackish aquatics	Hydrozoans		Sarsia	1.43	0.00	1.43	0.00
Salt/brackish aquatics	Hydrozoans		Bougainvilliidae	1.33	0.00	1.33	0.00
Salt/brackish aquatics	Hydrozoans		Haleciidae	2.44	0.00	2.44	0.00
Salt/brackish aquatics	Hydrozoans		Sertulariidae	2.44	0.00	2.44	0.00
Mixed aquatics	Fishes		Euteleosteomorpha	0.90	0.00	0.90	0.00
Mixed aquatics	Fishes		Ovalentaria	2.80	0.00	1.87	0.93
Mixed aquatics	Cnidarians	Cnidarians	Cnidaria	1.17	0.00	1.17	0.00
Mixed aquatics	Hydrozoans	Hydrozoans	Hydrozoa	0.82	0.00	0.82	0.00
Mixed aquatics	Hydrozoans		Hydroidolina	2.60	0.00	1.73	0.87
Mixed aquatics	Hydrozoans	Thecate hydroids	Leptothecata	7.32	0.00	4.88	2.44
Terrestrial	**Mammals**	**Goat subfamily**	**Caprinae**	**0.33**	**0.00**	**0.00**	**0.33**
Terrestrial	Nematodes		Caenorhabditis	8.59	0.00	0.00	8.59
Terrestrial	Dipterans		melanogaster group	19.81	0.00	19.81	0.00
Terrestrial	Dipterans		obscura subgroup	5.15	0.00	0.00	5.15
Mixed	Animals	Animals	Metazoa	0.25	0.00	0.25	0.00
Mixed	Birds	Perching birds	Passeriformes	0.76	0.00	0.76	0.00
Mixed	Invertebrates		Protostomia	0.52	0.00	0.52	0.00
Mixed	Dipterans		Muscomorpha	3.24	0.00	3.24	0.00

Embryophyta

Aquatics

All three samples contain aquatics. Only 174 contains freshwater taxa, but samples with more data would be expected to have greater taxonomic richness, so this may not necessarily mean an absence of freshwater taxa elsewhere. However, the marine *Zostera* (eelgrasses) and its family Zosteraceae are present throughout the core. The only other genus in Zosteraceae is the Pacific *Phyllospadix* (Short *et al.* 2007), so Zosteraceae will be interpreted as more *Zostera*. *Zostera* occurs in fully submerged marine habitats such as lagoons and estuaries, extending into shallow coastal waters with a maximum depth of around 12 m (Packham and Willis 1997, Short *et al.* 2007).

174 is the first sample discussed that associates *Zostera* with a very high read count. Throughout the cores, the majority of high-yield samples have a dominant ecological category of saltwater aquatics, which are invariably mostly *Zostera*. This pattern could be explained by the beneficial effect of salt on DNA preservation (Lindahl and Nyberg 1972, Kistler *et al.* 2017).

Every sample in ELF027 also contains what will be referred to as the reed group: the mixed aquatic Arundinoideae (reeds) or its child taxon Molinieae (common reed tribe), which are found in most cores along with *Phragmites* (common reeds). The only European *Phragmites* is *P. australis*, a cosmopolitan reed of brackish water often found in saltmarsh (Packham and Willis 1997). Many cores contain both *Zostera* and the reed group, suggesting an interface between open water and more terrestrial coastal environments.

Halophytes and xerophytes

Supporting the salt signal are a variety of halophytes, which are generally rare across the cores. They are present here in all three samples and make up a significant proportion of the deepest, 177, despite it returning only a small amount of data.

There is some overlap between halophytes and xerophytes because high-salt and arid habitats require similar adaptations. A significant xerophyte in ELF027 is Chenopodiaceae (the goosefoot family) in samples 174 and 160, which is used as an aridity indicator in pollen studies (Lu *et al.* 2019) despite containing halophytes and some plants adapted to less extreme environments. In this core, halophytes are unusually common and sample 174 also contains *Chenopodium*, so Chenopodiaceae could be interpreted as more of a coastal indicator than of aridity.

Woody taxa

The trees/shrubs ecological category also has a clear presence throughout the core, despite being relatively infrequent in sample 174. Most prominent is what will be referred to as the willow group: *Salix* (willows), *Populus* (poplars), and their parent taxa Saliceae and Salicaceae. As was typical for most cores, *Salix* is more frequent than *Populus*. Willow group genera are predominantly associated with mire or floodplain forests in Europe, as are the the *Betula* (birches) and *Alnus* (alders) found in multiple samples in this core (Pividori *et al.* 2016).

Prunus (stone fruit trees) also has a particularly high biogenomic mass in this and many other cores. The four European *Prunus* species prefer relatively open areas, such as scrub or woodland clearings, or are pioneers (San-Miguel-Ayanz *et al.* 2016). However, *Prunus* contains several well-studied cultivated species. Despite PIA attempting to account for over-represented taxa in reference databases, it is possible that some reads have been over-assigned. Indeed, *Prunus* is not dominant in any European forest types, so a high biomass would be surprising (Pividori *et al.* 2016).

Another notable tree is *Tilia* (limes) and its subfamily Tilioideae in samples 174 and 160. *Tilia* is a lowland tree that requires relatively warm temperatures to set seed (Eaton *et al.* 2016) and, according to pollen data, colonised Britain rather late after the last glaciation (∼8,000 BP, Brewer *et al.* 2017). *Tilia* may be an interesting link between date and climate and is examined further in Chapter 6 (Mesophilic taxa and human disturbance indicators).

Trees/shrubs and herbs

The category 'trees/shrubs and herbs' is not often useful for palaeoenvironmental reconstruction because of the diversity of its members. However, there are several in ELF027 that can be interpreted alongside more informative taxa. Chenopodioideae (the goosefoot subfamily) in 177 is likely linked to *Chenopodium*. Malvales, Malvaceae, and Malvoideae are the mallow order, family, and subfamily respectively. The three members native to Great Britain are the shrub *Daphne* (family Thymelaeaceae), herb *Malva* (mallow), and tree *Tilia* (lime; subfamily Tilioideae). *Malva* and *Tilia* are present in these cores. There is a clear *Tilia* signal in ELF027, so the parent taxa could be interpreted as such, suggesting that lime trees were present in all samples and possibly the dominant tree in 160.

Finally, sample 177 contains a very small signal from Viscaceae, the mistletoe family. The only British species is *Viscum album*, which despite being a hemiparasite of trees, prefers more open habitats to woodland (Briggs 2011). Viscaceae is absent from the other samples despite 177 having the least data, suggesting a genuine absence that may indicate a more closed habitat further up the core.

Grasses and relatives

All samples contain Poales. The PACMAD clade is particularly frequent in sample 174. This higher taxon contains the reed group, so may be interpreted as more reeds, but also *Sporobolus* and its parent taxa Zoysieae and Chloridoideae, which have clear signals across the core. The only European genus in Zoysieae is *Sporobolus*, so it can be interpreted as such. Unfortunately, the many *Sporobolus* species live in a range of habitats, so it is not very ecologically informative.

The PACMAD clade also contains what will be referred to as the Panicoideae group: Panicoideae itself and its child taxa Andropogoneae and Paniceae, along with occasional lower taxa in other cores. The group is found across ELF027 (and many other cores) at relatively low frequency. Several taxa are native to southern Europe, but none to Great Britain. This could suggest a warmer climate in ELF027 than at its location today. It is possible that these samples were at the surface during a warmer period, such as the Holocene Thermal Maximum around 8,000 BP (Martin *et al.* 2020). However, Panicoideae also appeared in some negative controls, suggesting possible modern contamination, and it contains the heavily-studied maize and sorghum, which may have resulted in reads being mis-assigned. These taxa are over-represented in the sequence database, so are likely to attract BLAST hits from taxa besides themselves. While PIA is designed to discard reads that apparently match over-represented taxa, a moderate number of spurious maize or sorghum hits may be enough to shift the assignment of a read incorrectly to Panicoideae while not flagging it for removal. Arundoideae, the reed subfamily, is a sister taxon to Panicoideae and a more parsimonious taxon for these reads to derive from. Detailed inspection of the Panicoideae group in Chapter 6 (Mesophilic taxa and human disturbance indicators) was inconclusive regarding assignment, but could confirm that the reads appear ancient, so are very unlikely to derive from modern contamination. There is not enough evidence to reject the Panicoideae group at present, so it will be tentatively accepted in this chapter.

Ferns

Samples 174 and 160 have a small fern signal. All of the ferns studied have large genome sizes (appendix D), so even relatively high read counts translate to small biogenomic masses. Ferns are rarely ecologically informative, spanning a wide range of habitats, but sample 174 does contain Pteridoideae, which has no British natives but at least three genera in southern Europe. Again, this could possibly indicate a warmer climate, but the signal is weak.

Mixed

Taxa in the mixed category are usually very high and unhelpful for palaeoecological reconstruction. However, they can constitute a significant proportion of biogenomic mass in some samples, of which sample 174 is typical. The most frequent taxon is *Zostera*. The mixed aquatic Alismatales and higher taxa Mesangiospermae (core angiosperms) and Liliopsida (monocots) are close behind. All are parent taxa of *Zostera*, and the prevalence of *Zostera* compared to other genera in this sample suggests that many of their reads could have derived from *Zostera* but were assigned higher, although they could also derive from related taxa absent from the database. The numbers of reads (in addition to biogenomic mass, which differs by taxon because of genome size) assigned to *Zostera* parent taxa vary considerably, which means that the parent taxa are affected by factors besides the amount of *Zostera*.

Metazoa

The two ELF027 with data contain terrestrial taxa but are also particularly diverse in cnidarians. 174 has the only saltwater taxa and a smaller terrestrial signal than 160, consistent with the more marine Embryophyta signal in 174, but the read counts are too small to infer a significant difference. Furthermore, all terrestrial taxa are closely related to well-studied organisms, so their reads may have been over-assigned. Overall, the Metazoa results are limited.

4.4.3 Pianka scores

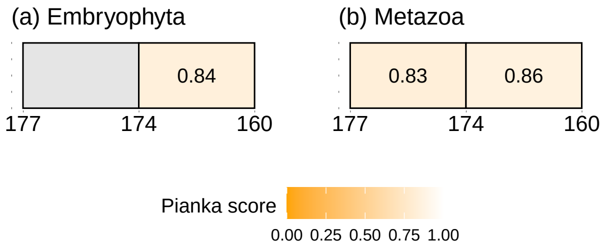


Figure 4.11: Pianka similarity scores between adjacent samples in ELF027 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores. Comparisons involving empty samples have no score.

The Pianka scores for Embryophyta are actually comparatively low for these cores, despite being >0.8. Sample 174 includes several very high taxa, which would inflate the scores, but also many unique taxa that make it appear less similar. Greater taxonomic richness would be expected from more data because it is a larger sample of the DNA pool. This makes it difficult to compare samples with very different read counts. However, it is clear from figure 4.10 that 174 is proportionally dominated by *Zostera* and the others not, so we can infer some community change.

The single Metazoa score (sample 177 has no reads) is of similar size. Not many taxa are present in both samples, but the many interrelated cnidarian taxa result in a relatively high degree of overlap.

4.4.4 Summary

ELF027 is notable for the extreme variation in data yield. The middle sample, 174, has one of the highest Embryophyta read counts across all cores and demonstrates a pattern common to many high-yield samples: the most frequent taxa are *Zostera* followed by its higher parent taxa, although these are not especially informative.

Looking across the core, all three samples have clear saltwater, reed, and halophyte signals, along with terrestrial and floodplain-associated trees. Together, these suggest a coastal environment with some freshwater influence, possibly an estuary. The additional *Zostera* in 174 may represent temporary submergence, perhaps by a wandering channel. *Tilia* may be more frequent than the willow group in 160 if Malvaceae can be interpreted as such, suggesting drier conditions, perhaps as the channel moved further away. However, there is little evidence of long-term ecological change.

4.5 ELF059

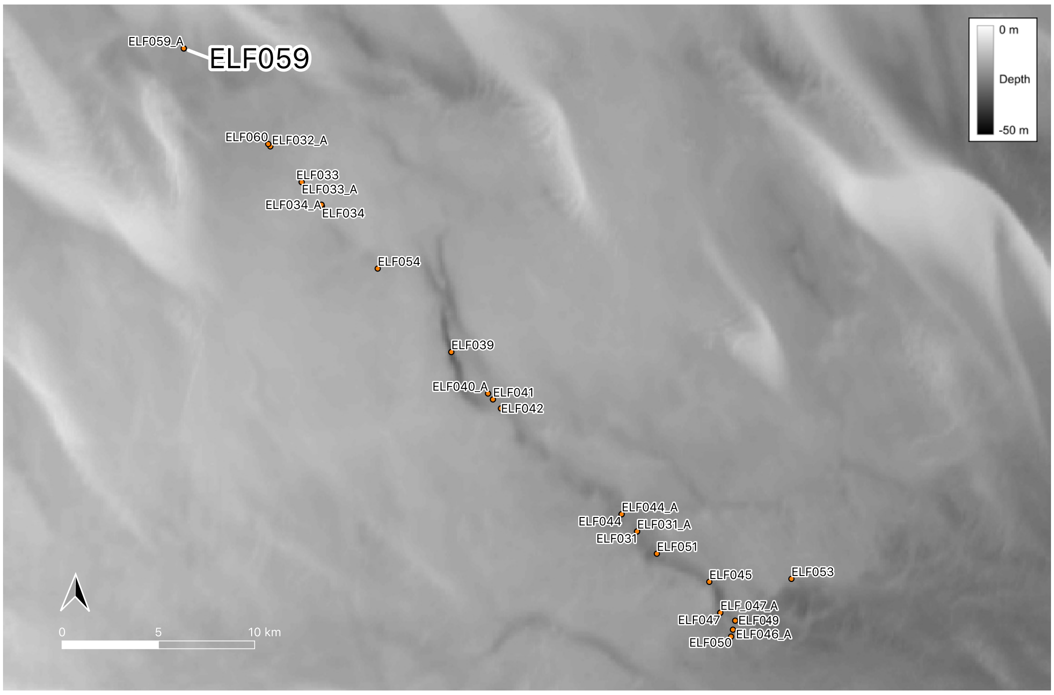


Figure 4.12: Bathymetry map of the palaeochannel transect highlighting ELF059. Darker grey indicates greater depth.

ELF059 is the first core in the palaeochannel transect, although as figure 4.12 shows, the channel only begins approximately half way along. ELF059 comes from higher ground towards the head of the valley.

4.5.1 Read counts

ELF059 also shows variance in Embryophyta read count, though less extreme than ELF027, ranging from around 100 in samples 378 and 280 to over 6,000 in 210. The Metazoa counts are far smaller but proportionally similar to their Embryophyta counterparts. There may be a slight trend of increasing data yield up the core.

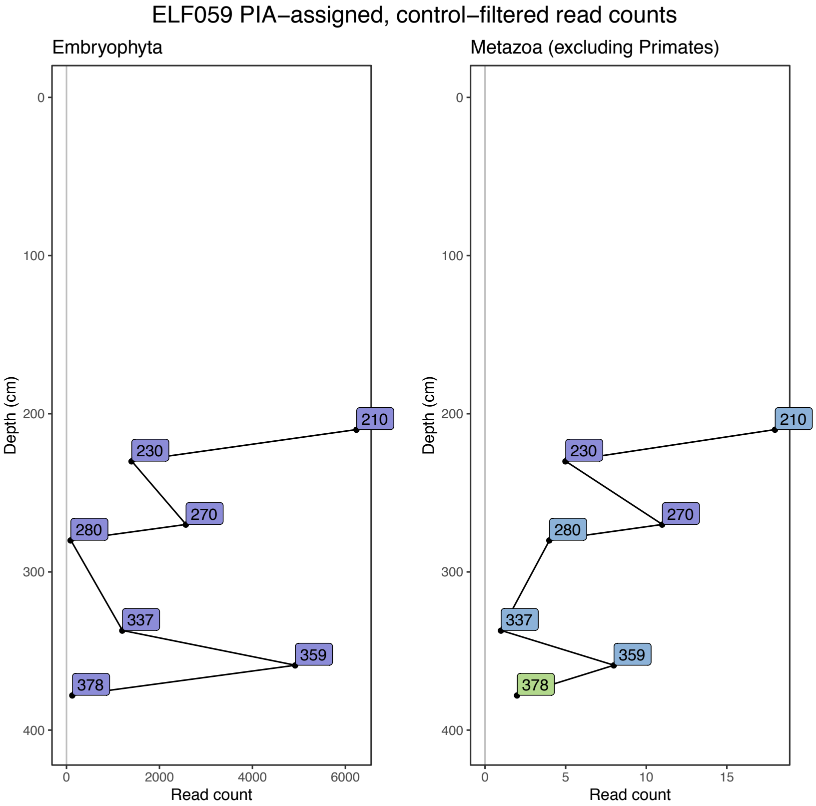
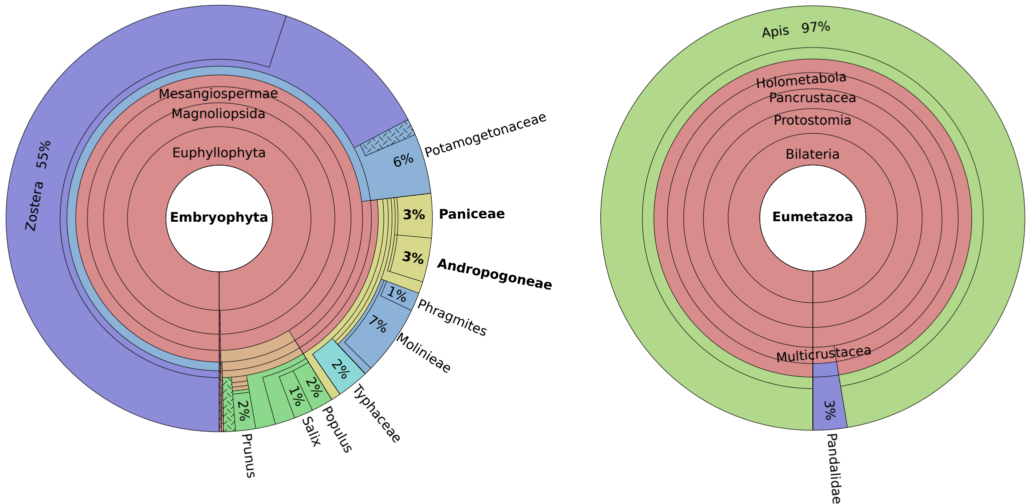


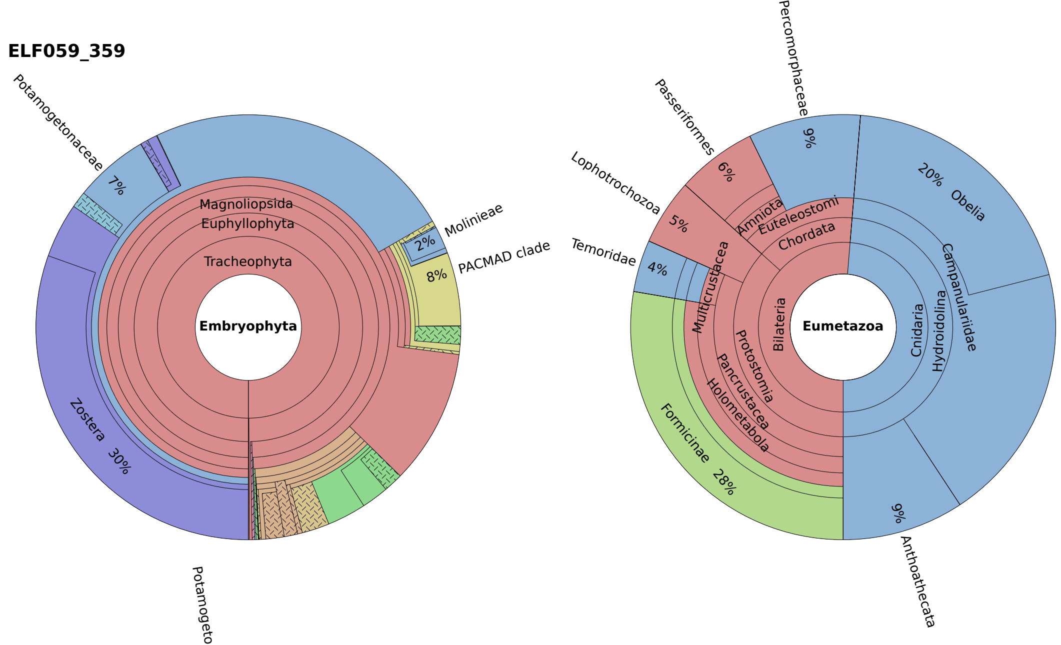
Figure 4.13: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF059. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.5.2 Taxonomic profiles

ELF059_378



ELF059_359



ELF059_337

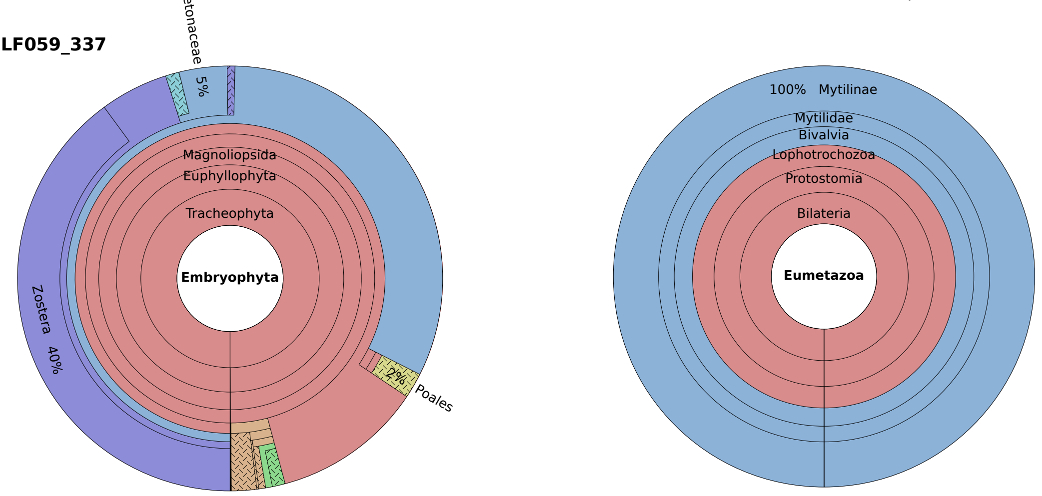


Figure 4.14: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF059, samples 378, 359, and 347. Continued in figure 4.15. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

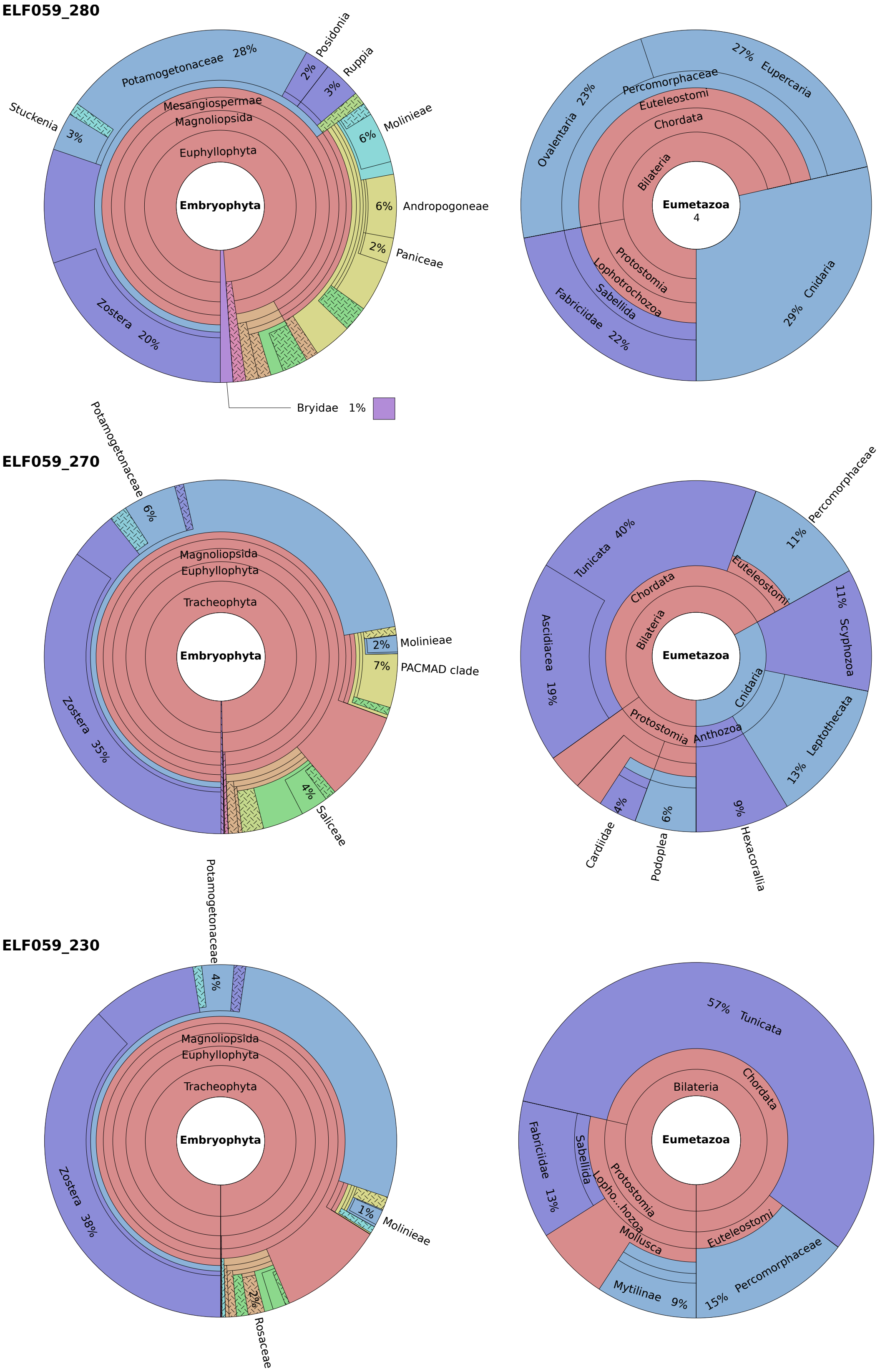


Figure 4.15: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF059, samples 280, 270, and 230. Continued in figure 4.16. See figure 4.4 for colour key.

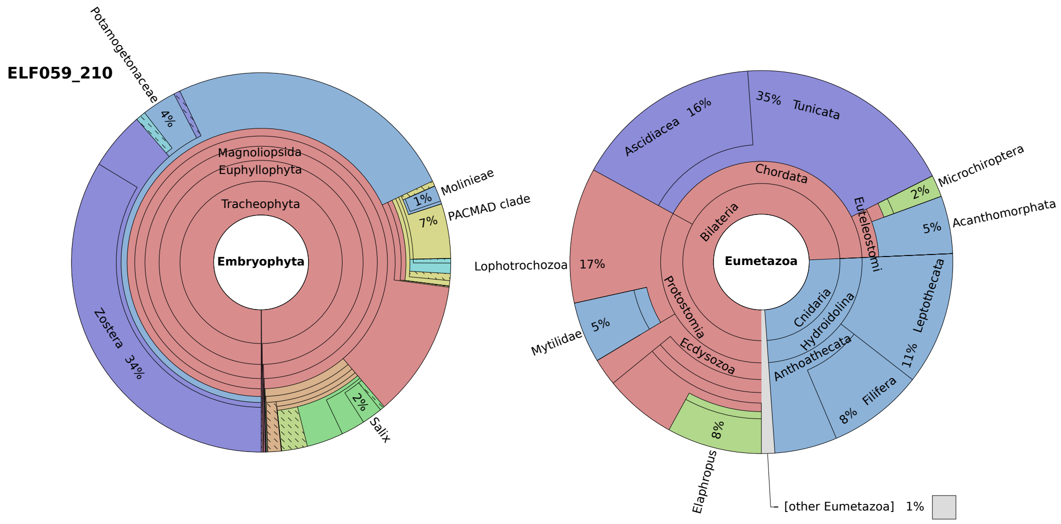


Figure 4.16: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF059, sample 210. See figure 4.4 for colour key.

Table 4.9: Biogenomic masses of European taxa in Embryophyta from ELF059. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	378	359	337	280	270	230	210
		Ingroup	5500.50	131.58	1600.04	355.28	91.01	878.58	459.04	1984.98
Freshwater aquatics	Coontails	Ceratophyllum	4.35	0.00	1.45	0.00	0.00	0.00	1.45	1.45
Freshwater aquatics	Loosestrifes	Lythrum	1.79	0.00	1.79	0.00	0.00	0.00	0.00	0.00
Freshwater aquatics	Water-milfoil family	Haloragaceae	6.92	0.00	4.62	0.00	0.00	2.31	0.00	0.00
Freshwater aquatics	Water-milfoils	Myriophyllum	18.46	0.00	6.92	2.31	0.00	0.00	0.00	9.23
Freshwater aquatics	Water-plantain family	Alismataceae	0.21	0.00	0.14	0.00	0.00	0.00	0.07	0.00
Freshwater aquatics	Arrowheads	Sagittaria	0.06	0.00	0.06	0.00	0.00	0.00	0.00	0.00
Freshwater aquatics	Duckweed family	Lemnoideae	1.22	0.00	0.00	0.00	1.22	0.00	0.00	0.00
Freshwater aquatics	Elodea family	Hydrocharitaceae	1.38	0.00	0.40	0.00	0.00	0.20	0.20	0.59
Freshwater aquatics	Pondweeds	Potamogeton	38.27	0.00	10.24	2.70	0.54	9.16	3.23	12.40
Freshwater aquatics	Bulrush family	Typhaceae	37.86	2.91	5.83	2.91	2.91	2.91	2.91	17.48
Freshwater aquatics	Bulrushes	Typha	11.32	0.00	3.77	0.00	0.00	0.00	0.00	7.55
Freshwater aquatics	Waterlilly family	Nymphaeaceae	2.65	0.00	1.59	0.00	0.00	0.00	0.00	1.06
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	23.69	0.00	11.85	2.09	0.00	3.48	2.09	4.18
Salt/brackish aquatics	Tasselweeds	Ruppia	21.89	0.00	9.01	0.00	1.29	3.00	2.58	6.01
Salt/brackish aquatics	**Neptune-grasses**	**Posidonia**	**0.96**	**0.00**	**0.00**	**0.00**	**0.64**	**0.32**	**0.00**	**0.00**
Salt/brackish aquatics	Eelgrass family	Zosteraceae	295.41	14.68	66.06	18.35	16.51	38.53	44.04	97.25
Salt/brackish aquatics	Eelgrasses	Zostera	1858.72	66.06	482.57	141.28	31.19	302.75	172.48	662.39
Mixed aquatics	Alismatids	Alismatales	1331.67	0.00	376.33	113.56	0.00	222.88	127.14	491.76
Mixed aquatics	Pondweed family	Potamogetonaceae	231.07	5.45	89.75	12.89	9.92	42.15	13.39	57.52
Mixed aquatics	Pondweeds	Stuckenia	23.31	1.49	9.42	0.99	1.49	4.96	0.50	4.46
Mixed aquatics	Reeds	Arundinoideae	14.85	0.71	6.01	0.35	0.35	1.41	1.06	4.95
Mixed aquatics	Common reed tribe	Molinieae	81.79	6.35	26.79	1.41	1.41	13.04	6.35	26.44
Mixed aquatics	Common reeds	Phragmites	5.46	1.68	2.10	0.00	0.42	0.42	0.42	0.42
Halophytes	Sea-blite subfamily	Suaedoideae	0.94	0.00	0.00	0.00	0.00	0.94	0.00	0.00
Halophytes	**Desert thumb**	**Cynomorium**	**0.19**	**0.00**	**0.00**	**0.19**	**0.00**	**0.00**	**0.00**	**0.00**
Trees/shrubs	Elder family	Adoxaceae	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13
Trees/shrubs		Viburnum	1.58	0.00	0.00	0.26	0.00	0.26	0.00	1.05
Trees/shrubs		Cornales	0.61	0.00	0.00	0.00	0.00	0.00	0.00	0.61
Trees/shrubs	Dogwood family	Cornaceae	0.59	0.00	0.59	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Olive family	Oleaceae	0.67	0.00	0.67	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Olive tribe	Oleeae	0.80	0.00	0.00	0.00	0.00	0.00	0.80	0.00
Trees/shrubs	Birch family	Betulaceae	23.76	1.08	1.08	0.00	0.00	6.48	0.00	15.12
Trees/shrubs	Alders	Alnus	9.68	0.00	0.00	0.00	0.00	0.00	4.84	4.84
Trees/shrubs	Birches	Betula	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.97
Trees/shrubs	Hazels	Corylus	6.82	0.00	2.27	0.00	0.00	2.27	0.00	2.27
Trees/shrubs	Oaks	Quercus	2.12	0.00	2.12	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Bayberry family	Myricaceae	1.57	0.00	0.00	0.00	0.00	0.00	0.00	1.57
Trees/shrubs	Willow family	Salicaceae	148.01	1.81	46.93	1.81	0.00	32.49	3.61	61.37
Trees/shrubs	Willow tribe	Saliceae	103.48	1.82	32.68	0.00	1.82	21.79	5.45	39.94
Trees/shrubs	Poplars	Populus	19.78	1.98	5.93	0.00	1.98	1.98	0.00	7.91
Trees/shrubs	Willows	Salix	67.47	1.73	17.30	3.46	1.73	6.92	1.73	34.60
Trees/shrubs		Amygdaleae	2.21	0.00	2.21	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Stone fruit trees	Prunus	40.18	1.83	10.96	1.83	0.00	7.30	5.48	12.78
Trees/shrubs	Apple tribe	Maleae	0.89	0.00	0.89	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Apples	Malus	0.99	0.00	0.00	0.00	0.00	0.99	0.00	0.00
Trees/shrubs	Elm family	Ulmaceae	1.51	0.00	0.00	0.00	0.00	0.00	0.00	1.51

Continued on next page

Table 4.9 continued

Ecological category	Common name	Taxon	Total	378	359	337	280	270	230	210
		Ingroup	5500.50	131.58	1600.04	355.28	91.01	878.58	459.04	1984.98
Trees/shrubs	Limes	Tilia	1.60	0.00	0.00	0.00	0.00	1.60	0.00	0.00
Trees/shrubs	**Plane trees**	**Platanus**	**0.62**	**0.00**	**0.62**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Trees/shrubs	Barberry family	Berberidaceae	0.23	0.00	0.23	0.00	0.00	0.00	0.00	0.00
Herbs	Primrose family	Primulaceae	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.35
Herbs	Borage family	Boraginaceae	4.06	0.00	3.39	0.00	0.00	0.00	0.00	0.68
Herbs		Cynoglossoideae	2.21	0.00	2.21	0.00	0.00	0.00	0.00	0.00
Herbs	Bedstraw family	Rubiaceae	1.01	0.00	1.01	0.00	0.00	0.00	0.00	0.00
Herbs	Bedstraw	Galium	0.73	0.00	0.73	0.00	0.00	0.00	0.00	0.00
Herbs	Catmint subfamily	Nepetoideae	4.52	0.00	2.26	0.00	0.00	0.00	0.00	2.26
Herbs	Mint tribe	Mentheae	6.14	0.00	2.46	2.46	0.00	0.00	0.00	1.23
Herbs	Mints	Mentha	1.52	0.00	1.52	0.00	0.00	0.00	0.00	0.00
Herbs	Goosefoot family	Chenopodiaceae	1.59	0.00	0.80	0.80	0.00	0.00	0.00	0.00
Herbs	Rhubarb family	Rumiceae	0.41	0.00	0.41	0.00	0.00	0.00	0.00	0.00
Herbs		Hologalegina	0.28	0.00	0.28	0.00	0.00	0.00	0.00	0.00
Herbs	Clover tribe	Trifolieae	13.21	0.00	13.21	0.00	0.00	0.00	0.00	0.00
Herbs	Cinquefoils	Potentilla	1.42	0.00	0.00	0.00	0.00	0.00	1.42	0.00
Herbs	Crane's-bills	Geranium	1.14	0.00	0.57	0.00	0.00	0.00	0.57	0.00
Herbs	Arum family	Araceae	0.92	0.00	0.55	0.00	0.00	0.18	0.00	0.18
Herbs		**Agavoideae**	**0.15**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.15**
Herbs	Bluebell family	Hyacinthaceae	0.08	0.00	0.00	0.00	0.00	0.00	0.08	0.00
Herbs	Irises	Iris	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.11
Herbs	Orchids	Orchidaceae	1.19	0.00	1.02	0.00	0.00	0.00	0.00	0.17
Herbs		Epidendroideae	0.32	0.00	0.00	0.00	0.00	0.00	0.00	0.32
Herbs		Dioscoreales	0.93	0.00	0.93	0.00	0.00	0.00	0.00	0.00
Herbs		Dioscorea	1.81	0.00	0.90	0.00	0.00	0.00	0.00	0.90
Herbs	Lily family	Liliaceae	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.15
Herbs	Bunchflower family	Melanthiaceae	0.03	0.00	0.03	0.00	0.00	0.00	0.00	0.00
Herbs	Buttercup subfamily	Ranunculoideae	0.43	0.00	0.11	0.00	0.00	0.11	0.00	0.22
Trees/shrubs and herbs		Gunneridae	2.35	0.00	0.94	0.00	0.00	0.00	0.00	1.41
Trees/shrubs and herbs		campanulids	1.85	0.00	0.31	0.00	0.00	0.92	0.00	0.62
Trees/shrubs and herbs	Carrot order	Apiales	0.93	0.00	0.47	0.00	0.00	0.47	0.00	0.00
Trees/shrubs and herbs	Ivy family	Araliaceae	3.53	0.00	1.01	0.00	0.50	1.01	0.00	1.01
Trees/shrubs and herbs		Asterales	0.57	0.00	0.29	0.00	0.00	0.00	0.00	0.29
Trees/shrubs and herbs	Daisy family	Asteraceae	0.81	0.00	0.00	0.00	0.00	0.00	0.27	0.54
Trees/shrubs and herbs		Asteroideae	0.50	0.00	0.00	0.00	0.00	0.00	0.25	0.25
Trees/shrubs and herbs		Dipsacales	0.87	0.00	0.00	0.00	0.00	0.44	0.00	0.44
Trees/shrubs and herbs		Ericales	0.71	0.00	0.00	0.00	0.00	0.36	0.00	0.36
Trees/shrubs and herbs		lamiids	5.48	0.00	5.48	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Boraginales	0.65	0.00	0.65	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Gentianales	0.90	0.00	0.00	0.00	0.00	0.90	0.00	0.00
Trees/shrubs and herbs	**Periwinkle tribe**	**Vinceae**	**0.78**	**0.00**	**0.78**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Trees/shrubs and herbs	Dead-nettle order	Lamiales	5.27	0.00	4.39	0.00	0.00	0.00	0.00	0.88
Trees/shrubs and herbs	Dead-nettle family	Lamiaceae	2.35	0.00	0.78	0.78	0.00	0.00	0.00	0.78
Trees/shrubs and herbs		Solanales	0.38	0.00	0.38	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Bindweed family	Convolvulaceae	0.70	0.00	0.35	0.00	0.00	0.00	0.00	0.35
Trees/shrubs and herbs	Nightshade subfamily	Solanoideae	0.95	0.00	0.47	0.00	0.00	0.00	0.00	0.47
Trees/shrubs and herbs	Knotweed family	Polygonaceae	0.55	0.00	0.55	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Rose subfamily	Rosoideae	5.56	0.00	0.00	0.00	1.39	0.00	0.00	4.17
Trees/shrubs and herbs	Mustard order	Brassicales	1.30	0.00	0.00	0.00	0.00	0.00	1.30	0.00

Continued on next page

Table 4.9 continued

Ecological category	Common name	Taxon	Total	378	359	337	280	270	230	210
		Ingroup	5500.50	131.58	1600.04	355.28	91.01	878.58	459.04	1984.98
Trees/shrubs and herbs		Malvales	0.58	0.00	0.00	0.00	0.00	0.00	0.58	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	3.26	0.00	0.65	0.00	0.00	1.31	0.65	0.65
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	1.11	0.00	0.56	0.56	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Myrtales	0.99	0.00	0.99	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Loosestrife family	Lythraceae	1.26	0.00	1.26	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Saxifragales	4.43	0.19	2.89	0.00	0.00	0.39	0.19	0.77
Trees/shrubs and herbs		Asparagales	0.88	0.00	0.64	0.00	0.00	0.16	0.00	0.08
Trees/shrubs and herbs	**Palms**	**Arecaceae**	**0.58**	**0.00**	**0.00**	**0.00**	**0.29**	**0.00**	**0.00**	**0.29**
Trees/shrubs and herbs		Liliales	0.07	0.00	0.03	0.00	0.00	0.00	0.00	0.03
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.31	0.00	0.16	0.16	0.00	0.00	0.00	0.00
Grasses and relatives	Grass order	Poales	22.54	0.88	8.49	0.29	0.88	2.63	0.59	8.78
Grasses and relatives	Sedges	Cyperaceae	6.17	0.00	1.54	0.00	0.00	3.08	0.00	1.54
Grasses and relatives		Cyperoideae	3.37	0.00	1.69	0.00	1.69	0.00	0.00	0.00
Grasses and relatives	True sedges	Carex	18.22	0.00	9.11	0.00	0.00	0.00	0.00	9.11
Grasses and relatives	Rushes	Juncaceae	1.04	0.00	0.00	0.00	0.00	0.00	0.00	1.04
Grasses and relatives		Coleanthinae	0.42	0.00	0.00	0.00	0.00	0.00	0.00	0.42
Grasses and relatives	Barlies	Hordeum	0.26	0.00	0.26	0.00	0.00	0.00	0.00	0.00
Grasses and relatives		PACMAD clade	222.25	0.00	88.30	0.00	0.00	42.66	0.00	91.28
Grasses and relatives		Chloridoideae	0.48	0.00	0.48	0.00	0.00	0.00	0.00	0.00
Grasses and relatives		**Panicoideae**	**16.19**	**1.08**	**4.86**	**0.00**	**2.16**	**2.70**	**2.16**	**3.24**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**13.97**	**3.99**	**0.80**	**1.20**	**2.00**	**1.20**	**2.00**	**2.79**
Grasses and relatives		**Paniceae**	**13.41**	**4.02**	**0.67**	**0.67**	**1.34**	**2.68**	**1.34**	**2.68**
Ferns	Ferns	Polypodiopsida	0.27	0.00	0.27	0.00	0.00	0.00	0.00	0.00
Ferns	Horsetails	Equisetum	0.18	0.00	0.09	0.00	0.00	0.05	0.00	0.05
Ferns	Leptosporangiate ferns	Polypodiidae	0.41	0.00	0.16	0.00	0.00	0.08	0.00	0.16
Ferns	Polypod ferns	Polypodiales	1.77	0.00	0.56	0.00	0.08	0.40	0.00	0.72
Ferns	Spleenwort suborder	Aspleniineae	3.78	0.09	1.26	0.00	0.00	1.35	0.18	0.90
Ferns		Athyriaceae	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.07
Ferns		Athyrium	0.56	0.00	0.00	0.00	0.00	0.42	0.00	0.14
Ferns	Bladder ferns	Cystopteris	0.14	0.00	0.14	0.00	0.00	0.00	0.00	0.00
Ferns		Thelypteridaceae	0.39	0.00	0.29	0.00	0.00	0.10	0.00	0.00
Ferns		Phegopteridoideae	0.28	0.00	0.14	0.00	0.00	0.14	0.00	0.00
Ferns		Thelypteridoideae	0.17	0.00	0.08	0.00	0.00	0.00	0.00	0.08
Ferns	Marsh fern	Thelypteris palustris	0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.00
Ferns	Bracken family	Dennstaedtiaceae	0.51	0.10	0.00	0.00	0.00	0.10	0.21	0.10
Bryophytes	Mosses	Bryophyta	1.94	0.00	0.00	0.00	0.00	0.00	0.00	1.94
Bryophytes		Bryidae	1.93	0.00	0.00	0.00	1.93	0.00	0.00	0.00
Bryophytes	Feather mosses	Hypnales	2.24	0.00	0.00	0.00	0.00	2.24	0.00	0.00
Mixed		Tracheophyta	3.98	0.00	1.38	0.17	0.00	0.35	0.17	1.90
Mixed		Euphyllophyta	4.47	0.00	3.10	0.00	0.00	0.52	0.00	0.86
Mixed	Monocots	Liliopsida	548.30	0.00	162.76	40.51	0.00	71.15	44.39	229.50

Table 4.10: European taxa in Metazoa (excluding Primates) from ELF059. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	378	359	337	280	270	230	210
			Ingroup	69.55	4.49	12.37	0.66	4.09	18.60	7.17	22.17
Salt/brackish aquatics	Tunicates	Tunicates	Tunicata	12.20	0.00	0.00	0.00	0.00	4.07	4.07	4.07
Salt/brackish aquatics	Tunicates	Sea squirts	Ascidiacea	6.90	0.00	0.00	0.00	0.00	3.45	0.00	3.45
Salt/brackish aquatics	Crustaceans	Pandalid shrimps	Pandalidae	0.12	0.12	0.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Annelids		Fabriciidae	1.80	0.00	0.00	0.00	0.90	0.00	0.90	0.00
Salt/brackish aquatics	Bivalves	Cockles	Cardiidae	0.66	0.00	0.00	0.00	0.00	0.66	0.00	0.00
Salt/brackish aquatics	Cnidarians	Hexacorals	Hexacorallia	1.62	0.00	0.00	0.00	0.00	1.62	0.00	0.00
Salt/brackish aquatics	Cnidarians	True jellyfish	Scyphozoa	2.10	0.00	0.00	0.00	0.00	2.10	0.00	0.00
Mixed aquatics	Fishes		Acanthomorphata	1.06	0.00	0.00	0.00	0.00	0.00	0.00	1.06
Mixed aquatics	Fishes		Percomorphaceae	4.22	0.00	1.06	0.00	0.00	2.11	1.06	0.00
Mixed aquatics	Fishes	Perch series	Eupercaria	1.09	0.00	0.00	0.00	1.09	0.00	0.00	0.00
Mixed aquatics	Fishes		Ovalentaria	0.93	0.00	0.00	0.00	0.93	0.00	0.00	0.00
Mixed aquatics	Copepods		Temoridae	0.49	0.00	0.49	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Copepods		Podoplea	1.05	0.00	0.00	0.00	0.00	1.05	0.00	0.00
Mixed aquatics	Bivalves	Mussels	Mytilidae	1.12	0.00	0.00	0.00	0.00	0.00	0.00	1.12
Mixed aquatics	Bivalves		Mytilinae	1.32	0.00	0.00	0.66	0.00	0.00	0.66	0.00
Mixed aquatics	Cnidarians	Cnidarians	Cnidaria	1.17	0.00	0.00	0.00	1.17	0.00	0.00	0.00
Mixed aquatics	Hydrozoans		Anthoathecata	2.29	0.00	1.15	0.00	0.00	0.00	0.00	1.15
Mixed aquatics	Hydrozoans		Filifera	1.75	0.00	0.00	0.00	0.00	0.00	0.00	1.75
Mixed aquatics	Hydrozoans	Thecate hydroids	Leptothecata	4.88	0.00	0.00	0.00	0.00	2.44	0.00	2.44
Mixed aquatics	Hydrozoans		Campanulariidae	2.44	0.00	2.44	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Hydrozoans	Sea fur	Obelia	2.44	0.00	2.44	0.00	0.00	0.00	0.00	0.00
Terrestrial	Mammals	Microbats	Microchiroptera	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.40
Terrestrial	Coleopterans		Elaphropus	1.74	0.00	0.00	0.00	0.00	0.00	0.00	1.74
Terrestrial	Hymenopterans	Honey bees	Apis	4.37	4.37	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Hymenopterans		Formicinae	3.42	0.00	3.42	0.00	0.00	0.00	0.00	0.00
Mixed	Animals		Eumetazoa	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.25
Mixed	Birds	Perching birds	Passeriformes	0.76	0.00	0.76	0.00	0.00	0.00	0.00	0.00
Mixed	Invertebrates		Ecdysozoa	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.50
Mixed	Coleopterans	Beetles	Coleoptera	1.32	0.00	0.00	0.00	0.00	0.00	0.00	1.32
Mixed	Invertebrates	Lophotrochozoans	Lophotrochozoa	3.73	0.00	0.62	0.00	0.00	0.62	0.00	2.48
Mixed	Molluscs	Molluscs	Mollusca	0.97	0.00	0.00	0.00	0.00	0.49	0.49	0.00

Embryophyta

Aquatics

Freshwater taxa are diverse and found throughout ELF059, but have small biogenomic masses. Most consistent are *Potamogeton* (pondweed) and the wetland grasses *Typha* (bulrushes) and Typhaceae. Instead, the most frequent ecological category in all seven samples is saltwater aquatics: overwhelmingly *Zostera*, but also *Ruppia* (tasselweeds) and its parent family Cymodoceaceae in all but the deepest sample (378). Cymodoceaceae north of the Mediterranean can be interpreted as *Ruppia*. *Ruppia* grows in brackish water and is often found at low frequency with *Zostera*, suggesting a lagoon or estuary instead of fully marine environments.

Much more unusual is a small signal from *Posidonia* (Neptune-grasses), of which only *P. oceanica* is European. It is currently restricted to the Mediterranean Sea. Chefaoui *et al.* (2017) argue that *P. oceanica* was further confined to refugia on southern coasts during the last glacial maximum (LGM). If the northern Mediterranean was climatically unsuitable, it is difficult to imagine a northern population of *P. oceanica* (perhaps from a previous interglacial) surviving the LGM to colonise the coasts of Doggerland. If instead the Mediterranean population expanded out to Doggerland after the LGM, why is it not present on any Atlantic coasts today? Further investigation in Chapter 6 (Mesophilic taxa and human disturbance indicators) suggests that it may be more appropriate to interpret *Posidonia* reads as mis-assigned. They will not be highlighted in subsequent cores.

Finally, all samples contain similar mixed aquatics signals from pondweeds (*Stuckenia*, Potamogetonaceae), the reed group, and order Alismatales. Alismatales may represent harder-to-assign reads from any of its child taxa present, such as *Stuckenia* or *Zostera*, or unrepresented taxa.

Terrestrial

All samples have a relatively small but diverse woody signal mostly comprising the willow group. More consistent taxa include Betulaceae, its child taxon *Corylus* (hazels), *Prunus*, and possibly *Tilia* if Malvaceae can be interpreted as such. Sample 359 also contains another southern European taxon, *Platanus*, which is native to floodplain forest east of Italy but naturalised in Britain. If genuine, this result would be biogeographically interesting but not necessarily an environmental indicator.

Grasses and relatives are also found throughout the core. Most reads are assigned to the PACMAD clade, which could be interpreted as more reeds. The Panicoideae group may also be mis-assigned reads, or potentially an indicator of warm climate. Other taxa are less informative. Note that the *Hordeum* (barleys) signal in 359 does not necessarily represent domesticated barley, either from contamination or early Neolithic activity, because *H. murinum* (wall barley) and *H. marinum* (sea barley) are widespread British natives. Further investigation of cereals in Chapter 6 (Mesophilic taxa and human disturbance indicators) suggests that these reads are probably ancient, but cannot confirm their taxonomy.

Metazoa

Aquatics are most numerous and have most biogenomic mass in all but the deepest sample, which only returned two reads. Specifically saltwater biogenomic mass increases up the core, but the proportion remains similar, so this does not necessarily indicate ecological change.

Exclusively terrestrial taxa are fewer and more dubious. Sample 378 has a read from *Apis* (honey bees), a heavily-studied domesticate to which the read may have been over-assigned. 210 has Microchiroptera (microbats), which includes many native British species. However, modern bat material is frequently processed in a neighbouring laboratory, and there were insufficient reads for age-authentication, so these may be contamination.

4.5.3 Pianka scores

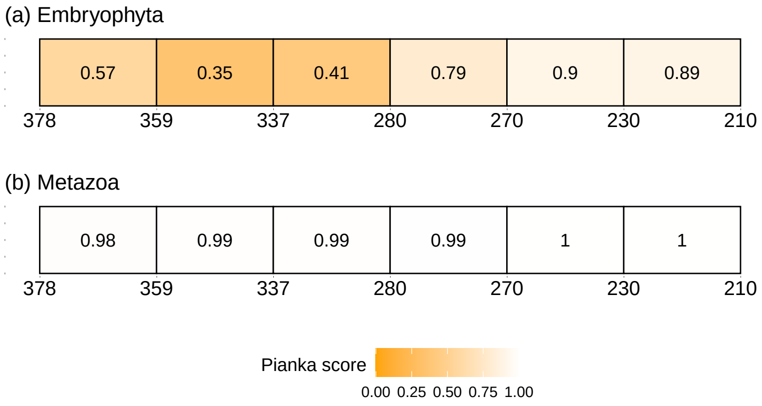


Figure 4.17: Pianka similarity scores between adjacent samples in ELF059 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores. Comparisons involving empty samples have no score.

The Pianka scores for Embryophyta (figure 4.17a) are very high, reaching the maximum of 1 between the last three samples. Despite inflation from higher taxa, this does suggest significant similarity across samples, and is consistent with the similar taxonomic profiles. The scores for Metazoa reach far lower; samples with less data are less likely to contain very higher taxa, so are probably less affected by inflation, and the amount of Metazoa data is usually small. These scores suggest that the first three samples are relatively distinct from their neighbours, with increasing similarity towards the top of the core. As there was little evidence of changes in ecological category from Metazoa, the apparently diverse taxa may simply be a result of limited sampling from the environment.

4.5.4 Summary

ELF059 is another *Zostera*-dominated core. However, the presence of *Ruppia* in most samples (the exception having relatively little data) and reeds throughout suggests a brackish environment rather than fully marine. The diverse terrestrial taxa are consistent with dry land nearby, although the taphonomy of sedaDNA in this context is poorly understood, so "nearby" is not well defined.

This core also contains two plants not native to Great Britain. *Platanus* is inconclusive but would not have much influence on ecological reconstruction. *Posidonia*, on the other hand, would suggest significantly warmer sea temperatures, although later analysis in Chapter 6 (Mesophilic taxa and human disturbance indicators) suggests it is most likely a false positive. PIA is not completely accurate and some manual discrimination is regrettably still required when interpreting these results.

4.6 ELF059A

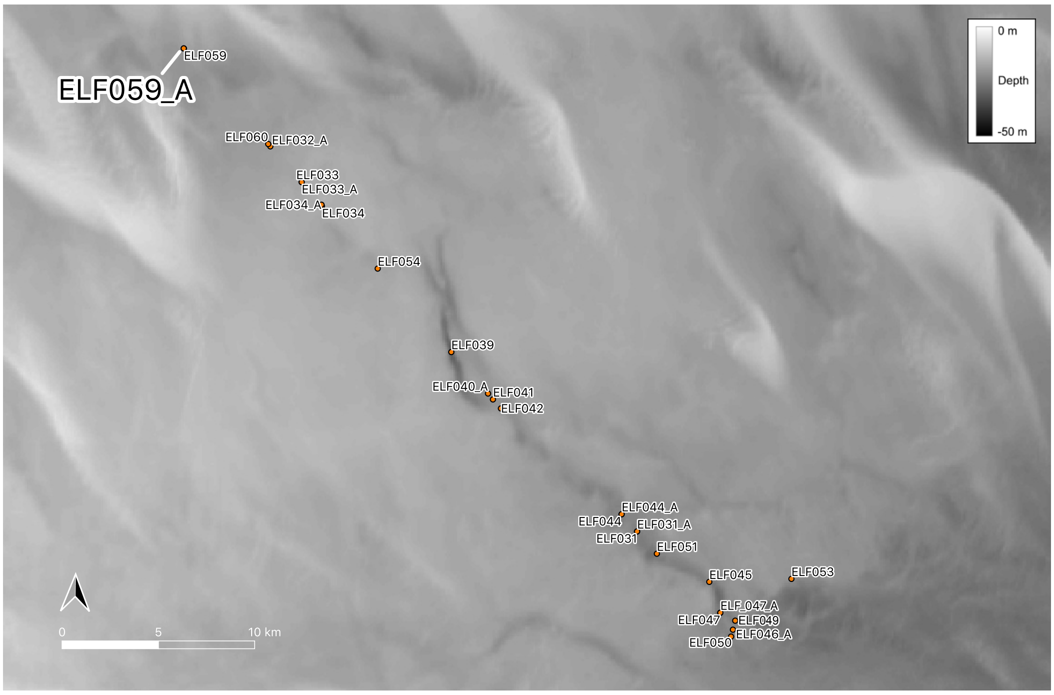


Figure 4.18: Bathymetry map of the palaeochannel transect highlighting ELF059A. Darker grey indicates greater depth.

ELF059A was taken at the same location as ELF059: at the start of the transect and ∼20 km northwest of the start of the palaeochannel.

4.6.1 Read counts

There is a loose trend of increasing read count up the core, from nearly 0 in sample 320 to over 27,000 Embryophyta reads in sample 135. The exception is the oldest sample, 355. However, the additional reads in 355 can be explained by greater sequencing effort. Recall that eight samples were deep-sequenced to investigate their possible associations with the Storegga tsunami (Chapter 2: Main materials and methods; section 2.1.5). These are marked in the biogenomic mass tables with a ”*”. The tsunami is not the focus of this thesis, but see Gaffney *et al.* (2020) for further discussion. One of those samples, ELF059A_355, had already undergone initial sequencing. I decided that the most consistent approach for incorporating both the initial and deep-sequencing data for sample ELF059A_355 was simply to merge the four replicates, as pairs of replicates were merged for other samples. As with the other deep-sequenced samples, any increased read count and taxonomic richness will be interpreted in light of the greater sequencing effort. For example, greater effort can explain the elevated read counts for sample 355: the original sequencing returned 3,956 Embryophyta and 7 animal reads, which is more consistent with the general trend.

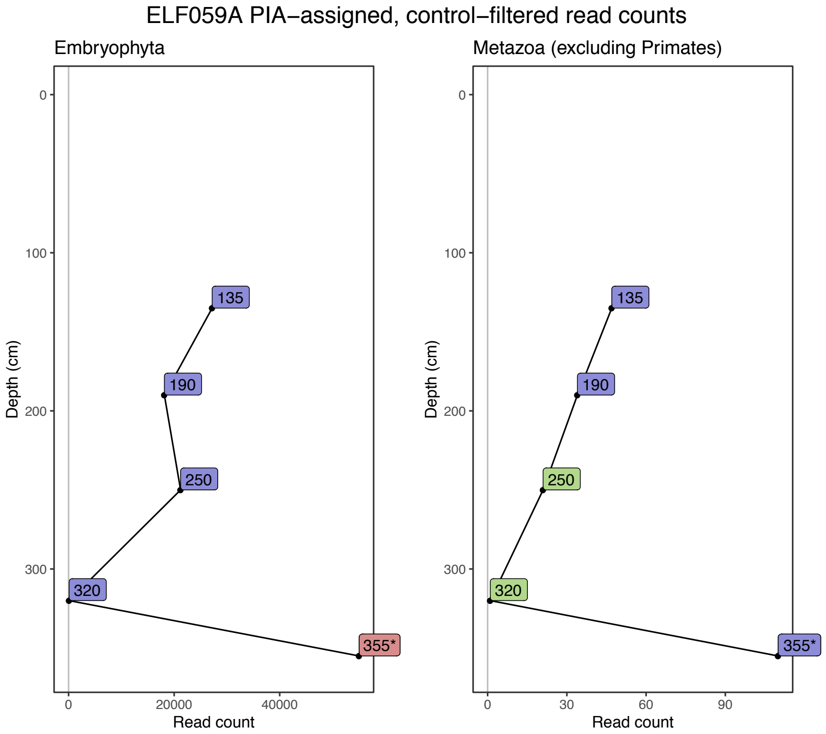


Figure 4.19: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF059A. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample 355(*) was sequenced in greater depth. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.6.2 Taxonomic profiles

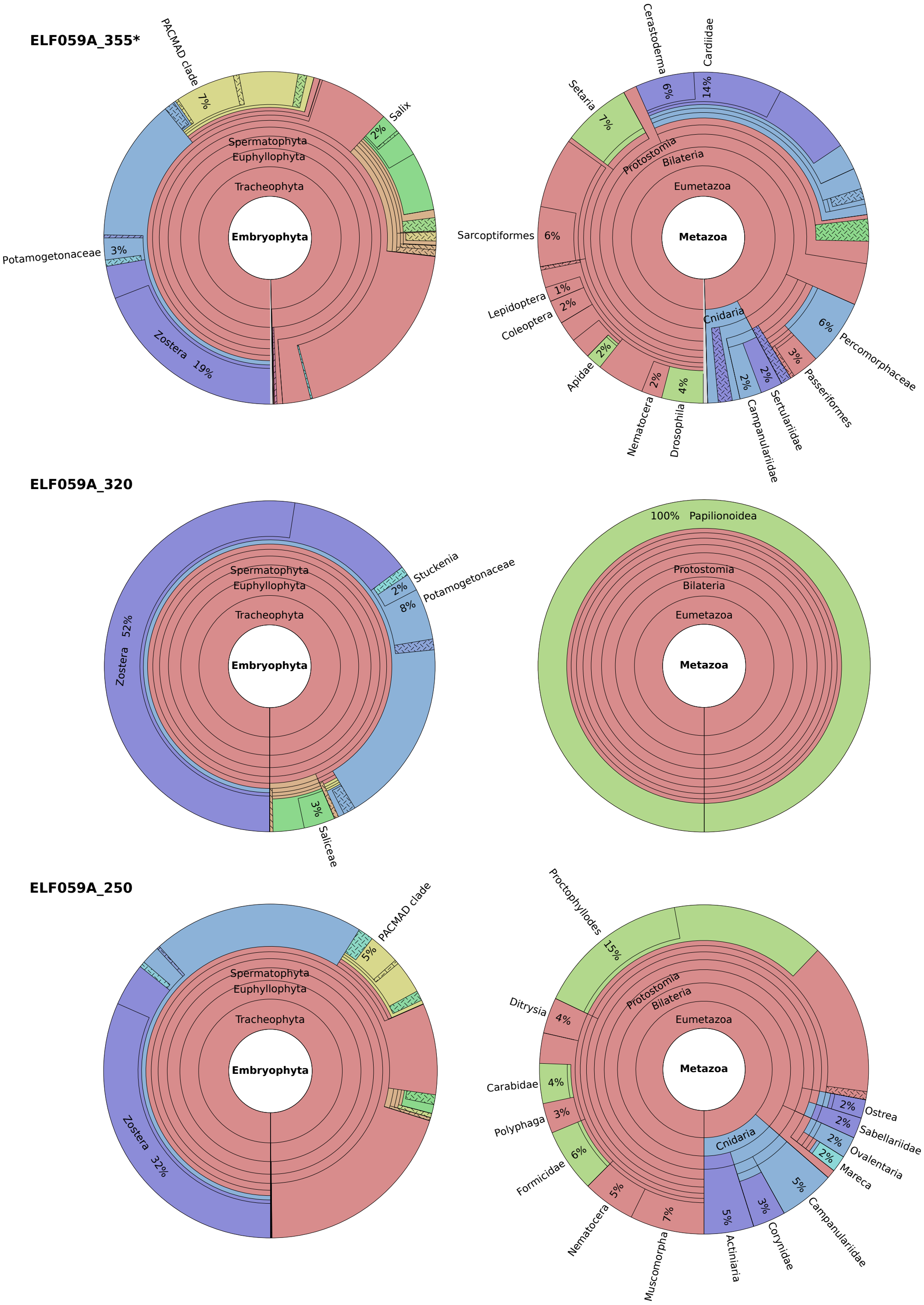


Figure 4.20: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF059A, samples 355, 320, and 250. 355 (*) was sequenced in greater depth. Continued in figure 4.21. See figure 4.4 for colour key.

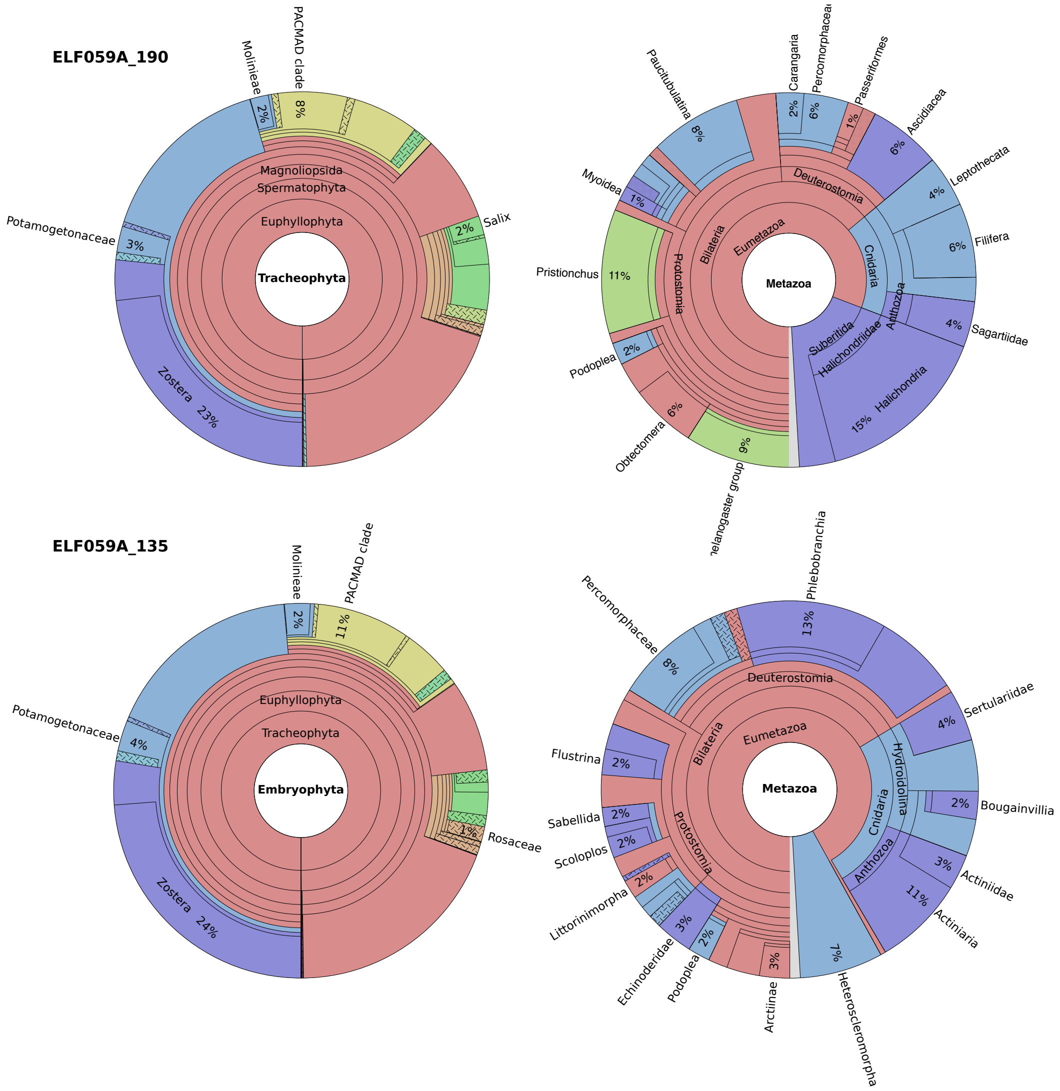


Figure 4.21: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF059A, samples 190 and 135. See figure 4.4 for colour key.

Table 4.11: Biogenomic masses of European taxa in Embryophyta from ELF059A. Samples are in cm. Sample 355 (*) was resequenced in greater depth. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	355*	320	250	190	135
		Ingroup	34841.66	15821.37	59.48	5868.44	5296.97	7795.40
Freshwater aquatics	Coontails	Ceratophyllum	1.45	0.00	0.00	1.45	0.00	0.00
Freshwater aquatics		Menyanthaceae	2.86	0.00	0.00	0.00	0.00	2.86
Freshwater aquatics	Gipsyworts	Lycopus	1.60	1.60	0.00	0.00	0.00	0.00
Freshwater aquatics	Bladderwort family	Lentibulariaceae	10.08	7.56	0.00	0.00	2.52	0.00
Freshwater aquatics	Bladderworts	Utricularia	17.18	12.89	0.00	0.00	0.00	4.30
Freshwater aquatics	Water-starwort	Callitriche	0.46	0.46	0.00	0.00	0.00	0.00
Freshwater aquatics	Water-milfoil family	Haloragaceae	34.62	23.08	0.00	4.62	0.00	6.92
Freshwater aquatics	Water-milfoils	Myriophyllum	69.23	57.69	0.00	0.00	6.92	4.62
Freshwater aquatics	Water-plantain family	Alismataceae	0.76	0.42	0.00	0.00	0.21	0.14
Freshwater aquatics	Arrowheads	Sagittaria	0.18	0.06	0.00	0.00	0.00	0.12
Freshwater aquatics	Duckweed family	Lemnoideae	6.10	1.22	0.00	1.22	1.22	2.44
Freshwater aquatics	Elodea family	Hydrocharitaceae	4.15	1.38	0.20	0.99	1.19	0.40
Freshwater aquatics	Pondweeds	Potamogeton	125.07	52.29	0.54	21.56	16.71	33.96
Freshwater aquatics		C3 Cyperus	2.22	2.22	0.00	0.00	0.00	0.00
Freshwater aquatics	Bulrush family	Typhaceae	119.42	26.21	0.00	29.13	26.21	37.86
Freshwater aquatics	Bur-reeds	Sparganium	4.00	0.00	0.00	0.00	2.00	2.00
Freshwater aquatics	Bulrushes	Typha	33.96	11.32	0.00	7.55	7.55	7.55
Freshwater aquatics	Waterlilly order	Nymphaeales	8.48	8.48	0.00	0.00	0.00	0.00
Freshwater aquatics	Waterlilly family	Nymphaeaceae	29.73	16.46	0.00	0.53	10.09	2.65
Freshwater aquatics	Waterlillies	Nuphar	2.84	1.78	0.00	0.00	0.71	0.36
Freshwater aquatics	Waterlillies	Nymphaea	2.32	1.55	0.00	0.00	0.77	0.00
Freshwater aquatics	Quillworts	Isoetes	0.15	0.00	0.00	0.00	0.00	0.15
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	45.99	20.21	0.00	8.36	8.36	9.06
Salt/brackish aquatics	Tasselweeds	Ruppia	52.36	20.60	0.43	6.01	10.73	14.59
Salt/brackish aquatics	**Neptune-grasses**	**Posidonia**	**1.92**	**0.00**	**0.00**	**0.32**	**0.32**	**1.28**
Salt/brackish aquatics	Eelgrass family	Zosteraceae	1225.69	499.08	7.34	242.20	183.49	293.58
Salt/brackish aquatics	Eelgrasses	Zostera	7937.61	2998.17	31.19	1845.87	1218.35	1844.04
Mixed aquatics	Alismatids	Alismatales	5524.77	2176.67	10.69	1198.76	814.96	1323.69
Mixed aquatics	Pondweed family	Potamogetonaceae	767.60	335.21	2.98	105.62	117.52	206.28
Mixed aquatics	Pondweeds	Stuckenia	106.61	44.13	0.99	10.91	17.85	32.73
Mixed aquatics		Fuireneae	5.36	5.36	0.00	0.00	0.00	0.00
Mixed aquatics	Reeds	Arundinoideae	87.71	35.01	0.35	9.90	14.50	27.94
Mixed aquatics	Common reed tribe	Molinieae	457.93	156.87	0.71	47.24	82.49	170.62
Mixed aquatics	Common reeds	Phragmites	12.61	2.10	0.00	3.36	2.52	4.62
Halophytes	Glasswort subfamily	Salicornioideae	0.46	0.00	0.00	0.46	0.00	0.00
Halophytes	Sea-blite subfamily	Suaedoideae	0.94	0.00	0.00	0.00	0.00	0.94
Halophytes	Leadwort family	Plumbaginaceae	0.21	0.00	0.00	0.00	0.00	0.21
Halophytes	**Desert thumb**	**Cynomorium**	**0.19**	**0.00**	**0.00**	**0.19**	**0.00**	**0.00**
Trees/shrubs		Cupressales	0.07	0.07	0.00	0.00	0.00	0.00
Trees/shrubs	Pines	Pinus	0.07	0.07	0.00	0.00	0.00	0.00
Trees/shrubs	Box family	Buxaceae	0.94	0.00	0.00	0.00	0.00	0.94
Trees/shrubs	Elder family	Adoxaceae	0.26	0.13	0.00	0.00	0.00	0.13
Trees/shrubs		Viburnum	0.79	0.00	0.00	0.00	0.26	0.53
Trees/shrubs		Cornales	3.07	1.23	0.00	1.23	0.61	0.00

Continued on next page

Table 4.11 continued

Ecological category	Common name	Taxon	Total	355*	320	250	190	135
		Ingroup	34841.66	15821.37	59.48	5868.44	5296.97	7795.40
Trees/shrubs	Dogwoods	Cornus	1.19	0.00	0.00	0.00	0.59	0.59
Trees/shrubs	Staff vine family	Celastraceae	1.20	1.20	0.00	0.00	0.00	0.00
Trees/shrubs	Beech order	Fagales	48.71	48.71	0.00	0.00	0.00	0.00
Trees/shrubs	Birch family	Betulaceae	118.79	42.12	0.00	7.56	25.92	43.20
Trees/shrubs	Alders	Alnus	20.97	3.23	0.00	0.00	3.23	14.52
Trees/shrubs	Birches	Betula	8.72	7.75	0.00	0.00	0.00	0.97
Trees/shrubs	Hazels	Corylus	13.64	6.82	0.00	2.27	2.27	2.27
Trees/shrubs	Oak family	Fagaceae	54.11	54.11	0.00	0.00	0.00	0.00
Trees/shrubs	Oaks	Quercus	20.10	2.12	0.00	1.06	4.23	12.70
Trees/shrubs	**Walnut family**	**Juglandaceae**	**1.37**	**0.00**	**0.00**	**0.00**	**0.00**	**1.37**
Trees/shrubs	Willow family	Salicaceae	1308.66	893.50	1.81	50.54	207.58	155.23
Trees/shrubs	Willow tribe	Saliceae	606.35	386.69	1.82	30.86	119.82	67.17
Trees/shrubs	Poplars	Populus	73.19	43.52	0.00	3.96	13.85	11.87
Trees/shrubs	Willows	Salix	468.86	292.39	0.00	20.76	86.51	69.20
Trees/shrubs		Amygdaleae	2.21	0.00	0.00	0.00	0.00	2.21
Trees/shrubs	Stone fruit trees	Prunus	116.88	3.65	0.00	5.48	21.91	85.83
Trees/shrubs	Apple tribe	Maleae	17.83	8.02	0.00	1.78	2.67	5.35
Trees/shrubs	Hawthorns	Crataegus	0.86	0.00	0.00	0.00	0.86	0.00
Trees/shrubs	Roses	Rosa	1.32	0.00	0.00	0.00	0.00	1.32
Trees/shrubs	Blackberry and relatives	Rubus	4.44	2.22	0.00	2.22	0.00	0.00
Trees/shrubs	Elm family	Ulmaceae	14.56	10.54	0.00	0.00	0.50	3.51
Trees/shrubs	Elms	Ulmus	1.29	1.29	0.00	0.00	0.00	0.00
Trees/shrubs		**Laurales**	**0.65**	**0.00**	**0.00**	**0.00**	**0.65**	**0.00**
Trees/shrubs	Barberry family	Berberidaceae	0.23	0.00	0.00	0.00	0.00	0.23
Trees/shrubs	Barberry subfamily	Berberidoideae	0.25	0.25	0.00	0.00	0.00	0.00
Trees/shrubs	Barberry	Berberis	0.70	0.00	0.00	0.00	0.00	0.70
Herbs		Apioideae	1.75	1.75	0.00	0.00	0.00	0.00
Herbs	Water-parsnip tribe	Oenantheae	0.79	0.79	0.00	0.00	0.00	0.00
Herbs	Hemp-agrimony	Eupatorium	0.35	0.00	0.00	0.00	0.00	0.35
Herbs		Inulinae	0.63	0.00	0.00	0.00	0.00	0.63
Herbs	Thistle tribe	Cardueae	0.48	0.00	0.00	0.00	0.00	0.48
Herbs	Thistle subtribe	Carduinae	0.55	0.00	0.00	0.00	0.55	0.00
Herbs	Chicory subfamily	Cichorioideae	0.57	0.57	0.00	0.00	0.00	0.00
Herbs	Chicory tribe	Cichorieae	7.30	7.30	0.00	0.00	0.00	0.00
Herbs	Hawk's-beard subtribe	Crepidinae	0.28	0.28	0.00	0.00	0.00	0.00
Herbs	Sow-thistle subtribe	Hyoseridinae	1.28	1.28	0.00	0.00	0.00	0.00
Herbs	Sow-thistles	Sonchus	5.35	5.35	0.00	0.00	0.00	0.00
Herbs	Bellflower family	Campanulaceae	0.47	0.00	0.00	0.00	0.00	0.47
Herbs	Primrose family	Primulaceae	4.19	3.14	0.00	0.00	1.05	0.00
Herbs	Loosestrifes	Lysimachia	0.72	0.36	0.00	0.00	0.00	0.36
Herbs	Dead-nettle subfamily	Lamioideae	2.54	1.90	0.00	0.00	0.00	0.63
Herbs	Catmint subfamily	Nepetoideae	32.75	28.24	0.00	0.00	3.39	1.13
Herbs	Mint tribe	Mentheae	60.15	55.24	0.00	1.23	1.23	2.46
Herbs	Mints	Mentha	18.29	18.29	0.00	0.00	0.00	0.00
Herbs		Salvia incertae sedis	1.33	0.00	0.00	0.00	1.33	0.00
Herbs	Broomrapes	Phelipanche	0.23	0.23	0.00	0.00	0.00	0.00
Herbs	Yellow-rattle tribe	Rhinantheae	0.32	0.32	0.00	0.00	0.00	0.00
Herbs	Nightshades	Solanum	1.43	1.43	0.00	0.00	0.00	0.00
Herbs	Goosefoot family	Chenopodiaceae	2.39	0.80	0.00	0.80	0.00	0.80
Herbs	Rhubarb family	Rumiceae	1.24	0.83	0.00	0.00	0.41	0.00

Continued on next page

Table 4.11 continued

Ecological category	Common name	Taxon	Total	355*	320	250	190	135
		Ingroup	34841.66	15821.37	59.48	5868.44	5296.97	7795.40
Herbs		NPAAA clade	0.71	0.71	0.00	0.00	0.00	0.00
Herbs		Fabeae	0.29	0.15	0.00	0.00	0.00	0.15
Herbs	Spurge family	Euphorbiaceae	0.31	0.00	0.00	0.31	0.00	0.00
Herbs	Avens	Geum	0.63	0.00	0.00	0.00	0.00	0.63
Herbs		Arabideae	2.19	0.00	0.00	0.00	0.00	2.19
Herbs	Mustard tribe	Brassiceae	0.97	0.97	0.00	0.00	0.00	0.00
Herbs	Mustards	Brassica	1.08	0.00	0.00	0.00	0.00	1.08
Herbs		Camelineae	6.28	0.00	0.00	0.00	3.14	3.14
Herbs	Thale cresses	Arabidopsis	3.07	0.00	0.00	0.00	0.00	3.07
Herbs	Geranium family	Geraniaceae	1.62	1.62	0.00	0.00	0.00	0.00
Herbs	Arum family	Araceae	4.06	1.47	0.00	0.37	0.55	1.66
Herbs	Arum subfamily	Aroideae	0.14	0.14	0.00	0.00	0.00	0.00
Herbs	Onions, leeks, garlics	Allium	0.15	0.15	0.00	0.00	0.00	0.00
Herbs	Daffodil subfamiliy	Amaryllidoideae	0.05	0.00	0.00	0.05	0.00	0.00
Herbs	Asphodel family	Asphodelaceae	0.07	0.07	0.00	0.00	0.00	0.00
Herbs	Bluebell family	Hyacinthaceae	0.08	0.08	0.00	0.00	0.00	0.00
Herbs	Irises	Iris	0.11	0.11	0.00	0.00	0.00	0.00
Herbs	Orchids	Orchidaceae	1.19	0.00	0.17	0.00	0.17	0.85
Herbs		Epidendroideae	2.87	1.28	0.00	0.32	0.00	1.28
Herbs	Coralroot orchids	Corallorhiza	0.08	0.08	0.00	0.00	0.00	0.00
Herbs		Dioscoreales	2.80	1.87	0.00	0.00	0.00	0.93
Herbs		Dioscoreaceae	1.82	0.00	0.00	0.00	0.91	0.91
Herbs		Dioscorea	0.90	0.00	0.00	0.90	0.00	0.00
Herbs	Autumn crocus family	Colchicaceae	0.14	0.14	0.00	0.00	0.00	0.00
Herbs	Lily family	Liliaceae	0.03	0.00	0.00	0.00	0.00	0.03
Herbs	**Birthwort family**	**Aristolochiaceae**	**0.31**	**0.31**	**0.00**	**0.00**	**0.00**	**0.00**
Herbs	**Birthworts**	**Aristolochia**	**5.41**	**2.70**	**0.00**	**0.00**	**2.70**	**0.00**
Herbs	Buttercup subfamily	Ranunculoideae	0.54	0.43	0.00	0.00	0.00	0.11
Herbs	Baneberries	Actaea	0.10	0.10	0.00	0.00	0.00	0.00
Herbs	Buttercup tribe	Ranunculeae	0.63	0.16	0.00	0.00	0.47	0.00
Herbs	Buttercups	Ranunculus	0.74	0.15	0.00	0.15	0.45	0.00
Trees/shrubs and herbs	Gymnosperms	Acrogymnospermae	0.11	0.11	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Gunneridae	4.24	0.94	0.00	0.00	0.94	2.35
Trees/shrubs and herbs		campanulids	16.93	9.85	0.00	0.92	2.16	4.00
Trees/shrubs and herbs	Carrot order	Apiales	1.87	0.93	0.00	0.47	0.47	0.00
Trees/shrubs and herbs	Umbellifer family	Apiaceae	0.91	0.46	0.00	0.46	0.00	0.00
Trees/shrubs and herbs	Ivy family	Araliaceae	5.55	2.02	0.00	0.50	1.01	2.02
Trees/shrubs and herbs		Asterales	5.13	3.71	0.00	0.00	0.29	1.14
Trees/shrubs and herbs	Daisy family	Asteraceae	26.88	23.35	0.00	0.81	1.09	1.63
Trees/shrubs and herbs		Asteroideae	3.01	0.75	0.00	0.25	0.50	1.50
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	0.73	0.18	0.00	0.18	0.00	0.37
Trees/shrubs and herbs	Aster tribe	Astereae	5.18	1.04	0.00	0.52	1.55	2.07
Trees/shrubs and herbs	Ragwort tribe	Senecioneae	0.46	0.23	0.00	0.00	0.00	0.23
Trees/shrubs and herbs		Dipsacales	1.31	0.00	0.00	0.44	0.44	0.44
Trees/shrubs and herbs	Honeysuckle family	Caprifoliaceae	1.13	0.00	0.00	0.00	0.57	0.57
Trees/shrubs and herbs		lamiids	7.53	2.05	0.00	0.00	2.74	2.74
Trees/shrubs and herbs	Dead-nettle order	Lamiales	72.01	62.35	0.00	0.00	3.51	6.15
Trees/shrubs and herbs	Dead-nettle family	Lamiaceae	12.55	10.99	0.00	0.78	0.00	0.78
Trees/shrubs and herbs	Plantain family	Plantaginaceae	2.76	0.00	0.00	0.00	2.76	0.00
Trees/shrubs and herbs		Solanales	0.38	0.38	0.00	0.00	0.00	0.00

Continued on next page

Table 4.11 continued

Ecological category	Common name	Taxon	Total	355*	320	250	190	135
		Ingroup	34841.66	15821.37	59.48	5868.44	5296.97	7795.40
Trees/shrubs and herbs	Bindweed family	Convolvulaceae	1.05	1.05	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Nightshade family	Solanaceae	1.06	1.06	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Nightshade subfamily	Solanoideae	0.47	0.47	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Amaranth family	Amaranthaceae	0.67	0.67	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Goosefoot subfamily	Chenopodioideae	0.82	0.00	0.00	0.82	0.00	0.00
Trees/shrubs and herbs	Legume family	Fabaceae	0.48	0.48	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		50 kb inversion clade	0.40	0.40	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		core genistoids	0.94	0.94	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Malpighiales	122.83	122.83	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Rose subfamily	Rosoideae	16.69	5.56	0.00	2.78	4.17	4.17
Trees/shrubs and herbs		Rosoideae incertae sedis	1.88	1.88	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mustard family	Brassicaceae	1.36	1.36	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	36.54	6.53	0.00	0.00	9.14	20.88
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	5.00	2.22	0.00	0.00	1.11	1.67
Trees/shrubs and herbs		Myrtales	0.99	0.00	0.00	0.99	0.00	0.00
Trees/shrubs and herbs	Loosestrife family	Lythraceae	1.26	0.00	0.00	0.00	0.00	1.26
Trees/shrubs and herbs	**Citrus family**	**Rutaceae**	**1.44**	**1.44**	**0.00**	**0.00**	**0.00**	**0.00**
Trees/shrubs and herbs	Sandalwood order	Santalales	0.06	0.00	0.00	0.00	0.00	0.06
Trees/shrubs and herbs		Saxifragales	17.91	12.33	0.19	1.93	0.96	2.50
Trees/shrubs and herbs		Asparagales	2.56	0.80	0.08	0.40	0.32	0.96
Trees/shrubs and herbs	**Palms**	**Arecaceae**	**0.87**	**0.87**	**0.00**	**0.00**	**0.00**	**0.00**
Trees/shrubs and herbs		Liliales	0.28	0.07	0.00	0.07	0.03	0.10
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.16	0.16	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Buttercup family	Ranunculaceae	0.11	0.11	0.00	0.00	0.00	0.00
Grasses and relatives	Grass order	Poales	201.14	108.62	0.00	21.96	35.43	35.13
Grasses and relatives	Sedges	Cyperaceae	47.81	20.05	0.00	4.63	13.88	9.25
Grasses and relatives		Cyperoideae	33.74	32.06	0.00	1.69	0.00	0.00
Grasses and relatives	True sedges	Carex	77.45	36.45	0.00	13.67	11.39	15.94
Grasses and relatives		Cypereae	1.55	1.55	0.00	0.00	0.00	0.00
Grasses and relatives	Grass family	Poaceae	1696.53	897.90	0.00	186.66	299.17	312.80
Grasses and relatives		BOP clade	169.09	85.13	0.00	23.80	34.69	25.47
Grasses and relatives	Brome grasses	Bromus	0.25	0.25	0.00	0.00	0.00	0.00
Grasses and relatives		Poeae	0.55	0.00	0.00	0.00	0.00	0.55
Grasses and relatives	Oats	Avena	0.13	0.00	0.00	0.00	0.13	0.00
Grasses and relatives		Poeae Chloroplast Group 2 (Poeae type)	0.21	0.21	0.00	0.00	0.00	0.00
Grasses and relatives	Barley subtribe	Hordeinae	0.12	0.00	0.00	0.00	0.00	0.12
Grasses and relatives	Barlies	Hordeum	0.13	0.13	0.00	0.00	0.00	0.00
Grasses and relatives		PACMAD clade	2020.57	908.83	0.00	177.10	314.02	620.61
Grasses and relatives		Chloridoideae	2.88	1.92	0.00	0.48	0.48	0.00
Grasses and relatives	Dropseeds	Sporobolus	0.95	0.00	0.00	0.00	0.00	0.95
Grasses and relatives		Danthonioideae	0.42	0.00	0.00	0.00	0.00	0.42
Grasses and relatives		**Panicoideae**	**52.90**	**1.08**	**0.00**	**16.19**	**18.89**	**16.73**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**40.33**	**35.93**	**0.00**	**1.60**	**0.80**	**2.00**
Grasses and relatives	**Broomsedges**	**Andropogon**	**0.52**	**0.00**	**0.00**	**0.26**	**0.00**	**0.26**
Grasses and relatives		**Paniceae**	**27.48**	**6.03**	**0.00**	**8.04**	**8.04**	**5.36**
Grasses and relatives		**Cenchrinae**	**1.73**	**0.58**	**0.00**	**0.00**	**0.58**	**0.58**
Grasses and relatives	**Panicgrass**	**Panicum**	**1.29**	**0.64**	**0.00**	**0.00**	**0.64**	**0.00**
Ferns	Ferns	Polypodiopsida	0.74	0.47	0.00	0.07	0.00	0.20
Ferns	Horsetails	Equisetum	0.41	0.14	0.00	0.05	0.09	0.14

Continued on next page

Table 4.11 continued

Ecological category	Common name	Taxon	Total	355*	320	250	190	135
		Ingroup	34841.66	15821.37	59.48	5868.44	5296.97	7795.40
Ferns		Ophioglossidae	0.03	0.00	0.00	0.03	0.00	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	2.71	1.81	0.00	0.16	0.33	0.41
Ferns	Polypod ferns	Polypodiales	13.03	9.89	0.00	0.48	0.88	1.77
Ferns	Spleenwort suborder	Aspleniineae	22.78	16.57	0.00	1.08	1.89	3.24
Ferns		Athyriaceae	0.52	0.22	0.00	0.07	0.15	0.07
Ferns		Athyrium	0.56	0.42	0.00	0.00	0.14	0.00
Ferns	Chain fern family	Blechnaceae	0.08	0.00	0.00	0.00	0.00	0.08
Ferns	Bladder ferns	Cystopteris	0.14	0.14	0.00	0.00	0.00	0.00
Ferns		Thelypteridaceae	2.15	1.47	0.00	0.00	0.49	0.20
Ferns		Phegopteridoideae	0.69	0.42	0.00	0.00	0.14	0.14
Ferns		Thelypteridoideae	2.92	1.84	0.00	0.25	0.17	0.67
Ferns	Maiden ferns	Thelypteris	0.13	0.00	0.00	0.13	0.00	0.00
Ferns	Marsh fern	Thelypteris palustris	0.26	0.26	0.00	0.00	0.00	0.00
Ferns	Bracken family	Dennstaedtiaceae	0.51	0.00	0.00	0.10	0.31	0.10
Ferns	Buckler and male ferns	Dryopteris	0.08	0.00	0.00	0.00	0.08	0.00
Ferns		Polypodiaceae	0.08	0.00	0.00	0.00	0.08	0.00
Bryophytes		Bryophytina	3.90	1.95	0.00	0.00	1.95	0.00
Bryophytes		Bryopsida	3.97	0.00	0.00	1.98	1.98	0.00
Bryophytes		Funariaceae	2.15	2.15	0.00	0.00	0.00	0.00
Bryophytes	Liverworts	Marchantiophyta	0.53	0.00	0.00	0.53	0.00	0.00
Mixed	Land plants	Embryophyta	44.30	44.30	0.00	0.00	0.00	0.00
Mixed		Tracheophyta	17.30	10.03	0.00	1.56	1.21	4.50
Mixed		Euphyllophyta	20.65	13.60	0.00	1.89	2.58	2.58
Mixed	Seed plants	Spermatophyta	86.06	86.06	0.00	0.00	0.00	0.00
Mixed	Flowering plants	Magnoliopsida	427.71	427.71	0.00	0.00	0.00	0.00
Mixed	Core angiosperms	Mesangiospermae	6688.29	2999.47	0.00	1169.37	1033.44	1486.02
Mixed	Monocots	Liliopsida	2598.06	1079.71	0.00	516.30	390.06	612.00
Mixed		Petrosaviidae	36.46	36.46	0.00	0.00	0.00	0.00
Mixed		commelinids	92.68	92.68	0.00	0.00	0.00	0.00

Table 4.12: European taxa in Metazoa (excluding Primates) from ELF059A. Samples are in cm. Sample 355 (*) was resequenced in greater depth. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	355*	320	250	190	135
			Ingroup	277.33	115.03	2.79	45.96	55.30	58.24
Freshwater aquatics	Birds	Wigeon genus	Mareca	0.72	0.00	0.00	0.72	0.00	0.00
Salt/brackish aquatics	Tunicates	Tunicates	Tunicata	4.07	0.00	0.00	0.00	0.00	4.07
Salt/brackish aquatics	Tunicates	Sea squirts	Ascidiacea	3.45	0.00	0.00	0.00	3.45	0.00
Salt/brackish aquatics	Tunicates		Phlebobranchia	7.14	0.00	0.00	0.00	0.00	7.14
Salt/brackish aquatics	Echinoderms		Echinacea	1.18	1.18	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Crustaceans		Crangonidae	0.11	0.11	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Kinorhynchs		Echinoderidae	1.79	0.00	0.00	0.00	0.00	1.79
Salt/brackish aquatics	Annelids	Tube worms and relatives	Sabellida	0.90	0.00	0.00	0.00	0.00	0.90
Salt/brackish aquatics	Annelids	Honeycomb worm family	Sabellariidae	0.90	0.00	0.00	0.90	0.00	0.00
Salt/brackish aquatics	Annelids		Orbiniidae	1.01	0.00	0.00	0.00	0.51	0.51
Salt/brackish aquatics	Annelids		Scoloplos	1.52	0.51	0.00	0.00	0.00	1.01
Salt/brackish aquatics	Bryozoans		Cheilostomatida	2.45	1.22	0.00	0.00	0.00	1.22
Salt/brackish aquatics	Bryozoans		Flustrina	1.22	0.00	0.00	0.00	0.00	1.22
Salt/brackish aquatics	Bivalves		Myoidea	0.71	0.00	0.00	0.00	0.71	0.00
Salt/brackish aquatics	Bivalves	Cockle superfamily	Cardioidea	8.99	8.99	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bivalves	Cockles	Cardiidae	9.88	9.88	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bivalves		Cerastoderma	6.57	6.57	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bivalves		Mytilus	2.43	1.22	0.00	0.00	0.61	0.61
Salt/brackish aquatics	Bivalves		Ostrea	0.81	0.00	0.00	0.81	0.00	0.00
Salt/brackish aquatics	Cephalopods		Octopus	0.24	0.00	0.00	0.00	0.00	0.24
Salt/brackish aquatics	Sea anemones	Sea anemones	Actiniaria	6.50	0.00	0.00	2.17	0.00	4.33
Salt/brackish aquatics	Sea anemones		Actiniidae	1.67	0.00	0.00	0.00	0.00	1.67
Salt/brackish aquatics	Sea anemones		Sagartiidae	2.17	0.00	0.00	0.00	2.17	0.00
Salt/brackish aquatics	Cnidarians	Octocorals	Octocorallia	1.56	1.56	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Hydrozoans		Corynidae	1.43	0.00	0.00	1.43	0.00	0.00
Salt/brackish aquatics	Hydrozoans		Bougainvillia	1.33	0.00	0.00	0.00	0.00	1.33
Salt/brackish aquatics	Hydrozoans		Sertulariidae	4.88	2.44	0.00	0.00	0.00	2.44
Salt/brackish aquatics	Poriferans		Suberitida	1.70	0.00	0.00	0.00	1.70	0.00
Salt/brackish aquatics	Poriferans		Halichondria	8.33	0.00	0.00	0.00	8.33	0.00
Mixed aquatics	Fishes		Euteleosteomorpha	0.90	0.00	0.00	0.00	0.00	0.90
Mixed aquatics	Fishes		Percomorphaceae	13.73	7.39	0.00	0.00	2.11	4.22
Mixed aquatics	Fishes		Carangaria	1.32	0.00	0.00	0.00	1.32	0.00
Mixed aquatics	Fishes		Ovalentaria	0.93	0.00	0.00	0.93	0.00	0.00
Mixed aquatics	Fishes		Otomorpha	0.71	0.00	0.00	0.00	0.00	0.71
Mixed aquatics	Mammals	Whales	Cetacea	0.30	0.00	0.00	0.00	0.00	0.30
Mixed aquatics	Copepods		Podoplea	2.10	0.00	0.00	0.00	1.05	1.05
Mixed aquatics	Crustaceans	Caridean shrimps	Caridea	0.08	0.08	0.00	0.00	0.00	0.00
Mixed aquatics	Annelids	Bristle worms	Polychaeta	0.74	0.74	0.00	0.00	0.00	0.00
Mixed aquatics	Gastrotrichs		Paucitubulatina	4.35	0.00	0.00	0.00	4.35	0.00
Mixed aquatics	Bivalves	Bivalves	Bivalvia	1.74	1.16	0.00	0.00	0.00	0.58
Mixed aquatics	Bivalves		Euheterodonta	2.33	2.33	0.00	0.00	0.00	0.00
Mixed aquatics	Bivalves		Veneroida	2.95	2.95	0.00	0.00	0.00	0.00
Mixed aquatics	Bivalves	Mussels	Mytilidae	1.68	0.56	0.00	0.00	0.56	0.56
Mixed aquatics	Bivalves		Mytilinae	0.66	0.00	0.00	0.00	0.66	0.00
Mixed aquatics	Cnidarians	Cnidarians	Cnidaria	1.17	1.17	0.00	0.00	0.00	0.00
Mixed aquatics	Hydrozoans		Hydroidolina	0.87	0.87	0.00	0.00	0.00	0.00

Continued on next page

Table 4.12 continued

Ecological category	Group	Common name	Taxon	Total	355*	320	250	190	135
			Ingroup	277.33	115.03	2.79	45.96	55.30	58.24
Mixed aquatics	Hydrozoans		Anthoathecata	1.15	0.00	0.00	0.00	1.15	0.00
Mixed aquatics	Hydrozoans		Filifera	5.26	0.00	0.00	0.00	3.51	1.75
Mixed aquatics	Hydrozoans	Thecate hydroids	Leptothecata	4.88	0.00	0.00	0.00	2.44	2.44
Mixed aquatics	Hydrozoans		Campanulariidae	4.88	2.44	0.00	2.44	0.00	0.00
Mixed aquatics	Poriferans		Heteroscleromorpha	3.95	0.00	0.00	0.00	0.00	3.95
Terrestrial	**Mammals**	**Goat subfamily**	**Caprinae**	**0.33**	**0.00**	**0.00**	**0.00**	**0.00**	**0.33**
Terrestrial	**Mammals**	**Sheep genus**	**Ovis**	**0.33**	**0.00**	**0.00**	**0.00**	**0.33**	**0.00**
Terrestrial	Nematodes		Pristionchus	5.88	0.00	0.00	0.00	5.88	0.00
Terrestrial	Nematodes		Setaria	7.69	7.69	0.00	0.00	0.00	0.00
Terrestrial	Mites	Group of feather mites	Proctophyllodinae	6.67	0.00	0.00	6.67	0.00	0.00
Terrestrial	Mites	Group of feather mites	Proctophyllodes	6.67	0.00	0.00	6.67	0.00	0.00
Terrestrial	Butterflies	Butterflies	Papilionoidea	2.79	0.00	2.79	0.00	0.00	0.00
Terrestrial	Coleopterans	Ground beetles	Carabidae	1.74	0.00	0.00	1.74	0.00	0.00
Terrestrial	Dipterans	Fruit flies	Drosophila	4.63	4.63	0.00	0.00	0.00	0.00
Terrestrial	Dipterans		melanogaster group	4.95	0.00	0.00	0.00	4.95	0.00
Terrestrial	Hymenopterans	Honey bee family	Apidae	1.82	1.82	0.00	0.00	0.00	0.00
Terrestrial	Hymenopterans	Ants	Formicidae	2.77	0.00	0.00	2.77	0.00	0.00
Mixed	Animals	Animals	Metazoa	1.49	0.50	0.00	0.00	0.50	0.50
Mixed	Animals		Eumetazoa	0.25	0.00	0.00	0.00	0.00	0.25
Mixed	Animals	Chordates	Chordata	0.77	0.00	0.00	0.38	0.00	0.38
Mixed	Mammals	Rodents and lagomorphs	Glires	0.28	0.00	0.00	0.00	0.28	0.00
Mixed	Birds	Perching birds	Passeriformes	3.78	3.02	0.00	0.00	0.76	0.00
Mixed	Squamates	Colubrid snake superfamily	Colubroidea	0.46	0.46	0.00	0.00	0.00	0.00
Mixed	Invertebrates		Protostomia	5.19	4.67	0.00	0.00	0.00	0.52
Mixed	Invertebrates		Ecdysozoa	2.00	1.50	0.00	0.00	0.50	0.00
Mixed	Arthropods	Arthropods	Arthropoda	0.98	0.49	0.00	0.00	0.49	0.00
Mixed	Mites		Acariformes	7.84	7.84	0.00	0.00	0.00	0.00
Mixed	Mites	Biting mites	Sarcoptiformes	6.67	6.67	0.00	0.00	0.00	0.00
Mixed	Mites		Astigmata	6.67	0.00	0.00	6.67	0.00	0.00
Mixed	Spiders		RTA clade	0.35	0.00	0.00	0.35	0.00	0.00
Mixed	Crustaceans	Crustaceans	Crustacea	0.19	0.19	0.00	0.00	0.00	0.00
Mixed	Insects	Winged insects	Pterygota	0.92	0.00	0.00	0.00	0.00	0.92
Mixed	Insects	Metamorphosing insects	Holometabola	1.98	1.98	0.00	0.00	0.00	0.00
Mixed	Lepidopterans	Butterflies and moths	Lepidoptera	1.61	1.61	0.00	0.00	0.00	0.00
Mixed	Lepidopterans		Ditrysia	3.13	0.00	0.00	1.56	1.56	0.00
Mixed	Lepidopterans		Obtectomera	4.69	0.00	0.00	0.00	3.13	1.56
Mixed	Moths		Arctiinae	1.44	0.00	0.00	0.00	0.00	1.44
Mixed	Coleopterans	Beetles	Coleoptera	3.97	2.65	0.00	1.32	0.00	0.00
Mixed	Coleopterans		Polyphaga	1.31	0.00	0.00	1.31	0.00	0.00
Mixed	Dipterans	True flies	Diptera	5.33	5.33	0.00	0.00	0.00	0.00
Mixed	Dipterans		Muscomorpha	3.24	0.00	0.00	3.24	0.00	0.00
Mixed	Dipterans	Midge suborder	Nematocera	4.59	2.30	0.00	2.30	0.00	0.00
Mixed	Hymenopterans	Sawflies, wasps, ants and bees	Hymenoptera	2.34	2.34	0.00	0.00	0.00	0.00
Mixed	Hymenopterans	Stinging wasps, ants and bees	Aculeata	2.33	2.33	0.00	0.00	0.00	0.00
Mixed	Invertebrates	Lophotrochozoans	Lophotrochozoa	5.59	2.48	0.00	0.00	1.86	1.24
Mixed	Annelids	Annelid worms	Annelida	1.54	0.00	0.00	0.00	0.00	1.54
Mixed	Molluscs	Molluscs	Mollusca	1.46	0.49	0.00	0.00	0.00	0.97
Mixed	Gastropods		Littorinimorpha	0.89	0.00	0.00	0.00	0.00	0.89

Embryophyta

Aquatics

Freshwater aquatics are diverse, if proportionally minor. All samples have a strong saltwater signal that mostly consists of *Zostera*, although the mixed aquatic Alismatales and uninformative higher taxa are about as frequent in all but one sample. *Ruppia* and reeds suggest brackish water and land in the region.

Trees/shrubs

All samples have a minor but diverse tree signal mostly consisting of the willow groups, followed by Betulaceae and *Prunus*. Probably as a result of the large data yield, several samples contain unusual taxa. One of the more informative is *Pinus* (the pine genus) in 355. This is only expected in any quantity in Great Britain until ~7,500 BP (Cox and Moore 1993), so to find a weak signal in the deepest sample, which is likely to be the oldest, is encouraging.

Another is Juglandaceae (the walnut family) from sample 135, of which only *Juglans regia* may be native to Europe. Its long history of cultivation makes establishing native status difficult. Beer *et al.* (2008) review pollen and macrofossil evidence that suggests *J. regia* originated in Central Asia and was present across Europe at various times during the Quaternary. European populations are assumed to have retreated to southern refugia during glaciations and the earliest consistent post-glacial evidence is from south-eastern Europe 6,000 BP. Most subsequent evidence across Europe appears to be from cultivated forms. Therefore, *J. regia* in northern Europe before cultivation is biogeographically surprising, although if migration could occur fast enough, the naturalisation of *J. regia* in England today demonstrates a tolerance for this sort of climate (de Rigo *et al.* 2016). Age-authentication in Chapter 6 (Mesophilic taxa and human disturbance indicators) was inconclusive, so for now, Juglandaceae should be taken with caution.

Grasses and relatives

All samples apart from 320, of the smallest read count, contain a terrestrial grass signal at least as strong as that of trees/shrubs. Most biogenomic mass is from higher grass taxa or the Panicoideae group, which should be treated with caution.

Metazoa

Aquatics

All samples apart from 320 have a clear aquatic animal signal. The only freshwater taxon, however, is the duck genus *Mareca* (wigeon and relatives) in sample 250. Saltwater taxa are much more numerous but rarely found in multiple samples, which is probably a result of the low read counts. An exception is the mussel *Mytilus* and its mixed-aquatic parent taxa: this is one of the few consistent animals in many of these cores, suggesting a genuine prevalence of mussel beds.

Terrestrial

All samples have a small terrestrial signal, but many of the taxa are close to model or domesticated animals. Among the exceptions, sample 250 contains the mite groups Proctophyllodinae and Proctophyllodes. These are feather mites, bird parasites, so may be connected to *Mareca*.

4.6.3 Pianka scores

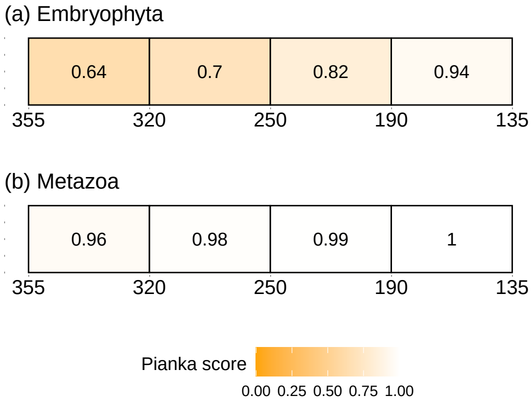


Figure 4.22: Pianka similarity scores between adjacent samples in ELF059A for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores. Comparisons involving empty samples have no score.

The pattern of Pianka scores resembles that of ELF059: the Embryophyta profiles all appear very similar, whereas the Metazoa profiles of deeper samples are more distinct. This may be due to the decrease in data down the core: samples with low read counts are less likely to overlap with others due to limited sampling from the environment. There was certainly no clear evidence of ecological change in the profiles.

4.6.4 Summary

The Embryophyta picture is again of an estuarine environment, combining open water and reeds, with nearby floodplain woodland and other more terrestrial plants. These may include Juglandaceae, which if genuine, could suggest a warmer climate than expected. Terrestrial animals may be more frequent than in ELF059, but most Metazoa biogenomic mass is still saltwater. There is no evidence of ecological change.

4.7 ELF060

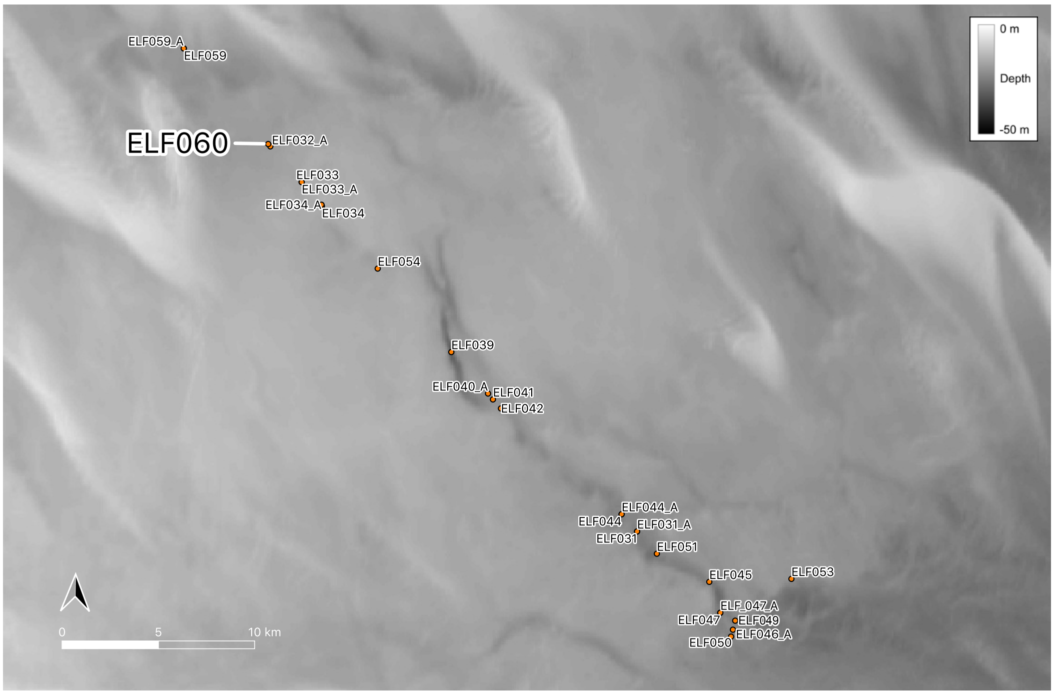


Figure 4.23: Bathymetry map of the palaeochannel transect highlighting ELF060. Darker grey indicates greater depth.

ELF060 is ∼10 km northwest of the visible palaeochannel and several km closer than ELF059 and ELF059A. It is very close to ELF032A.

4.7.1 Read counts

All four ELF060 samples have a relatively large number of Embryophyta reads but Metazoa data is scarce. For both groups, the two middle samples returned more data, but in Metazoa this only results in a single read versus zero. Plants will be more informative in this core than animals.

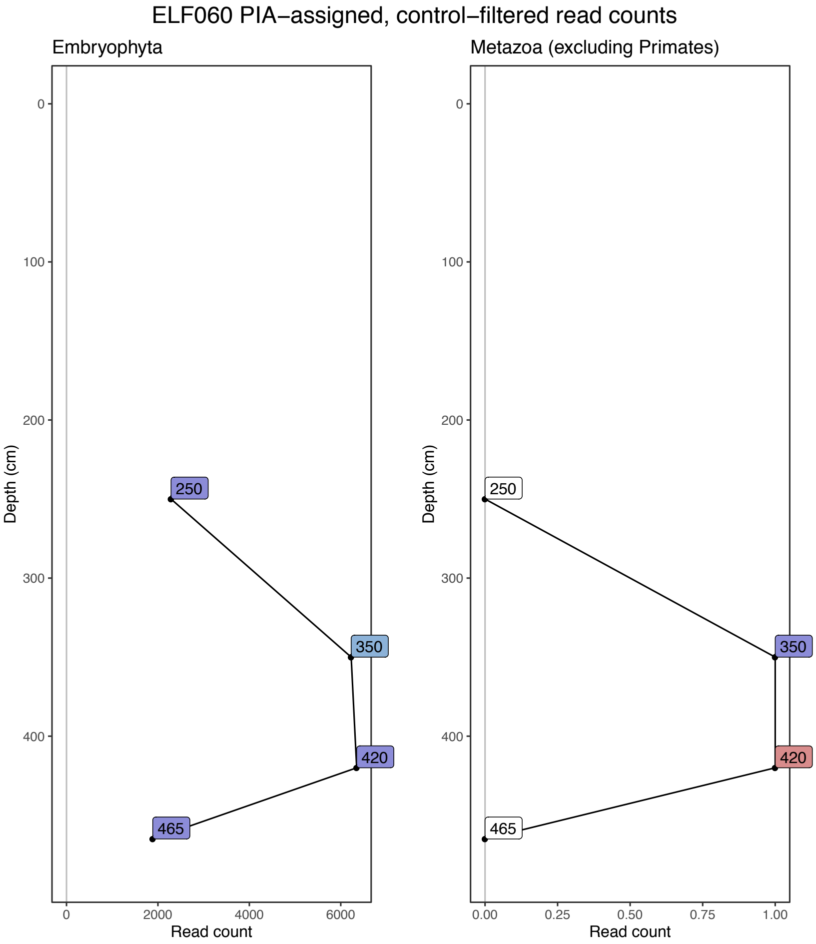


Figure 4.24: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF060. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.7.2 Taxonomic profiles

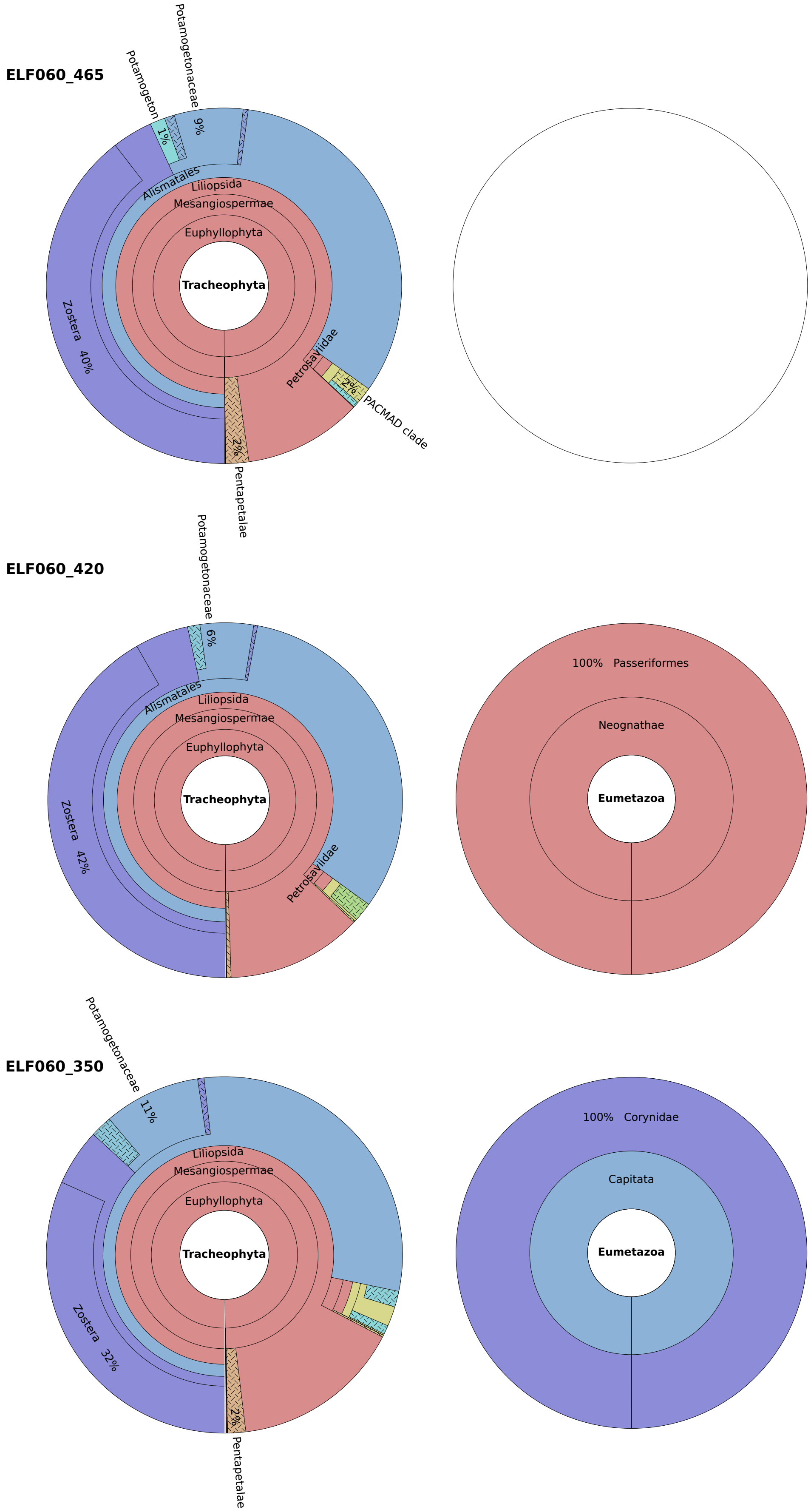


Figure 4.25: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF060, samples 465, 420, and 350. Continued in figure 4.26. See figure 4.4 for colour key.

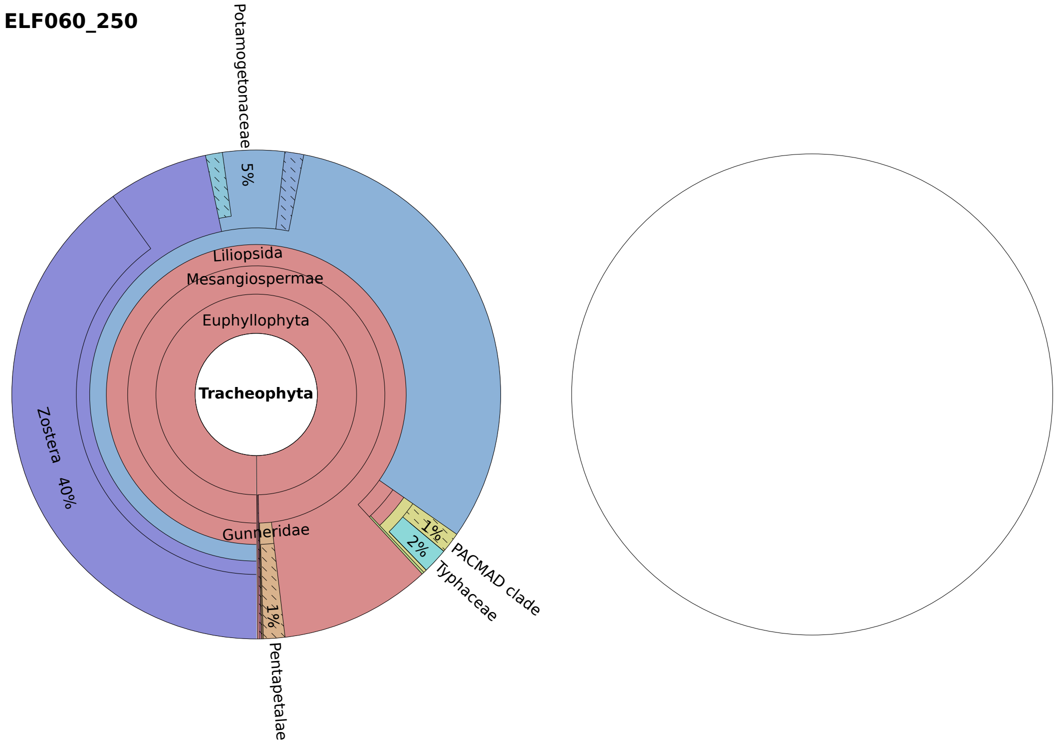


Figure 4.26: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF060, sample 250. See figure 4.4 for colour key.

Table 4.13: Biogenomic masses of European taxa in Embryophyta from ELF060. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	465	420	350	250
		Ingroup	4720.89	549.83	1828.77	1644.94	697.34
Freshwater aquatics	Water-milfoil family	Haloragaceae	2.31	0.00	0.00	2.31	0.00
Freshwater aquatics	Water-milfoils	Myriophyllum	9.23	2.31	0.00	2.31	4.62
Freshwater aquatics	Water-plantain family	Alismataceae	0.35	0.00	0.14	0.14	0.07
Freshwater aquatics	Duckweed family	Lemnoideae	1.22	0.00	0.00	0.00	1.22
Freshwater aquatics	Elodea family	Hydrocharitaceae	0.99	0.20	0.59	0.20	0.00
Freshwater aquatics	Water-nymphs	Najas	0.09	0.00	0.09	0.00	0.00
Freshwater aquatics	Pondweeds	Potamogeton	39.89	7.55	12.40	16.17	3.77
Freshwater aquatics	Pipeworts	Eriocaulon	0.38	0.00	0.38	0.00	0.00
Freshwater aquatics	Bulrush family	Typhaceae	29.13	2.91	8.74	5.83	11.65
Freshwater aquatics	Bulrushes	Typha	7.55	0.00	0.00	7.55	0.00
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	10.45	0.00	2.79	3.48	4.18
Salt/brackish aquatics	Tasselweeds	Ruppia	13.30	2.15	2.58	5.58	3.00
Salt/brackish aquatics	Eelgrass family	Zosteraceae	238.53	20.18	88.07	84.40	45.87
Salt/brackish aquatics	Eelgrasses	Zostera	1765.14	216.51	752.29	519.27	277.06
Mixed aquatics	Alismatids	Alismatales	1469.16	178.91	578.16	493.63	218.46
Mixed aquatics	Pondweed family	Potamogetonaceae	294.05	34.21	87.77	143.80	28.26
Mixed aquatics	Pondweeds	Stuckenia	33.22	4.96	7.93	16.36	3.97
Mixed aquatics	Reeds	Arundinoideae	6.01	0.71	1.41	2.48	1.41
Mixed aquatics	Common reed tribe	Molinieae	37.02	5.64	10.22	15.51	5.64
Mixed aquatics	Common reeds	Phragmites	2.94	0.42	0.84	1.68	0.00
Trees/shrubs	Birch family	Betulaceae	2.16	1.08	0.00	0.00	1.08
Trees/shrubs	Oaks	Quercus	1.06	0.00	0.00	0.00	1.06
Trees/shrubs	Willow family	Salicaceae	16.25	3.61	1.81	10.83	0.00
Trees/shrubs	Willow tribe	Saliceae	9.08	1.82	0.00	5.45	1.82
Trees/shrubs	Poplars	Populus	1.98	0.00	0.00	1.98	0.00
Trees/shrubs	Willows	Salix	3.46	0.00	1.73	1.73	0.00
Herbs	Mint tribe	Mentheae	1.23	1.23	0.00	0.00	0.00
Herbs	**Bridal-wreath tribe**	**Spiraeeae**	**1.87**	**0.00**	**1.87**	**0.00**	**0.00**
Herbs	Arum family	Araceae	0.55	0.18	0.00	0.18	0.18
Herbs	Orchids	Orchidaceae	0.68	0.17	0.17	0.34	0.00
Herbs		Epidendroideae	2.23	0.00	0.64	1.59	0.00
Herbs		Dioscoreales	1.87	0.00	0.93	0.00	0.93
Herbs		Dioscorea	1.81	0.00	0.90	0.90	0.00
Herbs	Buttercup subfamily	Ranunculoideae	0.22	0.00	0.00	0.00	0.22
Herbs	Baneberries	Actaea	0.10	0.00	0.10	0.00	0.00
Trees/shrubs and herbs		Gunneridae	0.47	0.00	0.00	0.00	0.47
Trees/shrubs and herbs		campanulids	0.31	0.00	0.00	0.31	0.00
Trees/shrubs and herbs		Asterales	0.57	0.00	0.29	0.29	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	0.27	0.00	0.00	0.00	0.27
Trees/shrubs and herbs		Asteroideae	0.25	0.25	0.00	0.00	0.00
Trees/shrubs and herbs		lamiids	0.68	0.68	0.00	0.00	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	1.76	0.00	0.88	0.88	0.00
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	1.11	0.56	0.56	0.00	0.00
Trees/shrubs and herbs		Saxifragales	1.93	0.00	0.00	0.96	0.96
Trees/shrubs and herbs		Asparagales	0.24	0.08	0.00	0.08	0.08
Trees/shrubs and herbs	**Palms**	**Arecaceae**	**0.29**	**0.00**	**0.00**	**0.29**	**0.00**
Trees/shrubs and herbs		Liliales	0.07	0.00	0.03	0.00	0.03
Trees/shrubs and herbs		**Proteales**	**0.32**	**0.00**	**0.00**	**0.00**	**0.32**
Grasses and relatives	Grass order	Poales	6.73	0.00	3.51	1.76	1.46
Grasses and relatives		PACMAD clade	29.27	0.00	0.00	29.27	0.00
Grasses and relatives		**Panicoideae**	**4.86**	**0.00**	**1.08**	**1.62**	**2.16**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**5.19**	**0.40**	**2.79**	**1.60**	**0.40**
Grasses and relatives		**Paniceae**	**8.04**	**1.34**	**6.03**	**0.67**	**0.00**
Ferns	Ferns	Polypodiopsida	0.14	0.00	0.00	0.07	0.07
Ferns	Leptosporangiate ferns	Polypodiidae	0.41	0.00	0.00	0.25	0.16
Ferns	Polypod ferns	Polypodiales	0.64	0.08	0.16	0.08	0.32
Ferns	Spleenwort suborder	Aspleniineae	0.90	0.00	0.36	0.36	0.18
Ferns		Thelypteridaceae	0.10	0.00	0.00	0.10	0.00
Ferns	Bracken family	Dennstaedtiaceae	0.10	0.00	0.00	0.00	0.10
Ferns		Polypodiaceae	0.08	0.00	0.08	0.00	0.00
Mixed		Tracheophyta	5.02	0.35	1.38	2.77	0.52
Mixed		Euphyllophyta	2.58	0.34	0.52	0.86	0.86
Mixed	Monocots	Liliopsida	604.97	59.08	223.41	253.32	69.15

Table 4.14: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF060.

Ecological category	Group	Common name	Taxon	Total	465	420	350	250
			Ingroup	2.18	0.00	0.76	1.43	0.00
Salt/brackish aquatics	Hydrozoans		Corynidae	1.43	0.00	0.00	1.43	0.00
Mixed	Birds	Perching birds	Passeriformes	0.76	0.00	0.76	0.00	0.00

Embryophyta

All ELF060 samples are again dominated by *Zostera* and Alismatales alongside small but consistent reeds and *Ruppia*. Freshwater aquatics are less frequent but quite diverse. This is in contrast to the terrestrial taxa. The samples all have relatively high read counts, but the richness of terrestrial plants in ELF060 is noticeably lower than a previous core with similar data yield, ELF059. Woody taxa are limited to the willow group with small signals from Betulaceae and *Quercus*, although Malvoideae could potentially be interpreted as *Tilia*. Herbs and grasses are similarly reduced.

Metazoa

There are only two Metazoa taxa: a passerine bird in sample 420 and a marine cnidarian in 350.

4.7.3 Pianka scores

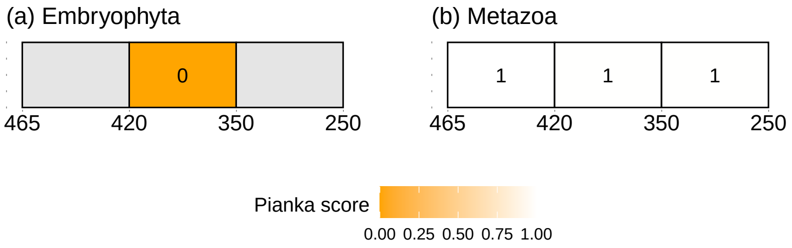


Figure 4.27: Pianka similarity scores between adjacent samples in ELF060 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores. Comparisons involving empty samples have no score.

The Embryophyta samples appear homogeneous, which is consistent with the very similar taxonomic profiles. The two samples with Metazoa data had a single and different taxon each, so a score of 0 is appropriate.

4.7.4 Summary

ELF060 appears overwhelmingly aquatic. The combination of saltwater, brackish, and freshwater taxa strongly suggests an estuarine environment. Other plants are surprisingly low in number and biogenomic mass, but they do demonstrate some terrestrial influence. Metazoa are uninformative apart from one marine taxon. These sediments appear to record a later stage of inundation than most other cores.

4.8 ELF032A

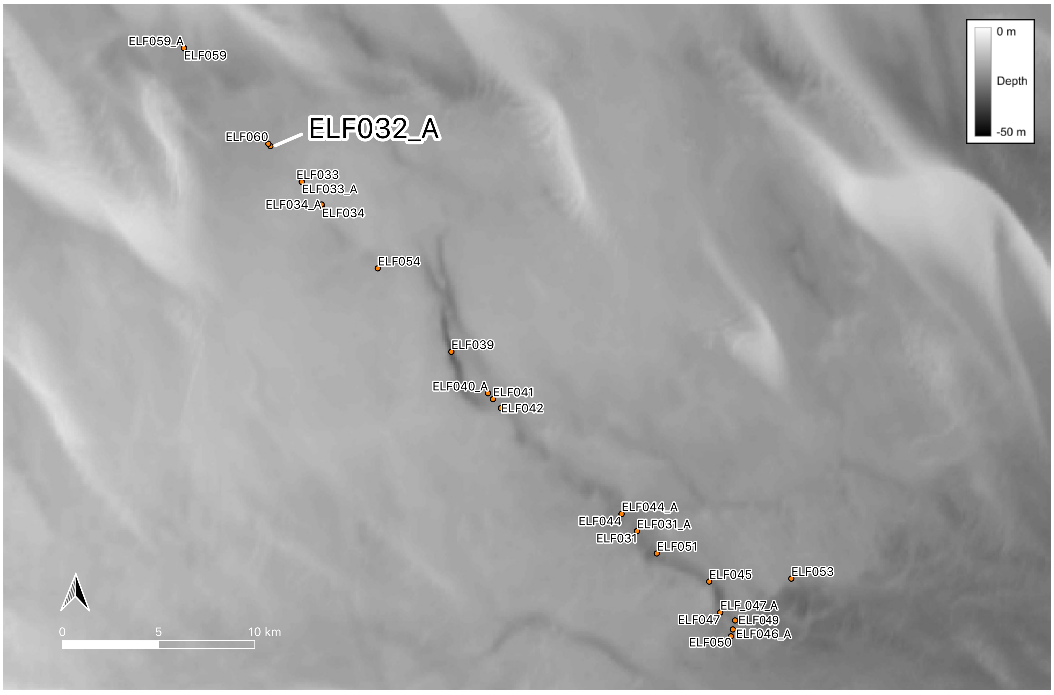


Figure 4.28: Bathymetry map of the palaeochannel transect highlighting ELF032A. Darker grey indicates greater depth.

ELF032A ~10 km northwest of the visible palaeochannel and very close to ELF060.

4.8.1 Read counts

ELF032A is one of the shorter cores at only 190 cm. There is variation in read count for Embryophyta, but no clear pattern and all counts are relatively high. Metazoa counts are <10 for most samples, but the oldest has nearly 80. This is surprisingly high for Metazoa, especially as the Embryophyta and Metazoa count plots usually show a similar pattern, despite the difference of scale. Though sample 177 has the highest Embryophyta count, it is not such an outlier.

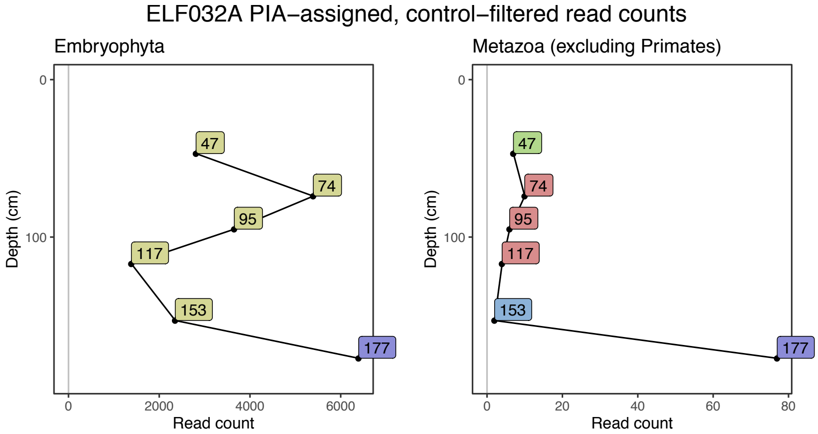


Figure 4.29: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF032A. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.8.2 Taxonomic profiles

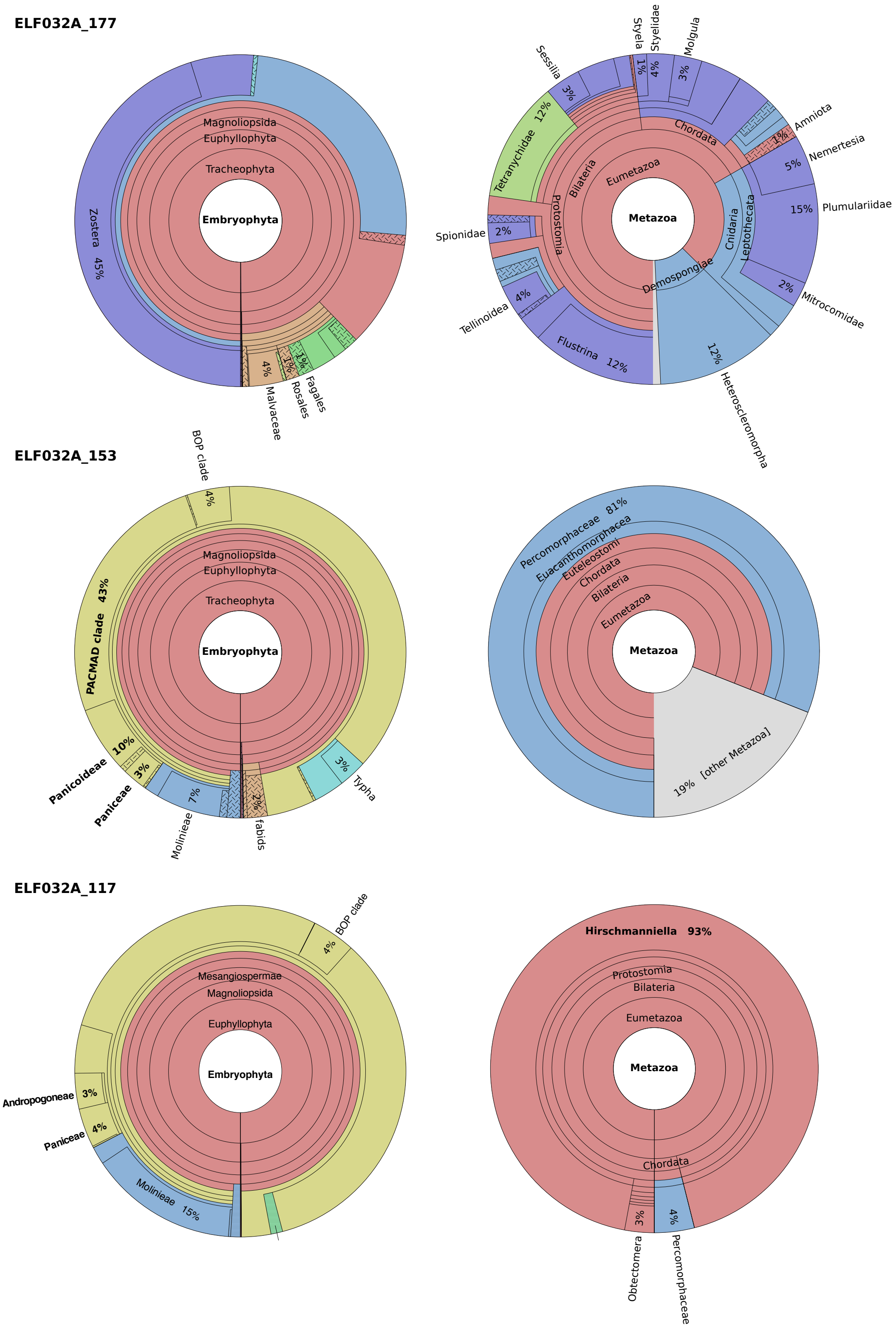


Figure 4.30: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF032A, samples 177, 153, and 117. Continued in figure 4.31. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

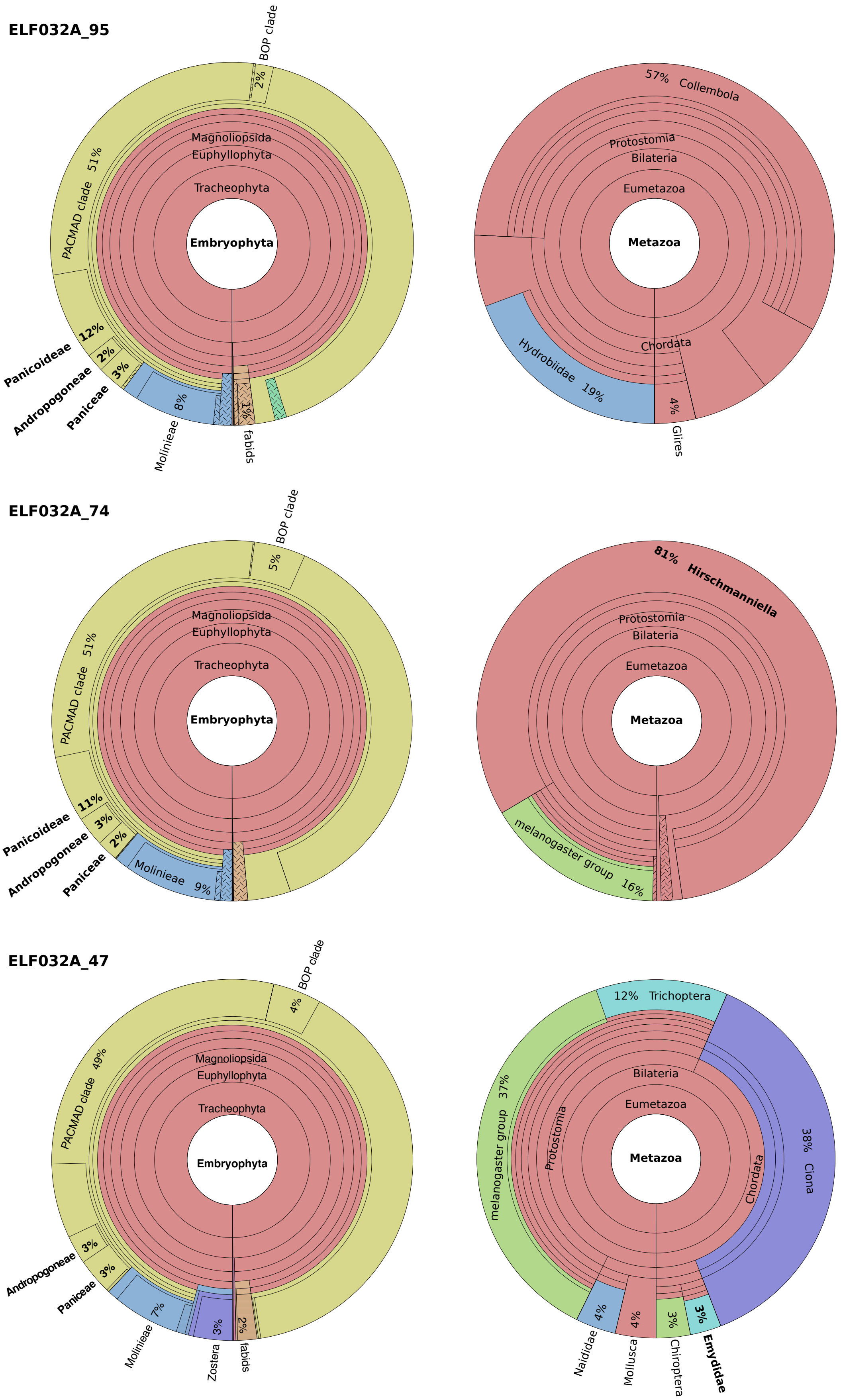


Figure 4.31: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF032A, samples 95, 74, and 47. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

Table 4.15: Biogenomic masses of European taxa in Embryophyta from ELF032A. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	177	153	117	95	74	47
		Ingroup	7041.38	2230.90	724.28	440.79	1120.78	1635.40	889.23
Freshwater aquatics	Water-plantain family	Alismataceae	0.14	0.00	0.00	0.00	0.14	0.00	0.00
Freshwater aquatics	Duckweed family	Lemnoideae	2.44	2.44	0.00	0.00	0.00	0.00	0.00
Freshwater aquatics	Pondweeds	Potamogeton	1.62	0.00	0.54	0.00	0.00	0.00	1.08
Freshwater aquatics	Rice tribe	Oryzeae	1.04	0.00	0.00	0.00	1.04	0.00	0.00
Freshwater aquatics	Bulrush family	Typhaceae	23.30	0.00	20.39	2.91	0.00	0.00	0.00
Freshwater aquatics	Bulrushes	Typha	30.19	3.77	18.87	0.00	7.55	0.00	0.00
Freshwater aquatics	Waterlily order	Nymphaeales	0.57	0.00	0.00	0.00	0.00	0.57	0.00
Freshwater aquatics	Waterlily family	Nymphaeaceae	1.06	0.00	0.53	0.00	0.53	0.00	0.00
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	0.70	0.70	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tasselweeds	Ruppia	1.29	0.86	0.00	0.00	0.00	0.00	0.43
Salt/brackish aquatics	**Neptune-grasses**	**Posidonia**	**0.32**	**0.32**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Salt/brackish aquatics	Eelgrass family	Zosteraceae	143.12	135.78	0.00	1.83	0.00	1.83	3.67
Salt/brackish aquatics	Eelgrasses	Zostera	1062.39	1001.83	7.34	1.83	11.01	12.84	27.52
Mixed aquatics	Alismatids	Alismatales	548.96	548.96	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Pondweed family	Potamogetonaceae	6.94	2.98	0.99	0.00	0.99	0.99	0.99
Mixed aquatics	Pondweeds	Stuckenia	1.98	1.98	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Reeds	Arundinoideae	57.65	0.71	9.55	6.72	14.50	18.39	7.78
Mixed aquatics	Common reed tribe	Molinieae	352.53	5.99	43.01	56.40	74.74	124.79	47.59
Mixed aquatics	Common reeds	Phragmites	25.21	0.42	5.04	0.84	5.04	7.56	6.30
Halophytes		Salicornia subg. Salicornia	0.72	0.72	0.00	0.00	0.00	0.00	0.00
Halophytes	Leadwort family	Plumbaginaceae	1.03	1.03	0.00	0.00	0.00	0.00	0.00
Halophytes	Sea-lavenders	Limonium	0.64	0.64	0.00	0.00	0.00	0.00	0.00
Xerophytes	Goosefoot family	Chenopodiaceae	2.39	2.39	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Elder family	Adoxaceae	0.13	0.13	0.00	0.00	0.00	0.00	0.00
Trees/shrubs		Viburnum	0.53	0.00	0.00	0.00	0.26	0.26	0.00
Trees/shrubs	Birch family	Betulaceae	24.84	19.44	2.16	0.00	2.16	1.08	0.00
Trees/shrubs	Alders	Alnus	1.61	1.61	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Birches	Betula	0.97	0.97	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Hazels	Corylus	11.36	4.55	2.27	0.00	0.00	4.55	0.00
Trees/shrubs	Oaks	Quercus	6.35	6.35	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Willow family	Salicaceae	66.79	52.35	1.81	0.00	5.42	5.42	1.81
Trees/shrubs	Willow tribe	Saliceae	38.12	29.05	0.00	0.00	1.82	1.82	5.45
Trees/shrubs	Poplars	Populus	9.89	5.93	0.00	0.00	1.98	0.00	1.98
Trees/shrubs	Willows	Salix	34.60	24.22	3.46	0.00	3.46	3.46	0.00
Trees/shrubs		Amygdaleae	2.21	0.00	0.00	0.00	0.00	0.00	2.21
Trees/shrubs	Stone fruit trees	Prunus	25.57	20.09	1.83	0.00	0.00	1.83	1.83
Trees/shrubs	Apple tribe	Maleae	3.57	3.57	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Elm family	Ulmaceae	3.01	2.01	0.00	0.00	0.00	0.50	0.50
Trees/shrubs	Limes	Tilia	4.00	4.00	0.00	0.00	0.00	0.00	0.00
Herbs		Rubioideae	0.72	0.72	0.00	0.00	0.00	0.00	0.00
Herbs	Mint tribe	Mentheae	2.46	0.00	0.00	0.00	2.46	0.00	0.00
Herbs	Nightshades	Solanum	0.71	0.71	0.00	0.00	0.00	0.00	0.00
Herbs	Hemp family	Cannabaceae	0.44	0.00	0.44	0.00	0.00	0.00	0.00
Herbs	Saxifrages	Saxifraga	1.39	1.39	0.00	0.00	0.00	0.00	0.00

Continued on next page

Table 4.15 continued

Ecological category	Common name	Taxon	Total	177	153	117	95	74	47
		Ingroup	7041.38	2230.90	724.28	440.79	1120.78	1635.40	889.23
Trees/shrubs and herbs		Gunneridae	0.94	0.47	0.00	0.00	0.47	0.00	0.00
Trees/shrubs and herbs		campanulids	0.92	0.31	0.62	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Carrot order	Apiales	0.93	0.47	0.00	0.00	0.47	0.00	0.00
Trees/shrubs and herbs	Ivy family	Araliaceae	1.51	0.00	0.00	0.00	0.50	0.00	1.01
Trees/shrubs and herbs		Asterales	1.43	0.86	0.57	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	0.27	0.27	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	0.37	0.37	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Aster tribe	Astereae	1.04	1.04	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Ericales	0.36	0.36	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		lamiids	0.68	0.68	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	0.88	0.00	0.88	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Oxalidales	0.48	0.48	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Rose subfamily	Rosoideae	4.17	2.78	1.39	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow order	Malvales	2.34	2.34	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	73.74	73.09	0.00	0.00	0.65	0.00	0.00
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	2.78	2.78	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Saxifragales	0.77	0.19	0.39	0.00	0.00	0.00	0.19
Trees/shrubs and herbs		Asparagales	0.08	0.08	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Liliales	0.03	0.03	0.00	0.00	0.00	0.00	0.00
Grasses and relatives	Grass order	Poales	123.55	3.22	31.91	11.13	18.74	57.68	0.88
Grasses and relatives	Sedges	Cyperaceae	6.17	1.54	1.54	1.54	1.54	0.00	0.00
Grasses and relatives		Cyperoideae	3.37	0.00	0.00	0.00	1.69	0.00	1.69
Grasses and relatives	Rushes	Juncus	1.52	1.52	0.00	0.00	0.00	0.00	0.00
Grasses and relatives	Grass family	Poaceae	1709.96	0.00	258.68	133.33	432.70	571.48	313.77
Grasses and relatives		BOP clade	164.74	0.00	28.32	16.42	16.76	68.88	34.36
Grasses and relatives	Rice subfamily	Oryzoideae	1.91	0.00	0.96	0.00	0.96	0.00	0.00
Grasses and relatives	False brome grasses	Brachypodium	1.82	0.00	0.00	0.00	0.00	1.82	0.00
Grasses and relatives	Barley subtribe	Hordeinae	0.24	0.00	0.00	0.12	0.00	0.12	0.00
Grasses and relatives	Barlies	Hordeum	0.13	0.00	0.13	0.00	0.00	0.00	0.00
Grasses and relatives		PACMAD clade	1274.45	0.00	173.63	108.64	309.56	451.44	231.18
Grasses and relatives		Chloridoideae	1.44	0.48	0.00	0.00	0.00	0.96	0.00
Grasses and relatives	Dropseeds	Sporobolus	1.90	1.90	0.00	0.00	0.00	0.00	0.00
Grasses and relatives		**Panicoideae**	**287.17**	**1.62**	**43.72**	**18.35**	**81.51**	**89.06**	**52.90**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**95.03**	**0.00**	**4.39**	**13.58**	**16.77**	**39.93**	**20.36**
Grasses and relatives		**Paniceae**	**109.26**	**0.00**	**15.42**	**14.75**	**24.80**	**28.82**	**25.47**
Grasses and relatives	**Barnyard grasses**	**Echinochloa**	**0.75**	**0.00**	**0.00**	**0.00**	**0.75**	**0.00**	**0.00**
Grasses and relatives		**Cenchrinae**	**2.30**	**0.00**	**0.58**	**0.00**	**1.15**	**0.00**	**0.58**
Grasses and relatives		**Melinidinae**	**0.87**	**0.00**	**0.00**	**0.00**	**0.87**	**0.00**	**0.00**
Grasses and relatives	**Panicgrass**	**Panicum**	**3.86**	**0.00**	**1.29**	**0.64**	**0.00**	**0.64**	**1.29**
Ferns	Ferns	Polypodiopsida	0.27	0.00	0.14	0.00	0.00	0.00	0.14
Ferns	Leptosporangiate ferns	Polypodiidae	0.41	0.08	0.08	0.00	0.08	0.08	0.08
Ferns	Polypod ferns	Polypodiales	0.97	0.16	0.00	0.00	0.08	0.32	0.40
Ferns	Spleenwort suborder	Aspleniineae	1.80	0.09	0.36	0.09	0.18	0.36	0.72
Ferns	Spleenworts	Asplenium	0.09	0.00	0.09	0.00	0.00	0.00	0.00
Ferns		Thelypteridaceae	0.10	0.00	0.00	0.00	0.00	0.00	0.10
Ferns		Phegopteridoideae	0.28	0.00	0.00	0.00	0.00	0.00	0.28
Ferns		Dryopteridoideae	0.08	0.00	0.08	0.00	0.00	0.00	0.00
Ferns		Polypodiaceae	0.08	0.08	0.00	0.00	0.00	0.00	0.00

Continued on next page

Table 4.15 continued

Ecological category	Common name	Taxon	Total	177	153	117	95	74	47
		Ingroup	7041.38	2230.90	724.28	440.79	1120.78	1635.40	889.23
Ferns		Pteridaceae	0.11	0.00	0.11	0.00	0.00	0.00	0.00
Bryophytes	Feather mosses	Hypnales	2.24	2.24	0.00	0.00	0.00	0.00	0.00
Bryophytes	Liverworts	Marchantiophyta	1.06	0.00	0.00	0.00	0.00	0.53	0.53
Mixed		Tracheophyta	1.38	0.52	0.17	0.00	0.17	0.17	0.35
Mixed		Euphyllophyta	2.93	1.03	0.69	0.34	0.69	0.17	0.00
Mixed	Monocots	Liliopsida	229.71	229.71	0.00	0.00	0.00	0.00	0.00

Table 4.16: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF032A. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	177	153	117	95	74	47
			Ingroup	242.18	99.93	1.30	53.68	7.61	66.38	13.29
Freshwater aquatics	**Testudines**	**Terrapins**	**Emydidae**	**0.37**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.37**
Freshwater aquatics	Trichopterans	Caddisflies	Trichoptera	1.58	0.00	0.00	0.00	0.00	0.00	1.58
Salt/brackish aquatics	Tunicates	Sea squirts	Ascidiacea	3.45	3.45	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tunicates	Vase tunicates	Ciona	5.00	0.00	0.00	0.00	0.00	0.00	5.00
Salt/brackish aquatics	Tunicates	Lightbulb sea squirts	Stolidobranchia	4.05	4.05	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tunicates	Sea grapes	Molgula	2.70	2.70	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tunicates		Styelidae	2.70	2.70	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tunicates		Styela	1.35	1.35	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Barnacles	Barnacles	Cirripedia	1.54	1.54	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Barnacles		Thoracica	3.64	3.64	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Barnacles	Acorn barnacles	Sessilia	3.48	3.48	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Annelids		Cirratulidae	0.77	0.77	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Annelids		Spionidae	2.00	2.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bryozoans		Cheilostomatida	2.45	2.45	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bryozoans		Flustrina	12.24	12.24	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bivalves		Tellinoidea	3.05	3.05	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bivalves		Abra	0.51	0.51	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bivalves		Mytilus	0.61	0.61	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Hydrozoans		Mitrocomidae	2.44	2.44	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Hydrozoans		Plumulariidae	9.76	9.76	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Hydrozoans		Nemertesia	4.88	4.88	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Euteleosteomorpha	0.90	0.90	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Euacanthomorphacea	1.06	1.06	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Percomorphaceae	4.22	1.06	1.06	2.11	0.00	0.00	0.00
Mixed aquatics	Birds	Waterfowl	Anatidae	0.72	0.72	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Copepods		Calanoida	0.22	0.00	0.00	0.00	0.00	0.22	0.00
Mixed aquatics	Annelids	Detritus worms	Naididae	0.48	0.00	0.00	0.00	0.00	0.00	0.48
Mixed aquatics	Bivalves	Bivalves	Bivalvia	1.16	1.16	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Bivalves		Euheterodonta	0.58	0.58	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Bivalves	Mussels	Mytilidae	0.56	0.56	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Gastropods	Mud snails	Hydrobiidae	1.47	0.00	0.00	0.00	1.47	0.00	0.00

Continued on next page

Table 4.16 continued

Ecological category	Group	Common name	Taxon	Total	177	153	117	95	74	47
			Ingroup	242.18	99.93	1.30	53.68	7.61	66.38	13.29
Mixed aquatics	Cnidarians	Cnidarians	Cnidaria	1.17	1.17	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Hydrozoans	Thecate hydroids	Leptothecata	2.44	2.44	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Poriferans		Heteroscleromorpha	11.86	11.86	0.00	0.00	0.00	0.00	0.00
Terrestrial	Mammals	Old World rats and mice	Murinae	0.31	0.31	0.00	0.00	0.00	0.00	0.00
Terrestrial	Mammals	Grey wolf	Canis lupus	0.36	0.36	0.00	0.00	0.00	0.00	0.00
Terrestrial	Mammals	Bats	Chiroptera	0.41	0.00	0.00	0.00	0.00	0.00	0.41
Terrestrial	Mites	Spider mites	Tetranychidae	11.76	11.76	0.00	0.00	0.00	0.00	0.00
Terrestrial	Dipterans		melanogaster group	14.86	0.00	0.00	0.00	0.00	9.90	4.95
Mixed	Animals	Animals	Metazoa	0.99	0.74	0.25	0.00	0.00	0.00	0.00
Mixed	Animals		Eumetazoa	0.25	0.00	0.00	0.00	0.00	0.25	0.00
Mixed	Animals	Chordates	Chordata	0.19	0.00	0.00	0.00	0.00	0.19	0.00
Mixed	Mammals	Rodents and lagomorphs	Glires	0.28	0.00	0.00	0.00	0.28	0.00	0.00
Mixed	Squamates	Colubrid snake superfamily	Colubroidea	0.46	0.00	0.00	0.00	0.00	0.46	0.00
Mixed	Invertebrates		Protostomia	1.04	0.00	0.00	0.00	0.52	0.52	0.00
Mixed	Invertebrates		Ecdysozoa	0.50	0.00	0.00	0.00	0.50	0.00	0.00
Mixed	**Nematodes**		**Hirschmanniella**	**100.00**	**0.00**	**0.00**	**50.00**	**0.00**	**50.00**	**0.00**
Mixed	Crustaceans	Crustaceans	Crustacea	0.19	0.19	0.00	0.00	0.00	0.00	0.00
Mixed	Crustaceans		Eumalacostraca	0.13	0.13	0.00	0.00	0.00	0.00	0.00
Mixed	Springtails	Springtails	Collembola	4.35	0.00	0.00	0.00	4.35	0.00	0.00
Mixed	Lepidopterans		Obtectomera	1.56	0.00	0.00	1.56	0.00	0.00	0.00
Mixed	Invertebrates	Lophotrochozoans	Lophotrochozoa	1.86	1.86	0.00	0.00	0.00	0.00	0.00
Mixed	Molluscs	Molluscs	Mollusca	2.44	1.46	0.00	0.00	0.49	0.00	0.49

Embryophyta

Aquatics

Freshwater aquatics are present in every sample but relatively scarce, with the exception of *Typha* and Typhaceae, which are among the most frequent taxa in sample 153. Saltwater aquatics (mostly *Zostera* with some *Ruppia*) are also present in all samples but only a substantial component of 177, the deepest: ELF032A is the first core discussed not dominated by *Zostera*. This contrasts with ELF060, taken in nearly the same location, which appears particularly aquatic.

Among mixed aquatics, Alismatales is only present in 177, which may be associated with the significant *Zostera* or greater read count. The reed group is present throughout and has greater biogenomic mass after 177.

Halophytes and xerophytes

A small number of halophytes and xerophytes have low biogenomic masses in 177 only. They may be rare taxa only recovered in this sample because of its greater data yield, but the correlation with *Zostera* suggests genuine environmental change, perhaps to a less saline environment.

Trees/shrubs

177 has a stronger woody signal than other samples. Even accounting for its greater data yield, biogenomic masses are higher in 177. However, the taxa involved are similar to the rest of the core: the willow group, *Corylus* and Betulaceae, and *Prunus*. The only clear taxonomic distinction for 177 is that its most frequent tree may be *Tilia*; the direct biogenomic mass is small, but there is a very large Malvaceae signal under Trees/shrubs and herbs. There is also a small signal from Malvoideae, which would suggest interpreting at least some Malvaceae as *Malva* instead. However, the estimated genome sizes of *Tilia* and *Malva* are similar (appendix D), so we can assume the same biogenomic mass from a given biomass, and it is easier to imagine large biomass from a tree than a small herb.

Grasses and relatives

While sample 177 shows the commonly-seen *Zostera*-dominated profile, other samples in ELF032A are dominated by grasses. This is unusual, but besides the increased biogenomic mass, the taxonomic profile of grasses is broadly similar to other samples. Higher taxa are most frequent and could be interpreted as wetland grasses, considering the prevalence of reeds and *Typha*. There is also a substantial signal from the common but somewhat questionable Panicoideae group. Other specific taxa are rare.

Metazoa

Aquatics

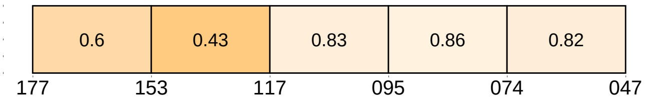
Every sample has mixed aquatic Metazoa, but the specific taxa are sharply divided. Saltwater aquatics, apart from a tunicate in 47, are restricted to sample 177 and are of high biogenomic mass. The only exclusively freshwater taxa are in sample 47. Both are rarely seen: Trichoptera (caddisflies) are in only two other cores and Emydidae (terrapins) nowhere else. The single European taxon in Emydidae, outside of Sicily, is *Emys orbicularis* (the European pond terrapin). *E. orbicularis* is limited to southern Europe today, where the summers are warm enough for egg incubation (Sommer *et al.* 2009). However, its range has fluctuated with climate, and two fossils confirm its presence Britain during the warmer Atlantic period (9100–5450 BP, Sommer *et al.* 2007 & 2009). The younger fossil was discovered in Norfolk, less than 100 km from ELF032A. Emydidae DNA is actually plausible, although the signal is weak.

Mixed

These animals are not particularly useful for reconstruction, but the extremely high biogenomic mass from the nematode *Hirschmanniella* deserves explanation. Only a single read was recovered in samples 117 and 74, but the estimated C-value of just 0.02 converted this to a biogenomic mass of 50. In reality, a single genus of nematode is unlikely to have been so ecologically dominant.

4.8.3 Pianka scores

(a) Embryophyta



(b) Metazoa

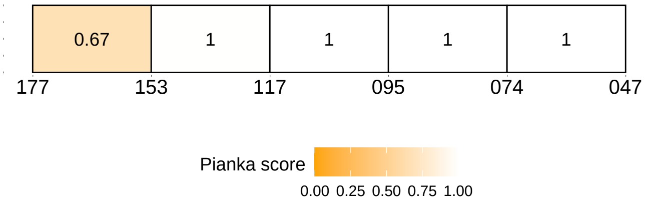


Figure 4.32: Pianka similarity scores between adjacent samples in ELF032A for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

The Embryophyta Pianka scores clearly reflect the profiles: sample 177 is very different from the others, which are themselves rather homogeneous. The Metazoa scores are more complex; most taxa are only present in 177, but the remainder are scattered across samples with low biogenomic masses, with the exception of *Hirschmanniella*. However, deeper samples do appear more distinct than those near the top.

4.8.4 Summary

ELF032A is an unusual core in many respects. The deepest sample shows a *Zostera*-dominated coastal environment with some brackish signal and floodplain trees, similar to many samples discussed previously. Subsequent samples abruptly change to a grass-dominated profile with strong reed and bulrush signals. The core may represent a previously open environment, perhaps an estuary, undergoing succession to reed bed. The marine-to-terrestrial transition implies that any inundation paused or even temporarily reversed at this location during this period.

The Metazoa data is unusually rich in the *Zostera*-dominated sample, particularly for saltwater taxa. The correlation between saltwater plants and animals and high read counts may be explained by the beneficial effect of salt on DNA preservation (Lindahl and Nyberg 1972, Kistler *et al.* 2017). Finally, the top sample contains one read attributed to *Emys orbicularis*. Although limited by climate to Southern Europe today, fossil evidence places this species near Doggerland during warmer parts of the Holocene.

4.9 ELF033

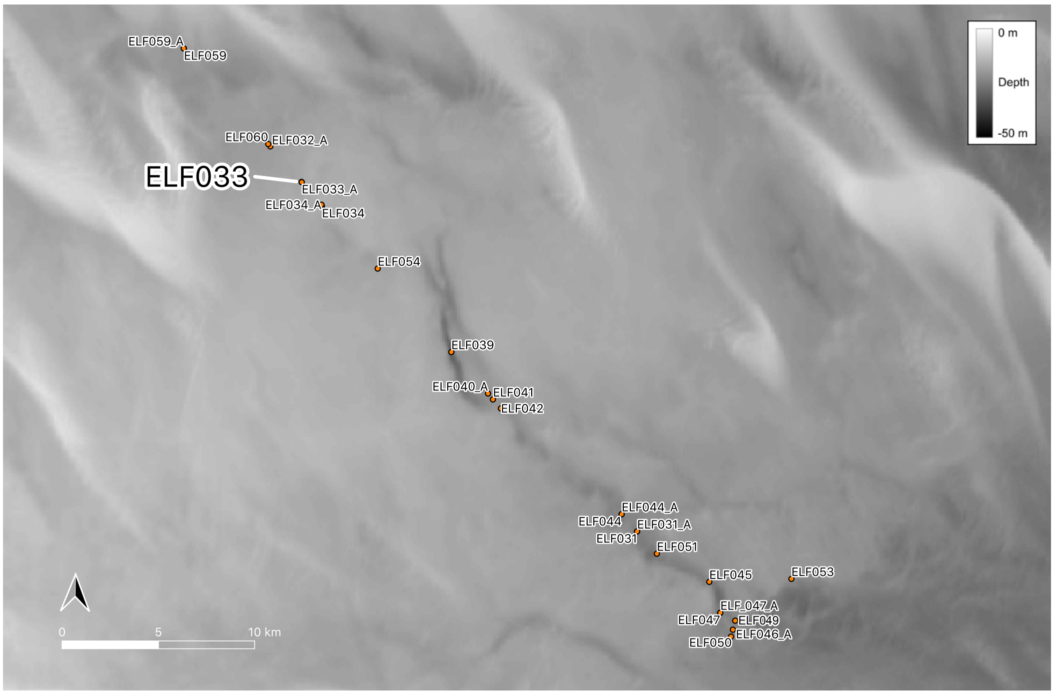


Figure 4.33: Bathymetry map of the palaeochannel transect highlighting ELF033. Darker grey indicates greater depth.

ELF033 is approximately half way between the start of the transect and the visible palaeochannel. ELF033A was taken at the same location.

4.9.1 Read counts

ELF033 was densely sampled and shows large variation in data recovery, making figure 4.34 challenging to read. Metazoa data is sparse. Most samples have relatively few Embryophyta reads, but sample 118 has over 3,000.

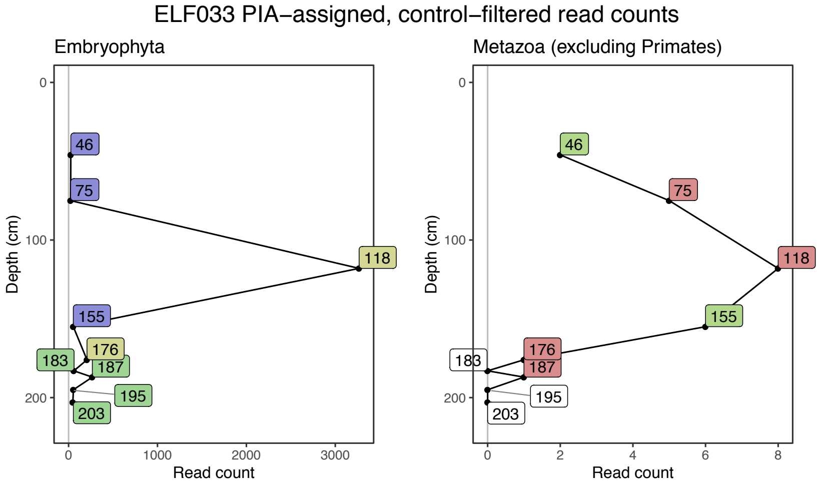
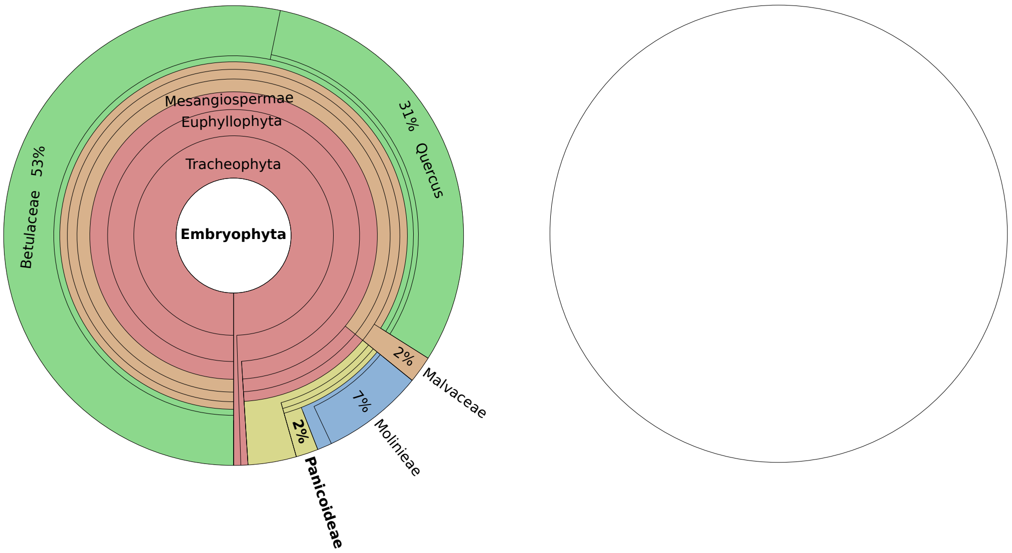


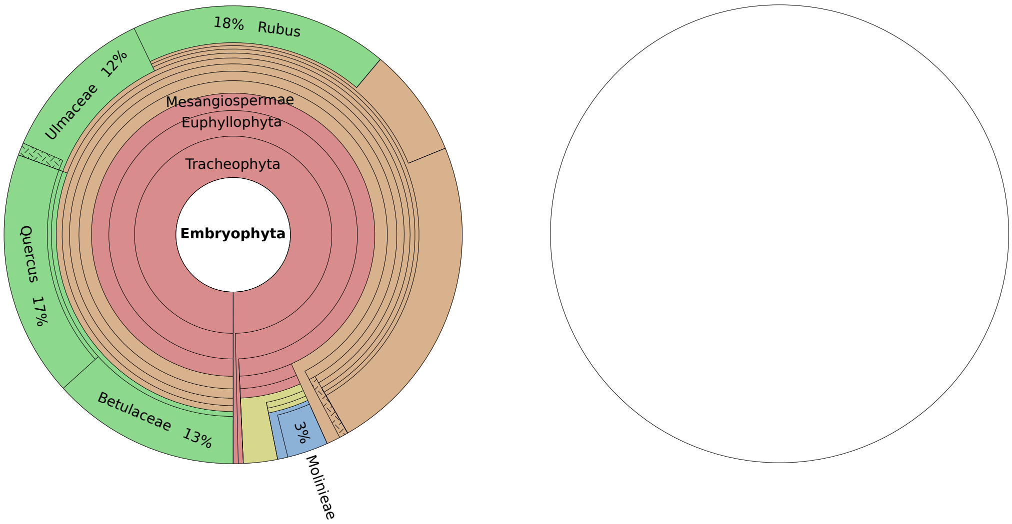
Figure 4.34: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF033. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4). Note that the label for 195 is attached with a grey line.

4.9.2 Taxonomic profiles

ELF033_203



ELF033_195



ELF033_187

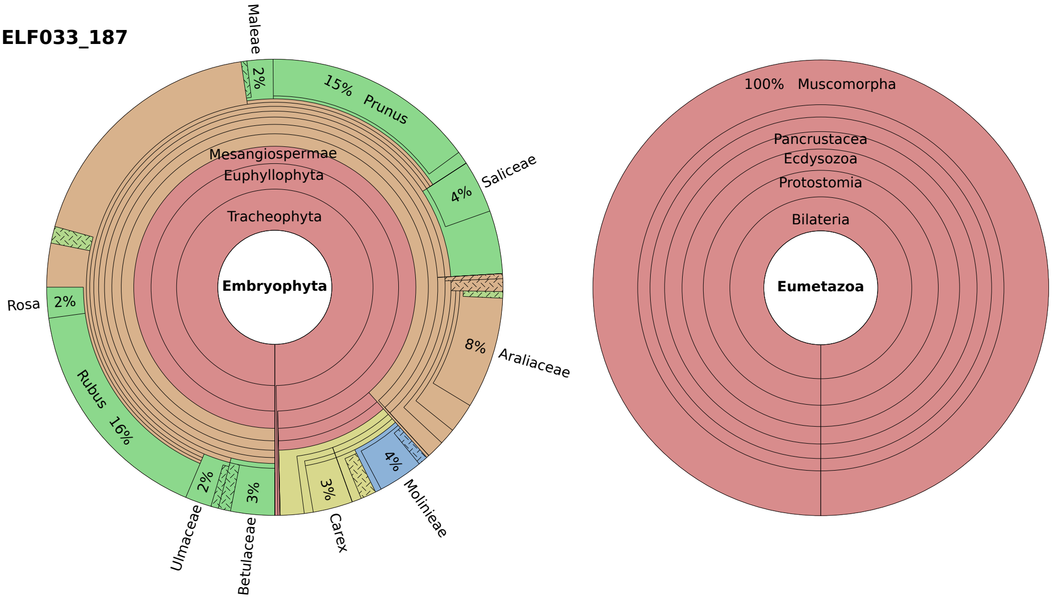


Figure 4.35: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF033, samples 203, 195, and 187. Continued in figure 4.36. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

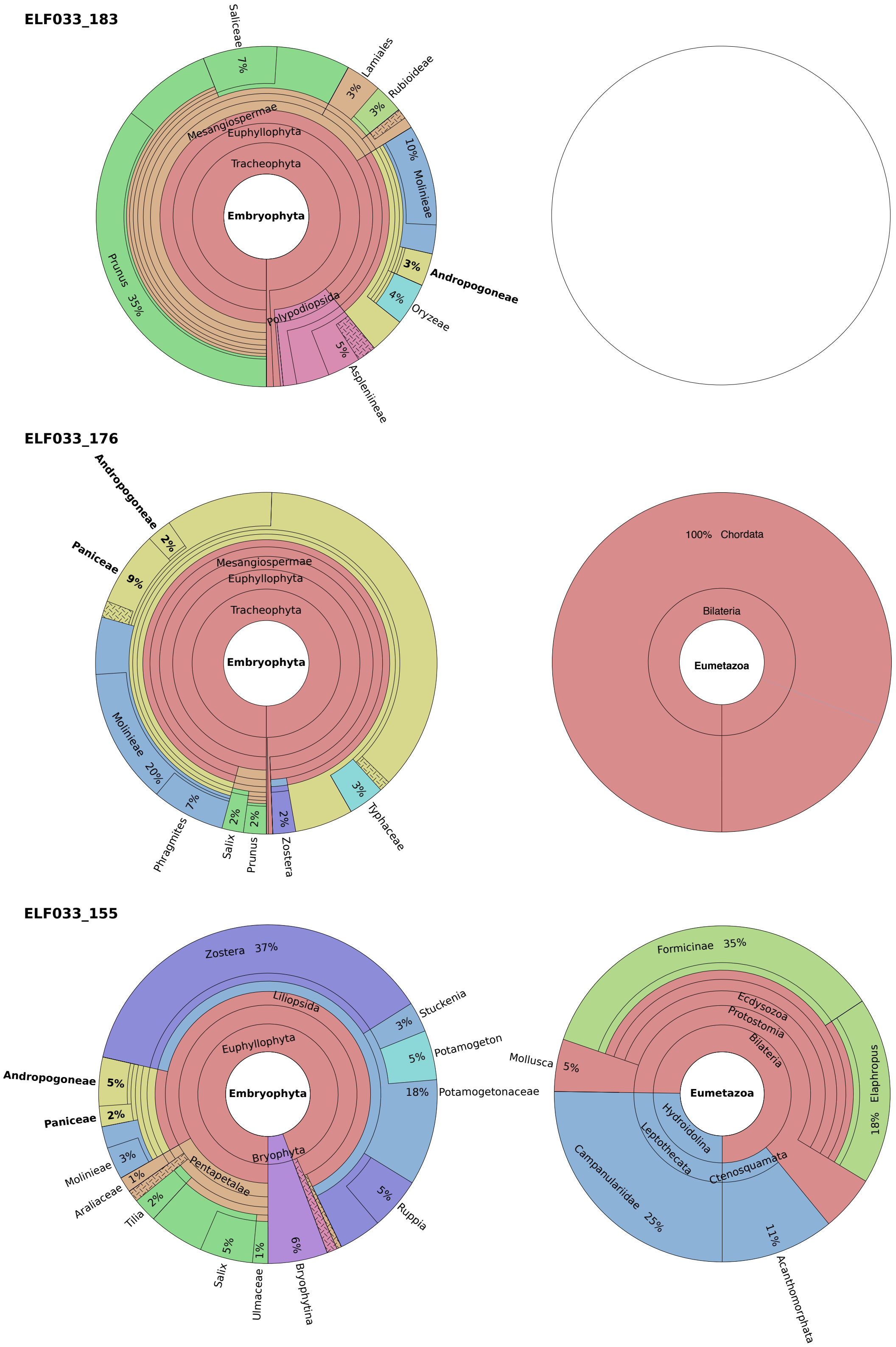
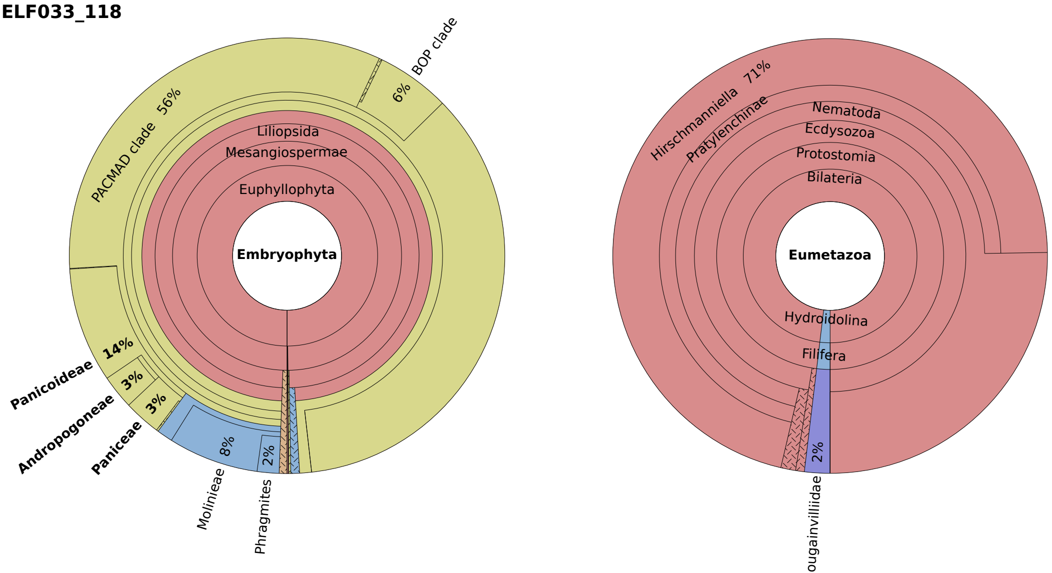
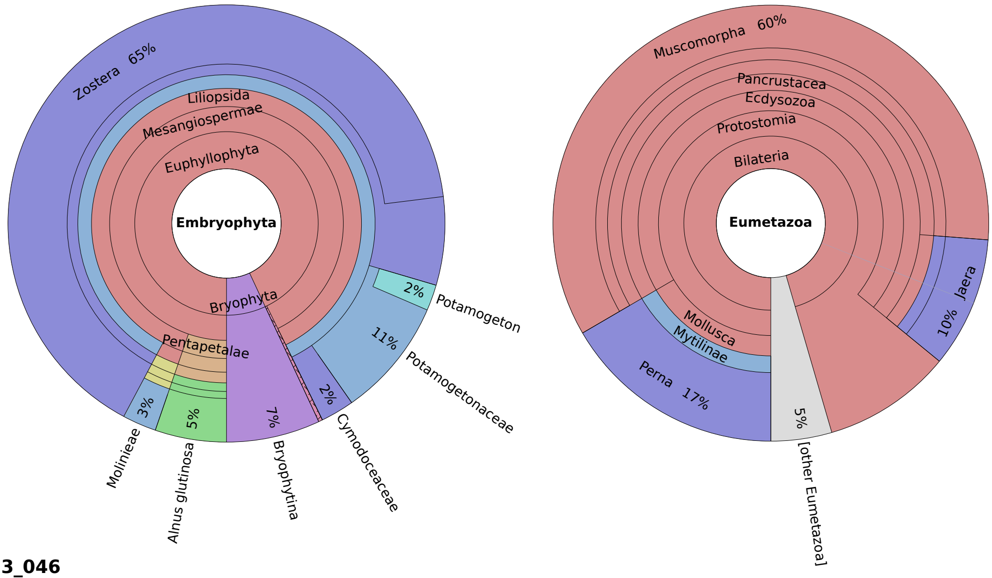


Figure 4.36: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF033, samples 183, 176, and 155. Continued in figure 4.37. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

ELF033_118



ELF033_075



ELF033_046

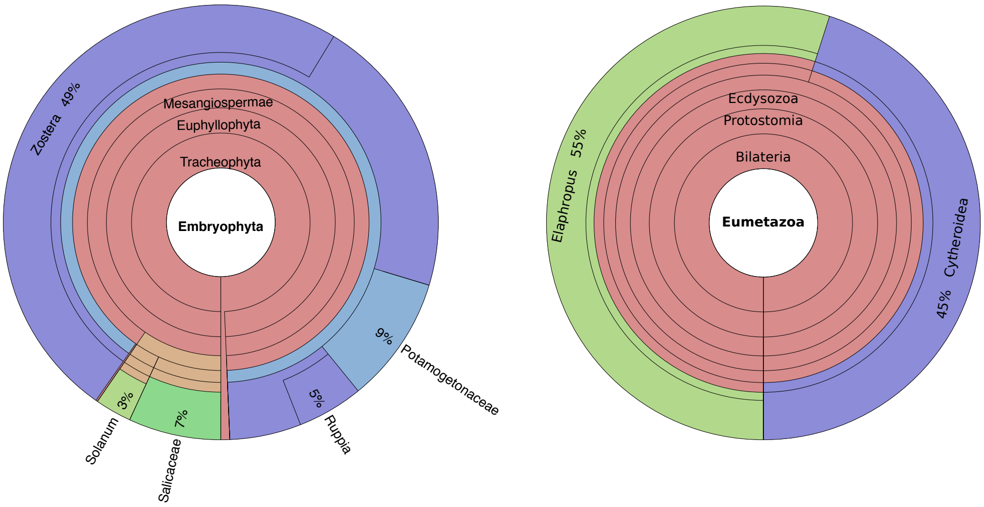


Figure 4.37: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF033, samples 118, 75, and 46. See figure 4.4 for colour key.

Table 4.17: Biogenomic masses of European taxa in Embryophyta from ELF033. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	203	195	187	183	176	155	118	75	46
		Ingroup	1584.94	37.43	49.31	250.29	26.99	105.48	38.42	1020.29	28.12	28.62
Freshwater aquatics	Coontails	Ceratophyllum	1.45	0.00	0.00	0.00	0.00	0.00	0.00	1.45	0.00	0.00
Freshwater aquatics	Water-milfoil family	Haloragaceae	2.31	0.00	0.00	0.00	0.00	0.00	0.00	2.31	0.00	0.00
Freshwater aquatics	Pondweeds	Potamogeton	2.16	0.00	0.00	0.00	0.00	0.00	1.62	0.00	0.54	0.00
Freshwater aquatics	Rice tribe	Oryzeae	1.04	0.00	0.00	0.00	1.04	0.00	0.00	0.00	0.00	0.00
Freshwater aquatics	Bulrush family	Typhaceae	2.91	0.00	0.00	0.00	0.00	2.91	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	3.48	0.00	0.00	0.00	0.00	0.00	1.39	0.00	0.70	1.39
Salt/brackish aquatics	Tasselweeds	Ruppia	3.00	0.00	0.00	0.00	0.00	0.00	1.72	0.00	0.00	1.29
Salt/brackish aquatics	Eelgrass family	Zosteraceae	7.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.83	5.50
Salt/brackish aquatics	Eelgrasses	Zostera	51.38	0.00	0.00	0.00	0.00	1.83	12.84	5.50	18.35	12.84
Mixed aquatics	Pondweed family	Potamogetonaceae	8.93	0.00	0.00	0.00	0.00	0.00	3.47	0.50	2.48	2.48
Mixed aquatics	Pondweeds	Stuckenia	0.99	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00
Mixed aquatics	Reeds	Arundinoideae	18.74	0.35	0.35	1.06	0.71	4.60	0.71	10.96	0.00	0.00
Mixed aquatics	Common reed tribe	Molinieae	89.19	2.47	1.41	7.76	2.47	10.93	1.06	62.40	0.71	0.00
Mixed aquatics	Common reeds	Phragmites	22.69	0.00	0.00	1.68	0.00	5.88	0.00	15.13	0.00	0.00
Trees/shrubs		Viburnum	1.05	0.00	0.00	1.05	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Dogwoods	Cornus	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.59	0.00	0.00
Trees/shrubs	Birch family	Betulaceae	32.40	18.36	6.48	7.56	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Black alder	Alnus glutinosa	1.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.49	0.00
Trees/shrubs	Oaks	Quercus	21.16	10.58	8.46	2.12	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Willow family	Salicaceae	16.25	0.00	0.00	10.83	1.81	0.00	0.00	1.81	0.00	1.81
Trees/shrubs	Willow tribe	Saliceae	12.71	0.00	0.00	9.08	1.82	0.00	1.82	0.00	0.00	0.00
Trees/shrubs	Willows	Salix	3.46	0.00	0.00	0.00	0.00	1.73	1.73	0.00	0.00	0.00
Trees/shrubs		Amygdaleae	4.42	0.00	0.00	2.21	2.21	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Stone fruit trees	Prunus	47.48	0.00	0.00	36.52	9.13	1.83	0.00	0.00	0.00	0.00
Trees/shrubs	Apple tribe	Maleae	4.46	0.00	0.00	4.46	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Apples	Malus	0.99	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Roses	Rosa	5.29	0.00	0.00	5.29	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Blackberry and relatives	Rubus	48.82	0.00	8.88	39.94	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Elm family	Ulmaceae	11.04	0.00	5.52	4.52	0.00	0.00	0.50	0.50	0.00	0.00
Trees/shrubs	Elms	Ulmus	1.72	0.00	0.43	1.29	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Limes	Tilia	0.80	0.00	0.00	0.00	0.00	0.00	0.80	0.00	0.00	0.00
Herbs	Ivies	Hedera	1.13	0.00	0.00	1.13	0.00	0.00	0.00	0.00	0.00	0.00
Herbs		Rubioideae	0.72	0.00	0.00	0.00	0.72	0.00	0.00	0.00	0.00	0.00
Herbs	Nightshades	Solanum	0.71	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.71
Herbs	Strawberry subtribe	Fragariinae	1.47	0.00	0.00	1.47	0.00	0.00	0.00	0.00	0.00	0.00
Herbs	Cinquefoils	Potentilla	1.42	0.00	0.00	1.42	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Gunneridae	0.94	0.00	0.47	0.47	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		campanulids	3.39	0.00	0.31	3.08	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Carrot order	Apiales	2.80	0.00	0.00	2.80	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Apiineae	5.13	0.00	0.00	5.13	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Ivy family	Araliaceae	19.69	0.00	0.00	19.18	0.00	0.00	0.50	0.00	0.00	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	0.54	0.00	0.00	0.27	0.27	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Asteroideae	0.25	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Dipsacales	0.87	0.00	0.00	0.87	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	1.76	0.00	0.00	0.88	0.88	0.00	0.00	0.00	0.00	0.00

Continued on next page

Table 4.17 continued

Ecological category	Common name	Taxon	Total	203	195	187	183	176	155	118	75	46
		Ingroup	1584.94	37.43	49.31	250.29	26.99	105.48	38.42	1020.29	28.12	28.62
Trees/shrubs and herbs	Rose subfamily	Rosoideae	55.64	0.00	11.13	44.51	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Rosoideae incertae sedis	11.29	0.00	3.76	7.53	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	0.65	0.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Saxifragales	0.58	0.00	0.00	0.00	0.00	0.00	0.39	0.19	0.00	0.00
Trees/shrubs and herbs		Liliales	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.16	0.00	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.00
Grasses and relatives	Grass order	Poales	20.20	1.17	1.17	4.10	0.88	4.68	0.00	8.20	0.00	0.00
Grasses and relatives	Sedges	Cyperaceae	1.54	0.00	0.00	1.54	0.00	0.00	0.00	0.00	0.00	0.00
Grasses and relatives	True sedges	Carex	6.83	0.00	0.00	6.83	0.00	0.00	0.00	0.00	0.00	0.00
Grasses and relatives	Grass family	Poaceae	330.90	0.00	0.00	0.00	0.00	0.00	0.00	330.90	0.00	0.00
Grasses and relatives		BOP clade	50.78	0.00	0.00	0.00	0.00	0.00	0.00	50.78	0.00	0.00
Grasses and relatives	False brome grasses	Brachypodium	1.82	0.00	0.00	0.00	0.00	0.00	0.00	1.82	0.00	0.00
Grasses and relatives		Aveninae	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.00
Grasses and relatives	**Feather grasses**	**Stipa**	**0.88**	**0.00**	**0.00**	**0.00**	**0.00**	**0.88**	**0.00**	**0.00**	**0.00**	**0.00**
Grasses and relatives	Barley subtribe	Hordeinae	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00
Grasses and relatives		PACMAD clade	337.34	0.00	0.00	0.00	0.00	31.25	0.00	306.09	0.00	0.00
Grasses and relatives		Chloridoideae	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.48	0.00	0.00
Grasses and relatives		**Panicoideae**	**90.14**	**0.54**	**0.00**	**1.62**	**0.00**	**8.64**	**0.00**	**79.35**	**0.00**	**0.00**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**29.95**	**0.00**	**0.00**	**2.00**	**0.80**	**2.00**	**1.60**	**23.56**	**0.00**	**0.00**
Grasses and relatives		**Paniceae**	**32.85**	**0.00**	**0.00**	**0.67**	**0.00**	**6.03**	**0.67**	**25.47**	**0.00**	**0.00**
Grasses and relatives	**Barnyard grasses**	**Echinochloa**	**0.75**	**0.00**	**0.00**	**0.00**	**0.00**	**0.75**	**0.00**	**0.00**	**0.00**	**0.00**
Grasses and relatives		**Cenchrinae**	**0.58**	**0.00**	**0.00**	**0.00**	**0.00**	**0.58**	**0.00**	**0.00**	**0.00**	**0.00**
Grasses and relatives	**Panicgrass**	**Panicum**	**1.29**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**1.29**	**0.00**	**0.00**
Ferns	Ferns	Polypodiopsida	0.07	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00
Ferns	Horsetails	Equisetum	0.09	0.00	0.00	0.05	0.00	0.00	0.05	0.00	0.00	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	0.41	0.00	0.00	0.00	0.33	0.00	0.00	0.00	0.08	0.00
Ferns	Polypod ferns	Polypodiales	0.88	0.00	0.00	0.00	0.80	0.00	0.08	0.00	0.00	0.00
Ferns	Spleenwort suborder	Aspleniineae	0.81	0.00	0.00	0.00	0.81	0.00	0.00	0.00	0.00	0.00
Ferns	Spleenworts	Asplenium	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00
Ferns		Athyriaceae	0.15	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00
Ferns		Thelypteridaceae	0.10	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00
Ferns		Phegopteridoideae	0.14	0.00	0.00	0.00	0.14	0.00	0.00	0.00	0.00	0.00
Ferns		Thelypteridoideae	0.17	0.00	0.00	0.00	0.08	0.00	0.08	0.00	0.00	0.00
Ferns	Marsh fern	Thelypteris palustris	0.13	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00
Bryophytes		Bryophytina	3.90	0.00	0.00	0.00	0.00	0.00	1.95	0.00	1.95	0.00
Mixed		Tracheophyta	1.38	0.17	0.17	0.52	0.17	0.17	0.00	0.00	0.00	0.17
Mixed		Euphyllophyta	2.24	0.17	0.17	0.34	0.17	0.34	0.00	1.03	0.00	0.00

Table 4.18: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF033. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	203	195	187	183	176	155	118	75	46
			Ingroup	92.78	0.00	0.00	3.24	0.00	0.19	9.67	71.08	5.43	3.17
Freshwater aquatics	Coleopterans	Water scavenger beetle family	Hydrophilidae	0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.77	0.00	0.00
Salt/brackish aquatics	Crustaceans		Jaera	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.53	0.00
Salt/brackish aquatics	Ostracods		Cytheroidea	1.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.43
Salt/brackish aquatics	Bivalves		Perna	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.90	0.00
Salt/brackish aquatics	Hydrozoans		Bougainvilliidae	1.33	0.00	0.00	0.00	0.00	0.00	0.00	1.33	0.00	0.00
Mixed aquatics	Fishes		Acanthomorphata	1.06	0.00	0.00	0.00	0.00	0.00	1.06	0.00	0.00	0.00
Mixed aquatics	Hydrozoans		Campanulariidae	2.44	0.00	0.00	0.00	0.00	0.00	2.44	0.00	0.00	0.00
Terrestrial	Coleopterans		Elaphropus	3.48	0.00	0.00	0.00	0.00	0.00	1.74	0.00	0.00	1.74
Terrestrial	Hymenopterans		Formicinae	3.42	0.00	0.00	0.00	0.00	0.00	3.42	0.00	0.00	0.00
Mixed	Animals		Eumetazoa	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00
Mixed	Animals	Chordates	Chordata	0.19	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.00	0.00
Mixed	Invertebrates		Protostomia	1.04	0.00	0.00	0.00	0.00	0.00	0.52	0.00	0.52	0.00
Mixed	Nematodes	Nematode worms	Nematoda	17.76	0.00	0.00	0.00	0.00	0.00	0.00	17.76	0.00	0.00
Mixed	**Nematodes**		**Hirschmanniella**	**50.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**50.00**	**0.00**	**0.00**
Mixed	Dipterans		Muscomorpha	6.47	0.00	0.00	3.24	0.00	0.00	0.00	0.00	3.24	0.00
Mixed	Molluscs	Molluscs	Mollusca	0.49	0.00	0.00	0.00	0.00	0.00	0.49	0.00	0.00	0.00
Mixed	Gastropods		Caenogastropoda	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.46	0.00	0.00

Embryophyta

Aquatics

The reed group is present in all but the top sample, whereas other aquatics only occur from sample 176 up. This is despite a relatively constant data yield, suggesting that the move from aquatics to reeds reflects environmental change.

Trees/shrubs

Most biogenomic mass in the four deepest samples is from woody taxa. The willow group are among the most frequent, but the signal is not as strong as in many other cores. Other prominent taxa include Betulaceae, *Quercus*, *Rubus* (blackberry and relatives), and *Ulmus* (elm) and Ulmaceae. Sample 187 also contains *Hedera* (ivies), classed here as a herb but supporting a woodland signal. The woody signal is proportionally smaller from sample 176 up. However, 75 contains a rarity: one read assigned to a species, *Alnus glutinosa*. PIA cannot assign to leaf taxa, so species are rare in this data. *A. glutinosa*, like the other two main European alders, prefers wet soils near watercourses or in marshes. The Betulaceae lower down the core may represent more.

Trees/shrubs and herbs

Samples 195 and 187 contain a strong signal from Rosoideae and Rosoideae *incertae sedis*, which are parent taxa of *Rubus*. *Rubus* was already the most frequent taxon in these samples. These reads suggest an even larger biomass.

Grasses and relatives

Grasses are present in all but the top two samples and dominant in 176 and 118. As is typically the case, most reads are assigned to either higher grass taxa or the Panicoideae group. As all samples but the top two had a significant reed signal, many of the higher-taxon reads could be interpreted as more reeds, although Panicoideae may be distinct.

Metazoa

Metazoa are absent from three of the lower samples, rare elsewhere, and mostly uninformative. Saltwater taxa are only in the topmost samples, although this distribution may be an artefact of low read counts, unlike in Embryophyta. Overall, there is no clear ecological signal.

4.9.3 Pianka scores

The Embryophyta Pianka scores correspond well with the taxonomic profiles. The first four samples, which returned mostly woody taxa, have high scores; the next four swap between grass and saltwater-dominated, and the top sample is very similar to the one below. However, there is too little data for such patterns to be visible in the Metazoa data.

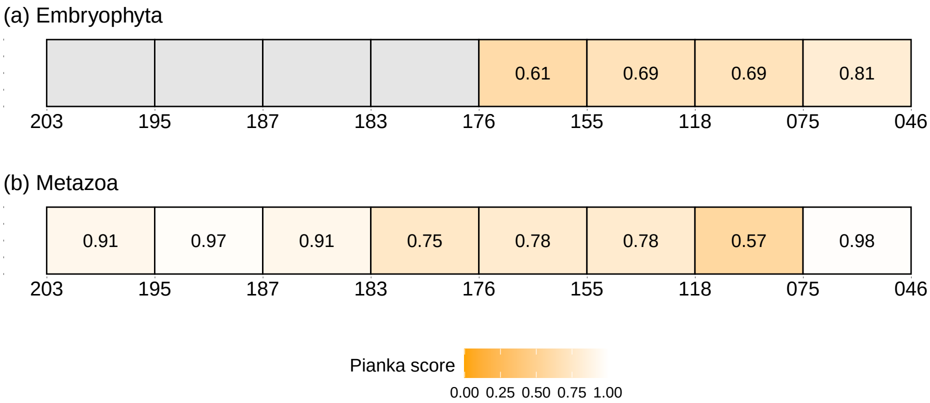


Figure 4.38: Pianka similarity scores between adjacent samples in ELF033 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

4.9.4 Summary

Samples in ELF033 generally have a low data yield, yet may capture the process of inundation. The deepest samples show a relatively stable assemblage of mostly woody taxa, including riparian, with some reeds and other grasses. Woody taxa diminish from sample 176 upward. 176 and 118 are dominated by grasses that may represent reed beds, and samples in between by *Zostera*. ELF033 could show a riparian environment becoming saltmarsh or another coastal wetland as the water table rises.

4.10 ELF033A

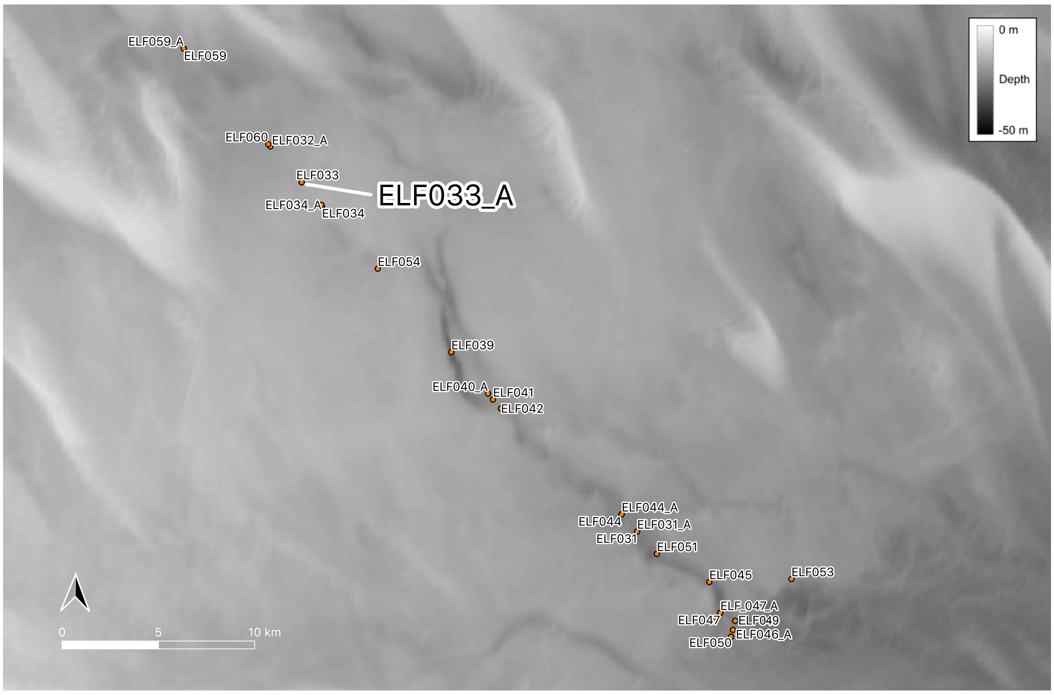


Figure 4.39: Bathymetry map of the palaeochannel transect highlighting ELF033A. Darker grey indicates greater depth.

ELF033A is the sister core of ELF033.

4.10.1 Read counts

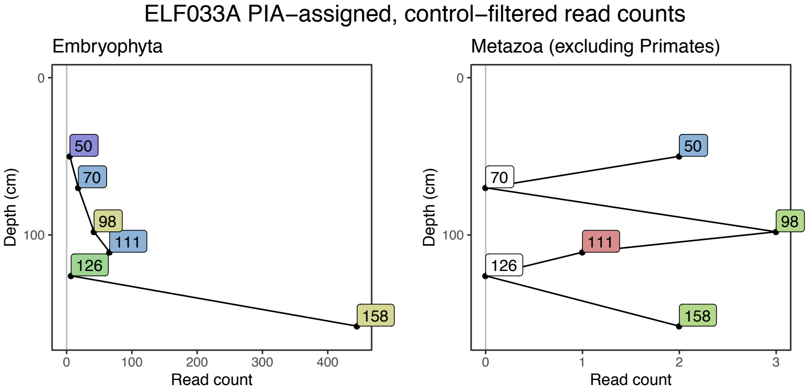
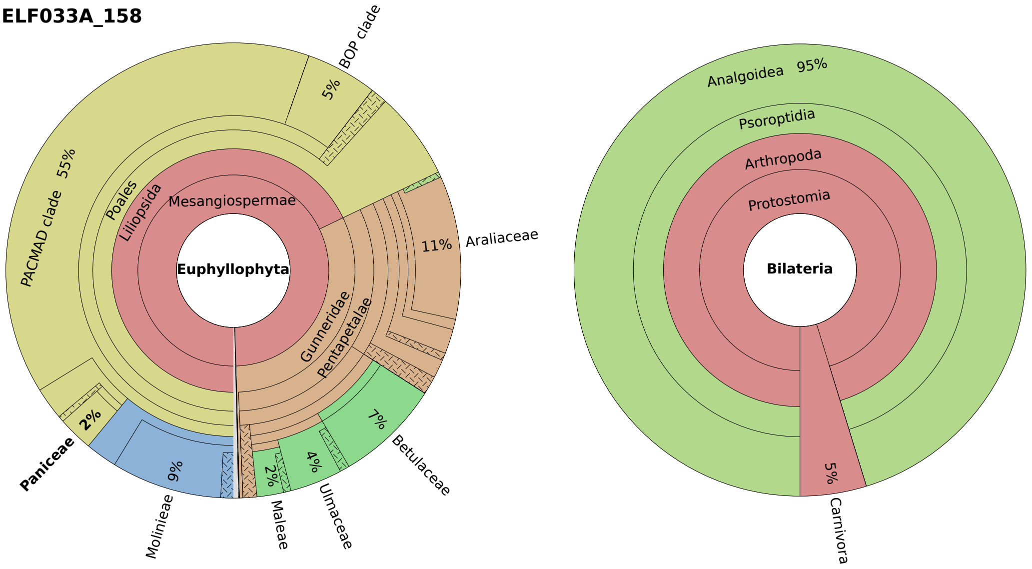


Figure 4.40: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF033A. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

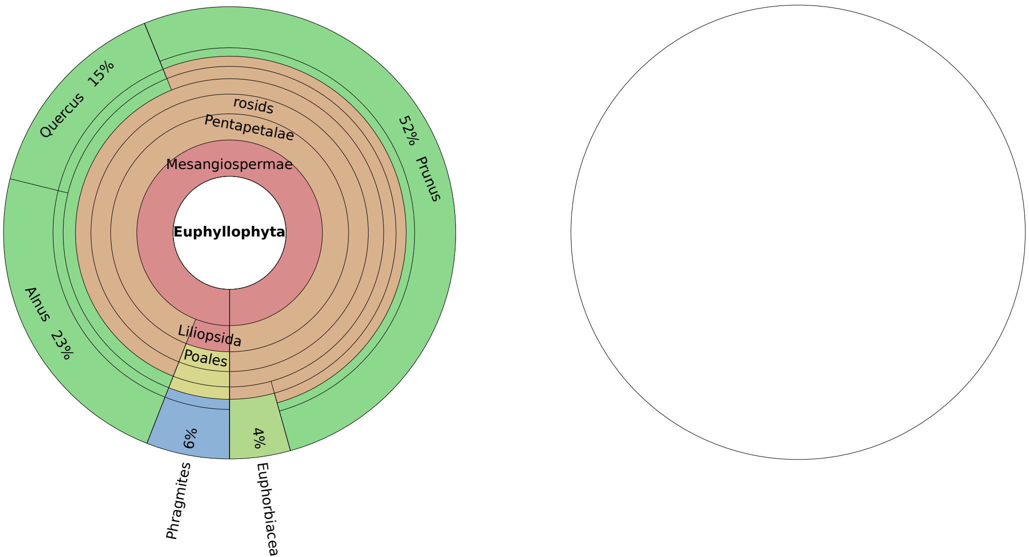
Like ELF033, ELF033A is a short core with regular samples, most of which have low read counts. There is one outlier for Embryophyta, sample 158, but with less than 500 reads it is still relatively poor.

4.10.2 Taxonomic profiles

ELF033A_158



ELF033A_126



ELF033A_111

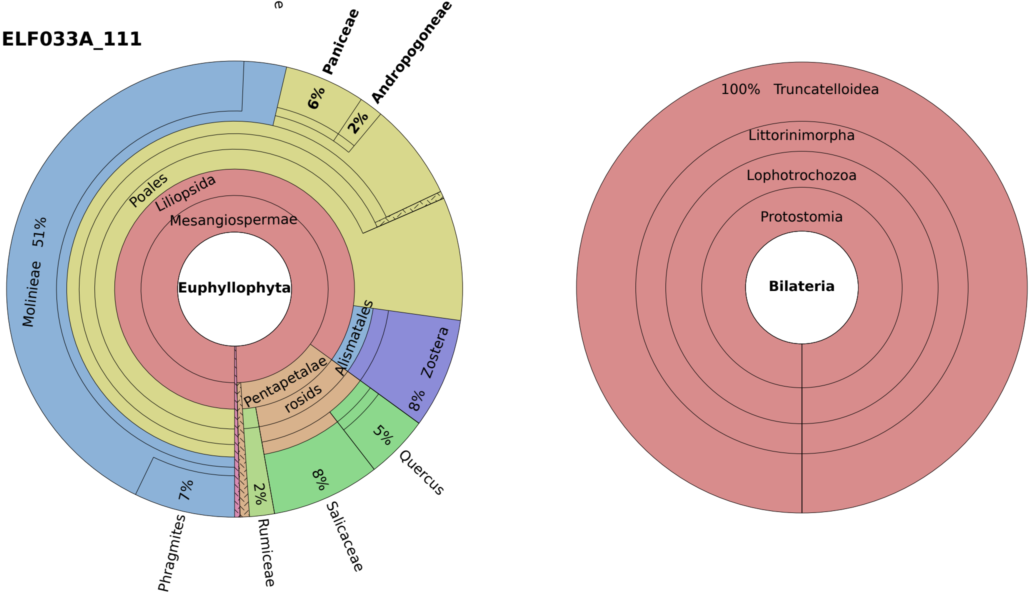


Figure 4.41: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF033A, samples 158, 126, and 111. Continued in figure 4.42. See figure 4.4 for colour key.

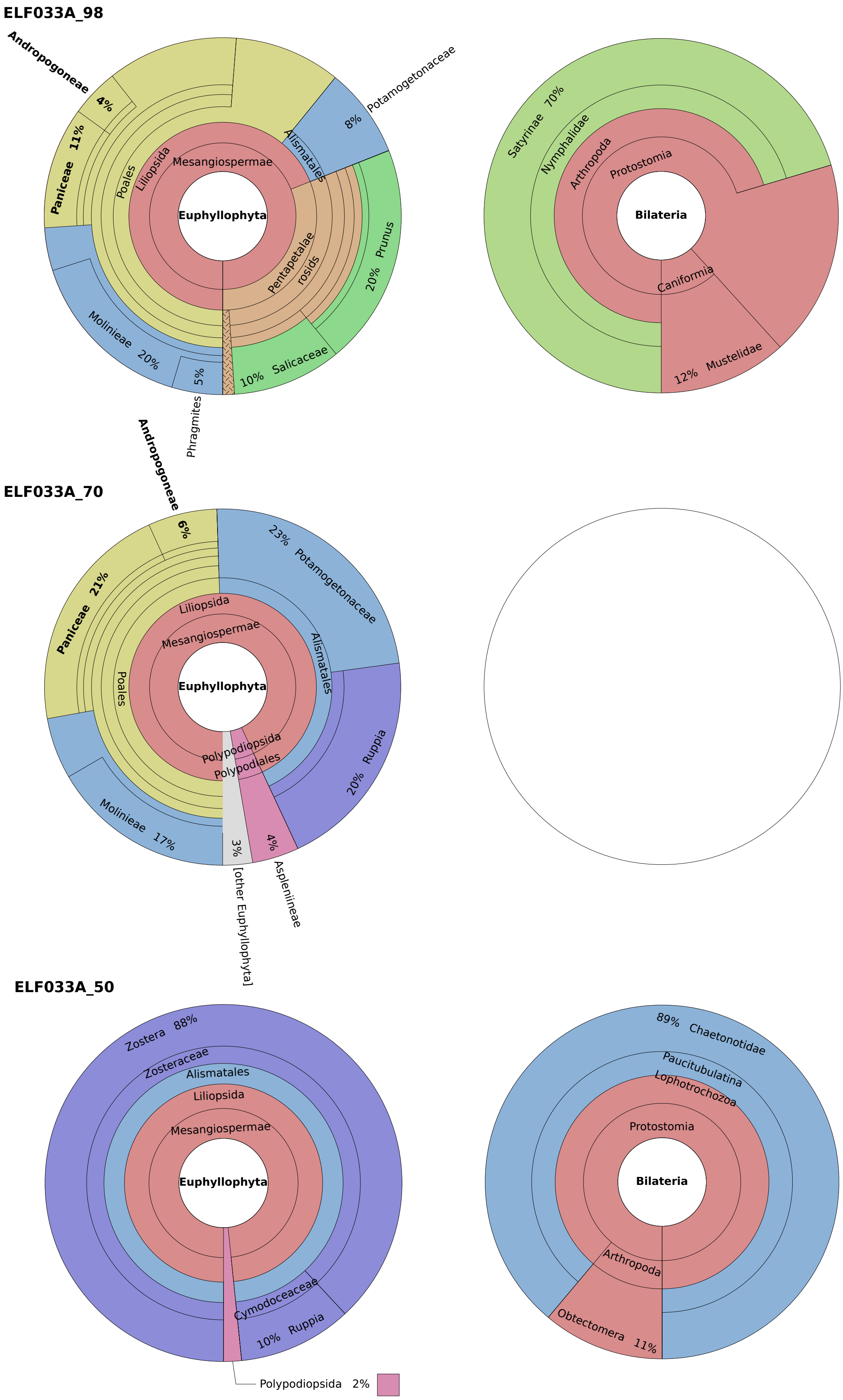


Figure 4.42: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF033A, samples 98, 70, and 50. See figure 4.4 for colour key.

Table 4.19: Biogenomic masses of European taxa in Embryophyta from ELF033A. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	158	126	111	98	70	50
		Ingroup	278.37	199.79	8.24	31.65	25.76	7.56	5.36
Salt/brackish aquatics	Tasselweeds	Ruppia	1.72	0.00	0.00	0.00	0.00	1.29	0.43
Salt/brackish aquatics	Eelgrasses	Zostera	5.50	0.00	0.00	1.83	0.00	0.00	3.67
Mixed aquatics	Pondweed family	Potamogetonaceae	2.98	0.00	0.00	0.00	1.49	1.49	0.00
Mixed aquatics	Reeds	Arundinoideae	6.01	4.24	0.00	0.71	0.71	0.35	0.00
Mixed aquatics	Common reed tribe	Molinieae	28.91	14.81	0.00	10.22	2.82	1.06	0.00
Mixed aquatics	Common reeds	Phragmites	4.62	1.68	0.42	1.68	0.84	0.00	0.00
Trees/shrubs		Viburnum	0.26	0.26	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Birch family	Betulaceae	14.04	14.04	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Alders	Alnus	1.61	0.00	1.61	0.00	0.00	0.00	0.00
Trees/shrubs	Oaks	Quercus	2.12	0.00	1.06	1.06	0.00	0.00	0.00
Trees/shrubs	Willow family	Salicaceae	3.61	0.00	0.00	1.81	1.81	0.00	0.00
Trees/shrubs	Stone fruit trees	Prunus	7.30	0.00	3.65	0.00	3.65	0.00	0.00
Trees/shrubs	Apple tribe	Maleae	3.57	3.57	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Apples	Malus	0.99	0.99	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Elm family	Ulmaceae	7.03	7.03	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Elms	Ulmus	1.29	1.29	0.00	0.00	0.00	0.00	0.00
Herbs	Ivies	Hedera	0.75	0.75	0.00	0.00	0.00	0.00	0.00
Herbs	Bellflower family	Campanulaceae	0.47	0.47	0.00	0.00	0.00	0.00	0.00
Herbs	Rhubarb family	Rumiceae	0.41	0.00	0.00	0.41	0.00	0.00	0.00
Herbs	Spurge family	Euphorbiaceae	0.31	0.00	0.31	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Gunneridae	0.47	0.47	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		campanulids	2.46	2.46	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Carrot order	Apiales	3.27	3.27	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Apiineae	1.40	1.40	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Ivy family	Araliaceae	19.18	19.18	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	0.27	0.27	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		lamiids	0.68	0.68	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Gentianales	0.90	0.90	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	0.88	0.88	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	1.31	1.31	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	0.56	0.56	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Saxifragales	0.19	0.00	0.00	0.00	0.19	0.00	0.00
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.16	0.00	0.00	0.16	0.00	0.00	0.00
Grasses and relatives	Grass order	Poales	15.52	11.71	0.00	2.05	1.76	0.00	0.00
Grasses and relatives	True sedges	Carex	2.28	2.28	0.00	0.00	0.00	0.00	0.00
Grasses and relatives		BOP clade	9.55	9.55	0.00	0.00	0.00	0.00	0.00
Grasses and relatives	Barley subtribe	Hordeinae	0.12	0.00	0.00	0.12	0.00	0.00	0.00
Grasses and relatives		PACMAD clade	73.92	73.92	0.00	0.00	0.00	0.00	0.00
Grasses and relatives		**Panicoideae**	**8.10**	**4.32**	**0.00**	**1.62**	**2.16**	**0.00**	**0.00**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**2.40**	**0.80**	**0.00**	**0.40**	**0.80**	**0.40**	**0.00**
Grasses and relatives		**Paniceae**	**9.38**	**4.69**	**0.00**	**1.34**	**2.01**	**1.34**	**0.00**
Ferns	Ferns	Polypodiopsida	0.14	0.07	0.00	0.00	0.00	0.00	0.07
Ferns	Horsetails	Equisetum	0.05	0.05	0.00	0.00	0.00	0.00	0.00
Ferns	Spleenwort suborder	Aspleniineae	0.36	0.00	0.00	0.09	0.00	0.27	0.00
Mixed		Euphyllophyta	0.86	0.69	0.00	0.00	0.00	0.17	0.00

Table 4.20: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF033A. Samples are in cm.

Ecological category	Group	Common name	Taxon	Total	158	126	111	98	70	50
			Ingroup	24.76	7.00	0.00	0.80	2.90	0.00	14.06
Mixed aquatics	Gastrotrichs		Chaetonotidae	12.50	0.00	0.00	0.00	0.00	0.00	12.50
Terrestrial	Mites		Analgoidea	6.67	6.67	0.00	0.00	0.00	0.00	0.00
Terrestrial	Butterflies	Browns	Satyrinae	2.04	0.00	0.00	0.00	2.04	0.00	0.00
Mixed	Mammals	Carnivorans	Carnivora	0.33	0.33	0.00	0.00	0.00	0.00	0.00
Mixed	Mammals	Weasel family	Mustelidae	0.34	0.00	0.00	0.00	0.34	0.00	0.00
Mixed	Invertebrates		Protostomia	0.52	0.00	0.00	0.00	0.52	0.00	0.00
Mixed	Lepidopterans		Obtectomera	1.56	0.00	0.00	0.00	0.00	0.00	1.56
Mixed	Gastropods		Truncatelloidea	0.80	0.00	0.00	0.80	0.00	0.00	0.00

Embryophyta

Aquatics

There are no freshwater aquatics, but saltwater appear from sample 111. Mixed aquatics, mostly in the reed group, are present in all but the top sample. This suggests a reed-to-saltwater transition as in ELF033, but with fewer reads and samples, the pattern is less clear.

Terrestrial

There is a clear woody signal in all but the top two samples, and trees are the most frequent ecological category in sample 126, but no individual taxa are particularly frequent or consistent. Supporting a general woodland signal is *Hedera* in the deepest sample, 158. *Hedera* is rarely seen in these cores but was also present in ELF033. Non-reed grasses have a clear signal in all samples but 126 and 55, which returned very few reads. Taxa are mostly higher or in the Panicoideae group.

Metazoa

Metazoa reads are very limited. The only confirmed aquatic taxon is found in the top sample and the only terrestrial below, but read counts are too low to infer genuine absence. There is no clear pattern.

4.10.3 Pianka scores

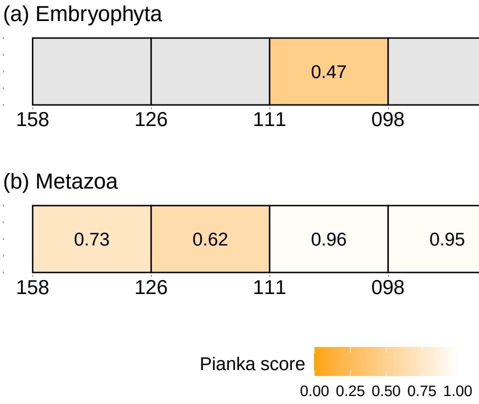


Figure 4.43: Pianka similarity scores between adjacent samples in ELF033A for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

The Embryophyta Pianka scores highlight the differences between samples 126 and 50 and their neighbours. This can probably be attributed to lack of data. The similarity between samples 111, 98, and 70, however, appears genuine. Their taxonomic profiles and read counts are indeed similar. There is too little data for informative Pianka comparisons of the Metazoa data.

4.10.4 Summary

ELF033A shares similarities with ELF033, but the even smaller amount of data makes inferring patterns more difficult. There may be a corresponding trend in Embryophyta from terrestrial to marine. ELF033A has more grass or reed than trees from the beginning, whereas ELF033 appears to move from trees to grass/reeds then marine, but ELF033A is a shorter core, so its deeper samples may correspond to the grass/reed strata in ELF033. The Metazoa data is too sparse to clarify any environmental changes.

4.11 ELF034A

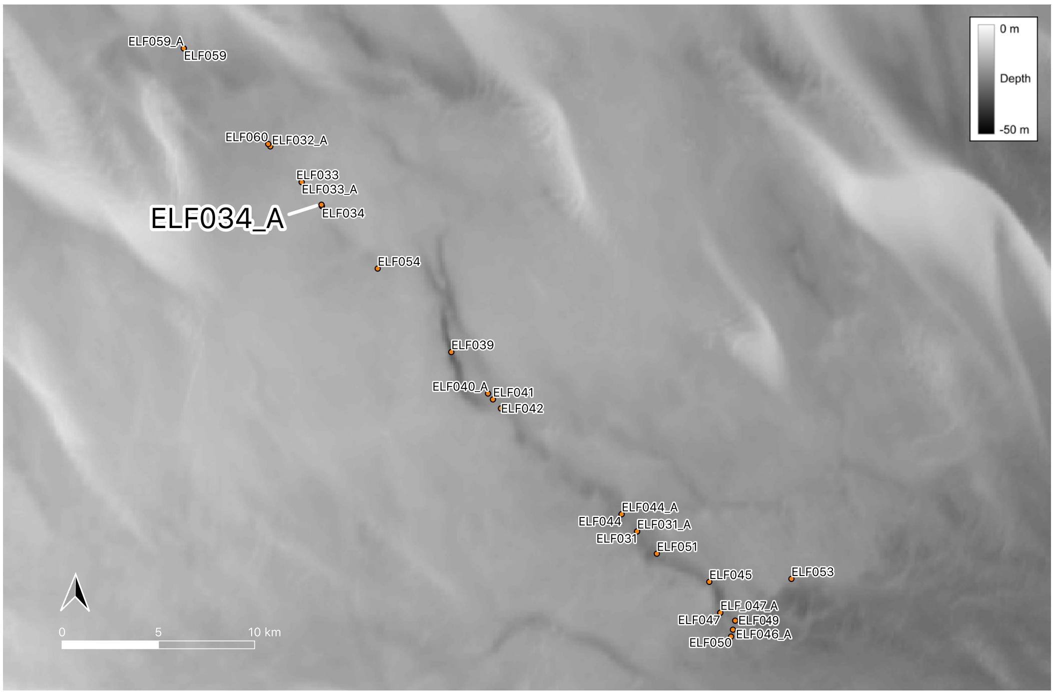


Figure 4.44: Bathymetry map of the palaeochannel transect highlighting ELF034A. Darker grey indicates greater depth.

The 'A' core in a pair of sister cores is typically discussed second, but ELF034A is slightly northwest of ELF034 (figure 4.51), approximately 2 km along the transect from ELF033 and ELF033A.

4.11.1 Read counts

ELF034A is one of the most heavily sampled cores. However, the deepest eight samples have very little data: the maximum Embryophyta count is 12. Fortunately, the top three samples all have much higher read counts, peaking at over 1,800 in 63 cm.

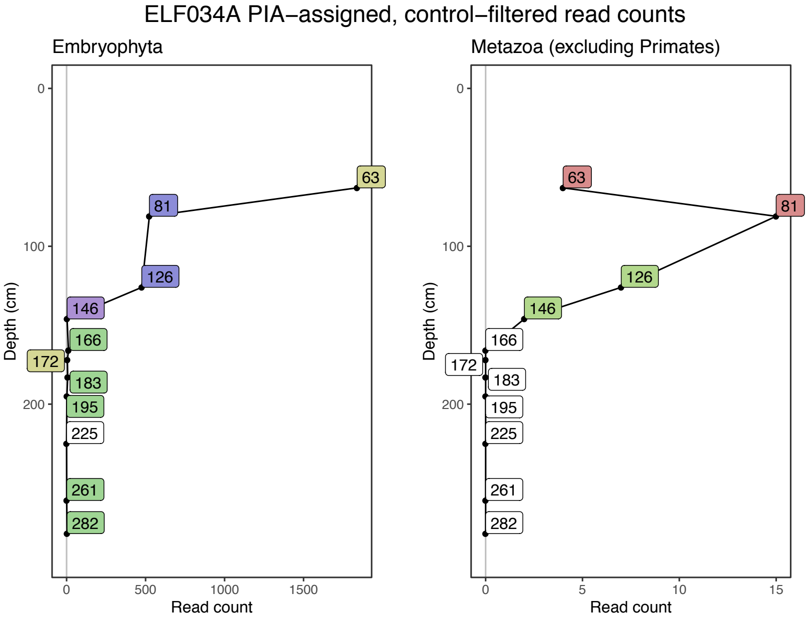
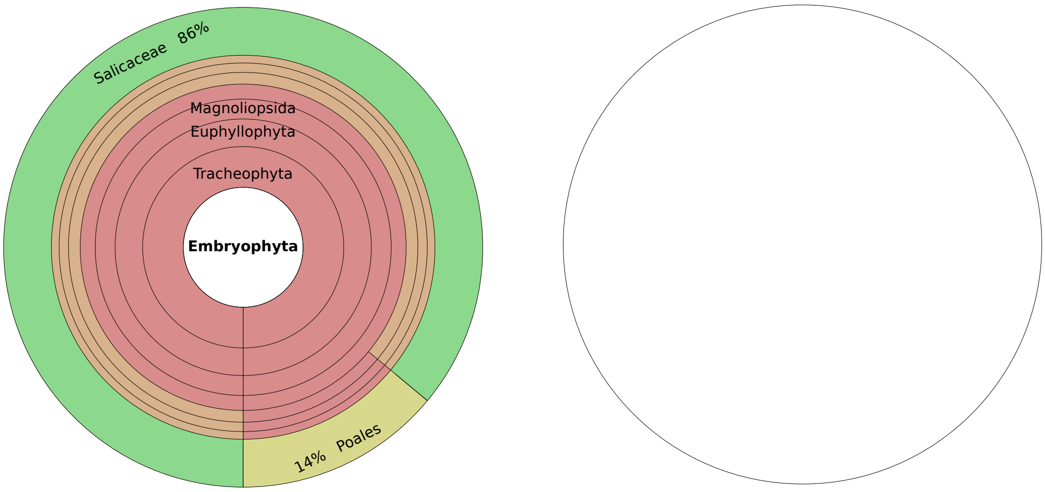


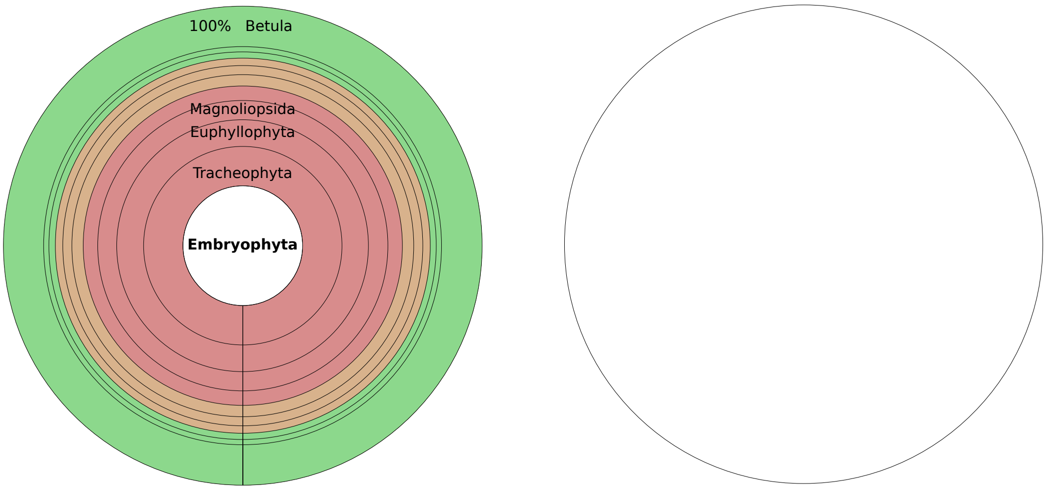
Figure 4.45: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF034A. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.11.2 Taxonomic profiles

ELF034A_282



ELF034A_261



ELF034A_225

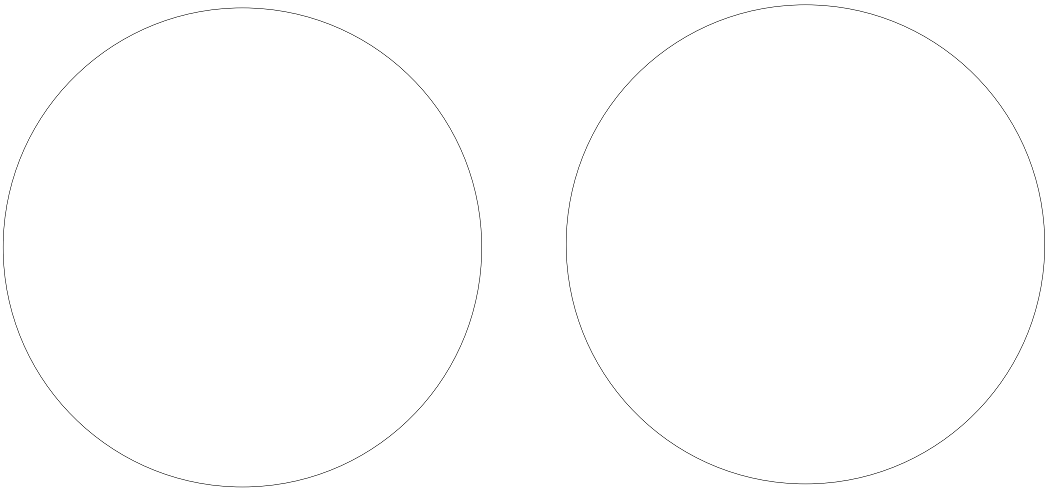
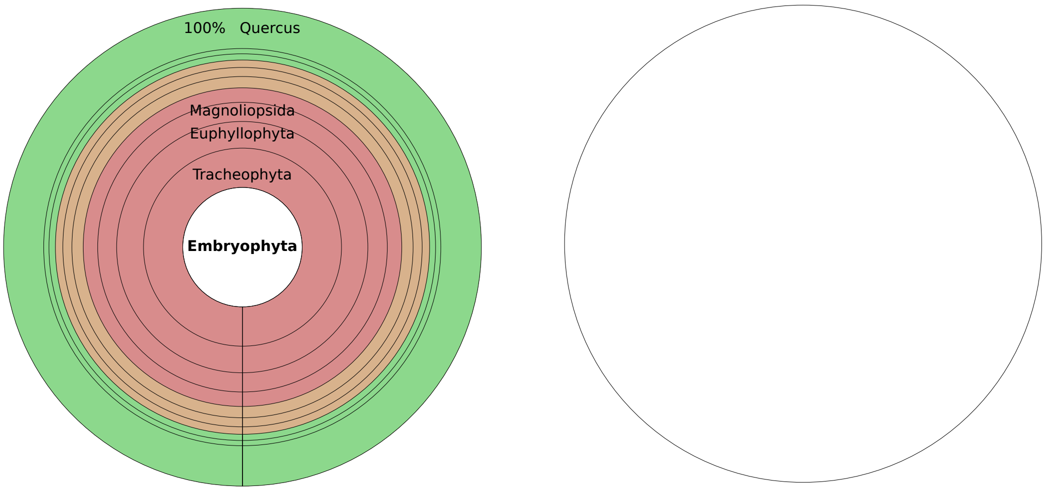
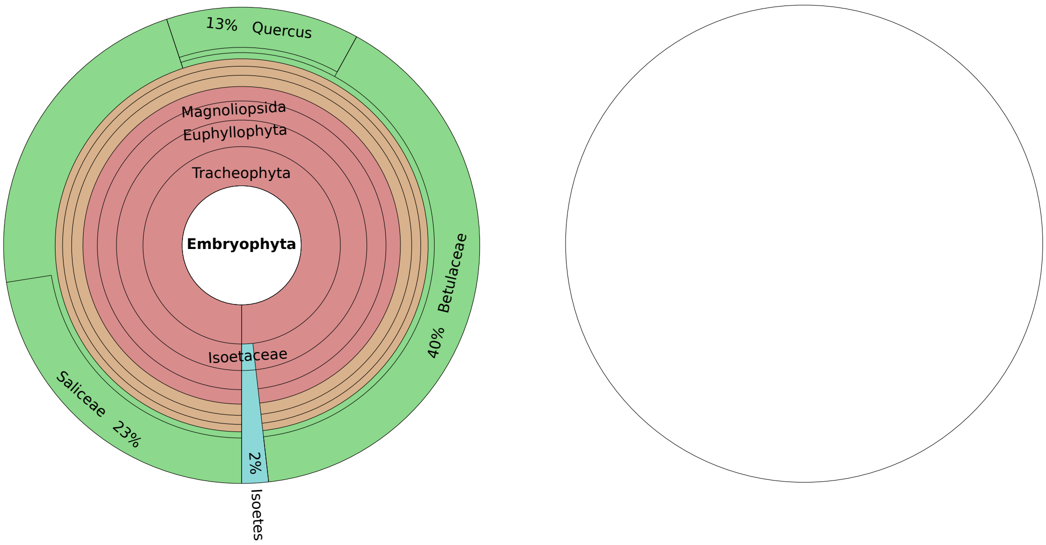


Figure 4.46: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF034A, samples 282, 261, and 225. Continued in figure 4.47. See figure 4.4 for colour key.

ELF034A_195



ELF034A_183



ELF034A_172

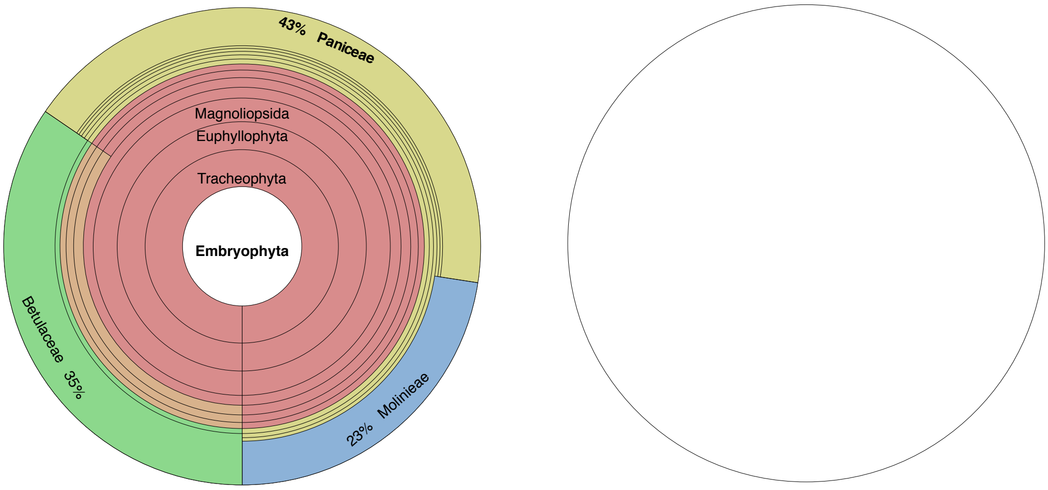
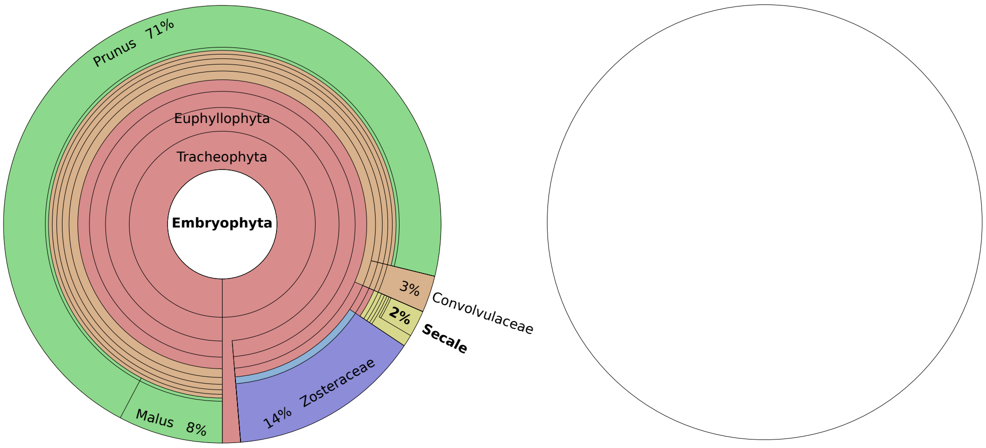
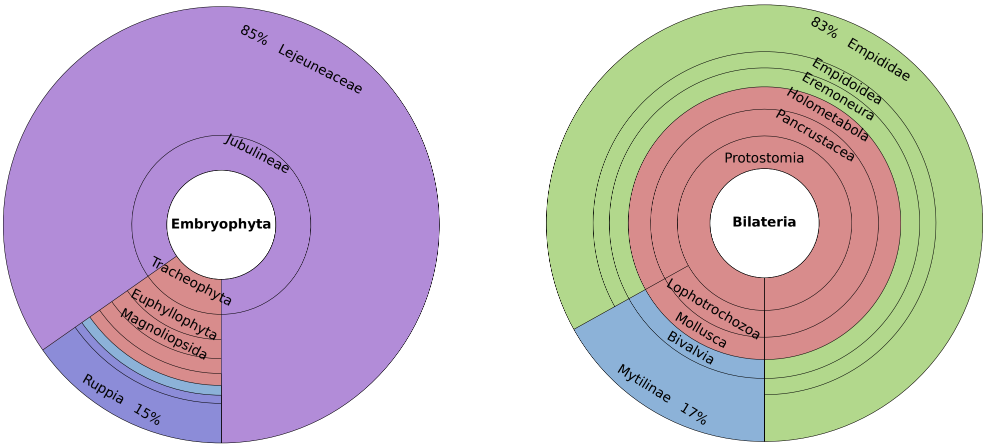


Figure 4.47: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF034A, samples 195, 183, and 172. Continued in figure 4.48. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

ELF034A_166



ELF034A_146



ELF034A_126

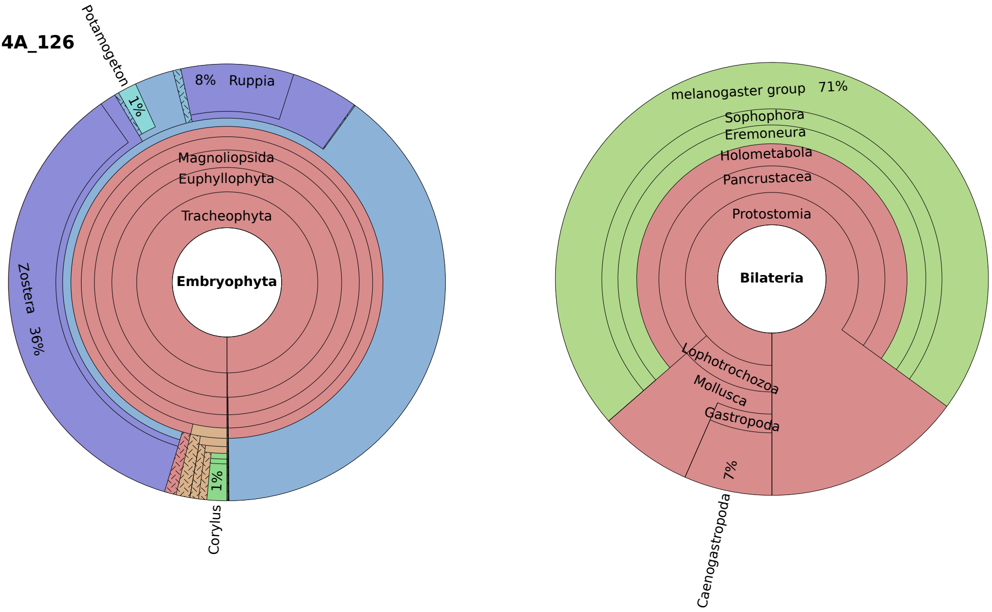


Figure 4.48: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF034A, samples 166, 146, and 126. Continued in figure 4.49. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

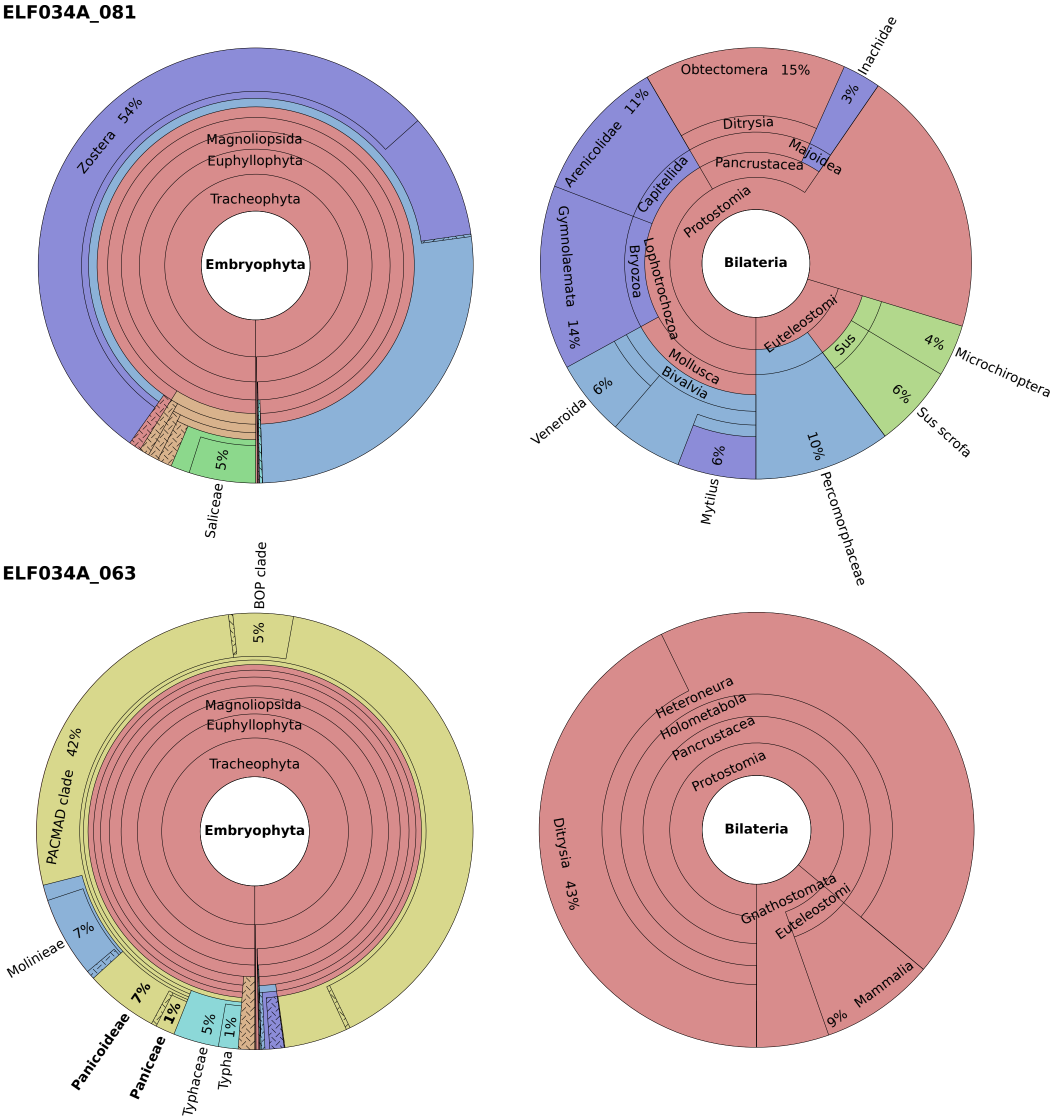


Figure 4.49: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF034A, samples 81 and 63. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

Table 4.21: Biogenomic masses of European taxa in Embryophyta from ELF034A. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	282	261	225	195	183	172	166	146	126	81	63
		Ingroup	1004.88	2.45	0.97	0.00	1.91	8.06	3.13	12.84	5.62	161.36	260.81	547.73
Freshwater aquatics	Water-plantain family	Alismataceae	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00
Freshwater aquatics	Arrowheads	Sagittaria	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00
Freshwater aquatics	Elodea family	Hydrocharitaceae	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.59	0.00	0.00
Freshwater aquatics	Pondweeds	Potamogeton	2.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.16	0.00	0.00
Freshwater aquatics	Rice tribe	Oryzeae	1.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.04
Freshwater aquatics	Bulrush family	Typhaceae	17.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	17.48
Freshwater aquatics	Bulrushes	Typha	7.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.55
Freshwater aquatics	Waterlily order	Nymphaeales	0.57	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.57	0.00
Freshwater aquatics	Waterlily family	Nymphaeaceae	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.53
Freshwater aquatics	Quillworts	Isoetes	0.44	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.15	0.00	0.15
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	7.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.67	0.00	0.00
Salt/brackish aquatics	Tasselweeds	Ruppia	13.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.86	12.88	0.00	0.00
Salt/brackish aquatics	**Neptune-grasses**	**Posidonia**	**0.32**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.32**	**0.00**	**0.00**
Salt/brackish aquatics	Eelgrass family	Zosteraceae	29.36	0.00	0.00	0.00	0.00	0.00	0.00	1.83	0.00	1.83	23.85	1.83
Salt/brackish aquatics	Eelgrasses	Zostera	198.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	55.05	137.61	5.50
Mixed aquatics	Alismatids	Alismatales	130.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	61.62	68.41	0.00
Mixed aquatics	Pondweed family	Potamogetonaceae	6.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.46	0.50	1.49
Mixed aquatics	Pondweeds	Stuckenia	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00
Mixed aquatics	Reeds	Arundinoideae	6.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	6.01
Mixed aquatics	Common reed tribe	Molinieae	32.43	0.00	0.00	0.00	0.00	0.00	0.71	0.00	0.00	0.35	0.35	31.02
Mixed aquatics	Common reeds	Phragmites	3.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.36
Xerophytes	Goosefoot subfamily	Chenopodioideae	0.82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.82	0.00
Trees/shrubs	Elder family	Adoxaceae	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13
Trees/shrubs		Cornales	0.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.61
Trees/shrubs	Birch family	Betulaceae	6.48	0.00	0.00	0.00	0.00	3.24	1.08	0.00	0.00	0.00	0.00	2.16
Trees/shrubs	Birches	Betula	0.97	0.00	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Hazels	Corylus	2.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.27	0.00	0.00
Trees/shrubs	Oaks	Quercus	2.12	0.00	0.00	0.00	1.06	1.06	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Willow family	Salicaceae	9.03	1.81	0.00	0.00	0.00	1.81	0.00	0.00	0.00	0.00	3.61	1.81
Trees/shrubs	Willow tribe	Saliceae	14.52	0.00	0.00	0.00	0.00	1.82	0.00	0.00	0.00	0.00	12.71	0.00
Trees/shrubs	Stone fruit trees	Prunus	10.96	0.00	0.00	0.00	0.00	0.00	0.00	9.13	0.00	0.00	1.83	0.00
Trees/shrubs	Apples	Malus	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.00
Trees/shrubs	Elm family	Ulmaceae	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Herbs	Plantains	Plantago	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.46	0.00	0.00
Herbs	Rhubarb family	Rumiceae	0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.41
Herbs	Clover tribe	Trifolieae	0.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.94	0.00
Herbs	Onion family	Allioideae	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00
Herbs	Asparagus family	Asparagaceae	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00
Herbs	Orchids	Orchidaceae	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00
Herbs		Epidendroideae	0.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.32	0.00	0.00
Trees/shrubs and herbs	Ivy family	Araliaceae	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00
Trees/shrubs and herbs	Ragwort tribe	Senecioneae	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.23	0.00	0.00
Trees/shrubs and herbs		Solanales	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.00
Trees/shrubs and herbs	Bindweed family	Convolvulaceae	1.05	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.00	0.70	0.00	0.00

Continued on next page

Table 4.21 continued

Ecological category	Common name	Taxon	Total	282	261	225	195	183	172	166	146	126	81	63
		Ingroup	1004.88	2.45	0.97	0.00	1.91	8.06	3.13	12.84	5.62	161.36	260.81	547.73
Trees/shrubs and herbs		core genistoids	0.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.94	0.00	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	1.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.96	0.00
Trees/shrubs and herbs		Myrtales	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00
Trees/shrubs and herbs		Saxifragales	0.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.19	0.19
Trees/shrubs and herbs	**Palms**	**Arecaceae**	**0.29**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.29**
Grasses and relatives	Grass order	Poales	24.59	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	24.30
Grasses and relatives	Sedges	Cyperaceae	3.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.54	1.54
Grasses and relatives	Grass family	Poaceae	207.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	207.49
Grasses and relatives		BOP clade	23.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	23.13
Grasses and relatives		Stipeae	0.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.63
Grasses and relatives		Triticodae	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00
Grasses and relatives	**Ryes**	**Secale**	**0.25**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.25**	**0.00**	**0.00**	**0.00**	**0.00**
Grasses and relatives		PACMAD clade	140.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	140.39
Grasses and relatives		**Panicoideae**	**28.07**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**28.07**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**2.40**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.40**	**0.00**	**2.00**
Grasses and relatives		**Paniceae**	**8.71**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**1.34**	**0.00**	**0.00**	**0.00**	**0.00**	**7.37**
Ferns	Ferns	Polypodiopsida	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07
Ferns	Polypod ferns	Polypodiales	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.08
Ferns	Spleenwort suborder	Aspleniineae	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.18
Ferns	Bracken family	Dennstaedtiaceae	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00
Bryophytes		Lejeuneaceae	4.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.76	0.00	0.00	0.00
Mixed		Tracheophyta	0.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.35
Mixed		Euphyllophyta	1.38	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.17	1.03

Table 4.22: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF034A. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	282	261	225	195	183	172	166	146	126	81	63
			Ingroup	29.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.90	11.38	10.33	3.65
Salt/brackish aquatics	Crabs		Inachidae	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.00
Salt/brackish aquatics	Annelids	Lugworms	Arenicolidae	1.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.11	0.00
Salt/brackish aquatics	Bryozoans		Gymnolaemata	1.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.43	0.00
Salt/brackish aquatics	Bivalves		Mytilus	0.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.61	0.00
Mixed aquatics	Fishes		Percomorphaceae	1.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.06	0.00
Mixed aquatics	Bivalves		Veneroida	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.59	0.00
Mixed aquatics	Bivalves	Mussels	Mytilidae	0.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.00
Mixed aquatics	Bivalves		Mytilinae	0.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.66	0.00	0.00	0.00
Terrestrial	Mammals	Wild boar and domestic pig	Sus scrofa	0.64	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.64	0.00
Terrestrial	Mammals	Microbats	Microchiroptera	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.00
Terrestrial	Dipterans		melanogaster group	4.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.95	0.00	0.00
Terrestrial	Dipterans	Dagger flies	Empididae	3.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.24	0.00	0.00	0.00
Mixed	Vertebrates	Jawed vertebrates	Gnathostomata	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20
Mixed	Mammals	Mammals	Mammalia	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31
Mixed	Invertebrates		Protostomia	3.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.04	2.08	0.00
Mixed	Lepidopterans		Heteroneura	1.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.58
Mixed	Lepidopterans		Ditrysia	1.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.56
Mixed	Lepidopterans		Obtectomera	1.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.56	0.00
Mixed	Molluscs	Molluscs	Mollusca	0.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.49	0.00	0.00
Mixed	Gastropods		Caenogastropoda	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.46	0.00	0.00

Embryophyta

The most consistent ecological category throughout the core is trees/shrubs, particularly Betulaceae and the willow group, though none with high biogenomic mass. Trees are the most frequent category in samples up to 146. Saltwater aquatics (*Zostera* and *Ruppia*) are present in the top half of the core, becoming dominant in 126 and 81. They give way to a majority of grasses in 63, at least some of which could be interpreted as reeds.

However, one interesting grass is also present in a deeper sample, 166, as two of only twelve reads: *Secale* (ryes). *Secale* is not native to Great Britain; although there are wild species in Europe, the cultivated *S. cereale* was introduced by early farmers (Colledge *et al.* 2005). *Secale* is absent from negative controls and age-authentication of general cereal reads in Chapter 6 (Mesophilic taxa and human disturbance indicators) was cautiously optimistic, but there is insufficient evidence to confirm cultivated rye in ELF034A at present, and the lack of DNA from this sample suggests that deeper sequencing may not be helpful. Other environmental proxies in Europe's Lost Frontiers may be more informative.

Metazoa

Metazoa are absent from the seven deepest samples and rare in the remainder. Terrestrial taxa are most frequent in samples 146 and 126, whereas sample 81 adds several saltwater and mixed aquatics. Taxa in the top sample are uninformative.

4.11.3 Pianka scores

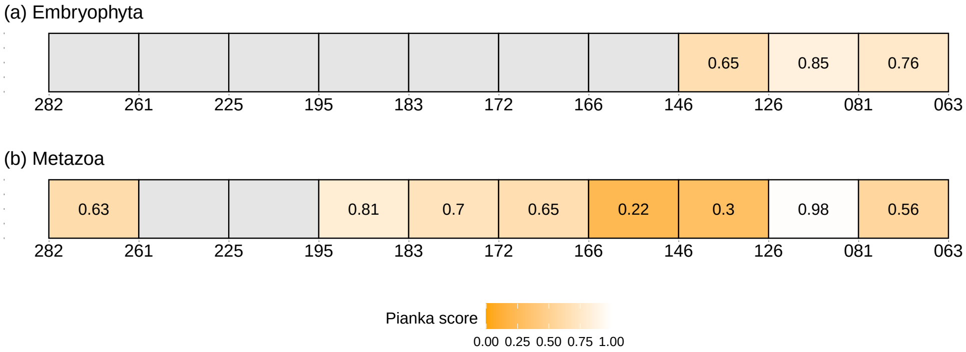


Figure 4.50: Pianka similarity scores between adjacent samples in ELF034A for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

Most Pianka comparisons involve too few reads to be informative. However, the Embryophyta scores between samples 126, 81, and 63 correspond well with the profiles. 126 and 81 are similarly dominated by *Zostera*, whereas 63 has mostly grasses.

4.11.4 Summary

It is difficult to separate differences along ELF034A that were due to lack of data or environmental change. However, there appears to be a terrestrial to saltwater trend, corresponding with an increase in read count, possibly due to the effects of salt on DNA preservation. What data is in the deeper samples is almost all from woody taxa, whereas the top samples appear estuarine with reed beds. Metazoa data appears in sample 146 with mostly terrestrial taxa and aquatics are added in sample 81. The low read counts prevent firm conclusions, but the data is consistent with inundation.

4.12 ELF034

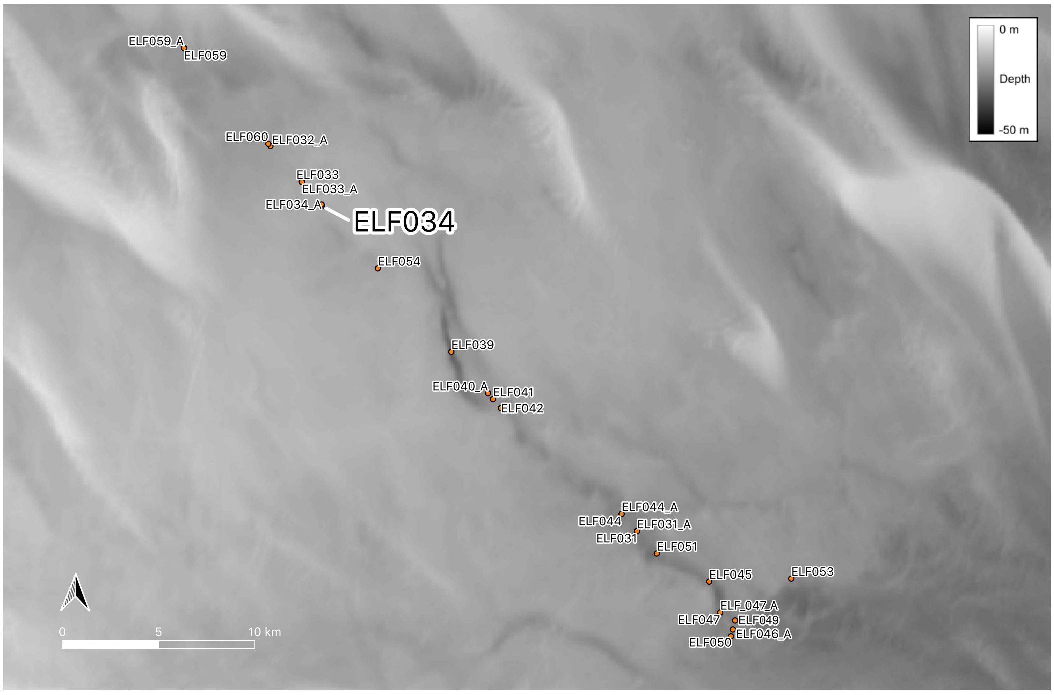


Figure 4.51: Bathymetry map of the palaeochannel transect highlighting ELF034. Darker grey indicates greater depth.

ELF034 is slightly further along the transect than ELF034A.

4.12.1 Read counts

ELF034 is another core with very few reads. There may be a trend for increasing read count with depth, but the maximum is only 27 for Embryophyta. Note that the single Metazoa reads returned by samples 157 and 94 were not assigned to European taxa, so those samples are treated as empty.

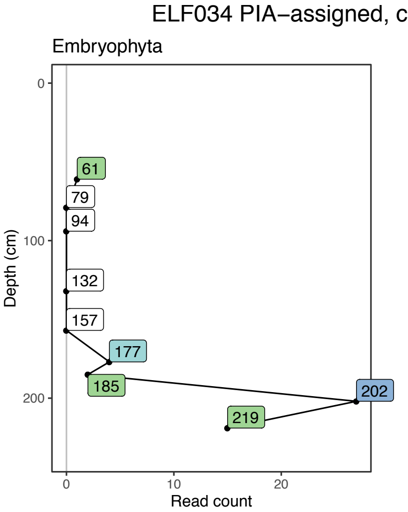
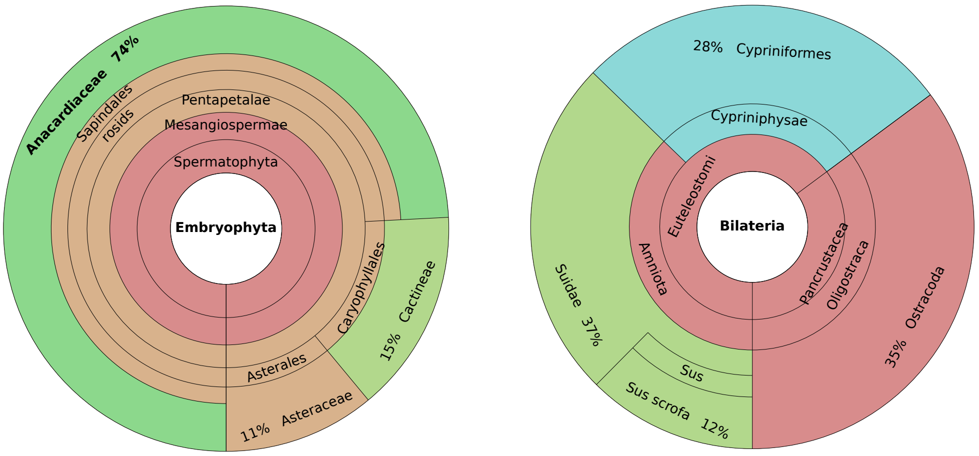


Figure 4.52: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF034. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.12.2 Taxonomic profiles

ELF034_219



ELF034_202

ELF034_185

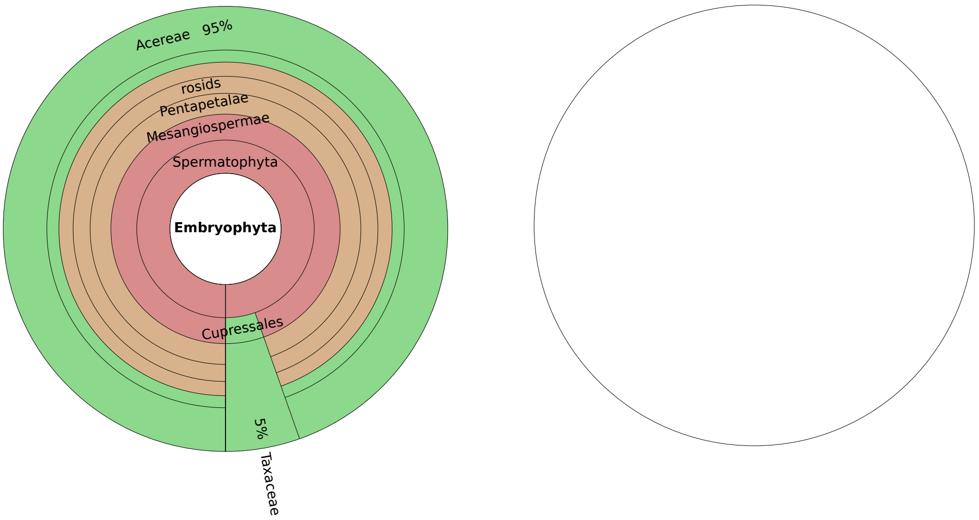


Figure 4.53: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF034, samples 219, 202, and 185. Continued in figure 4.54. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

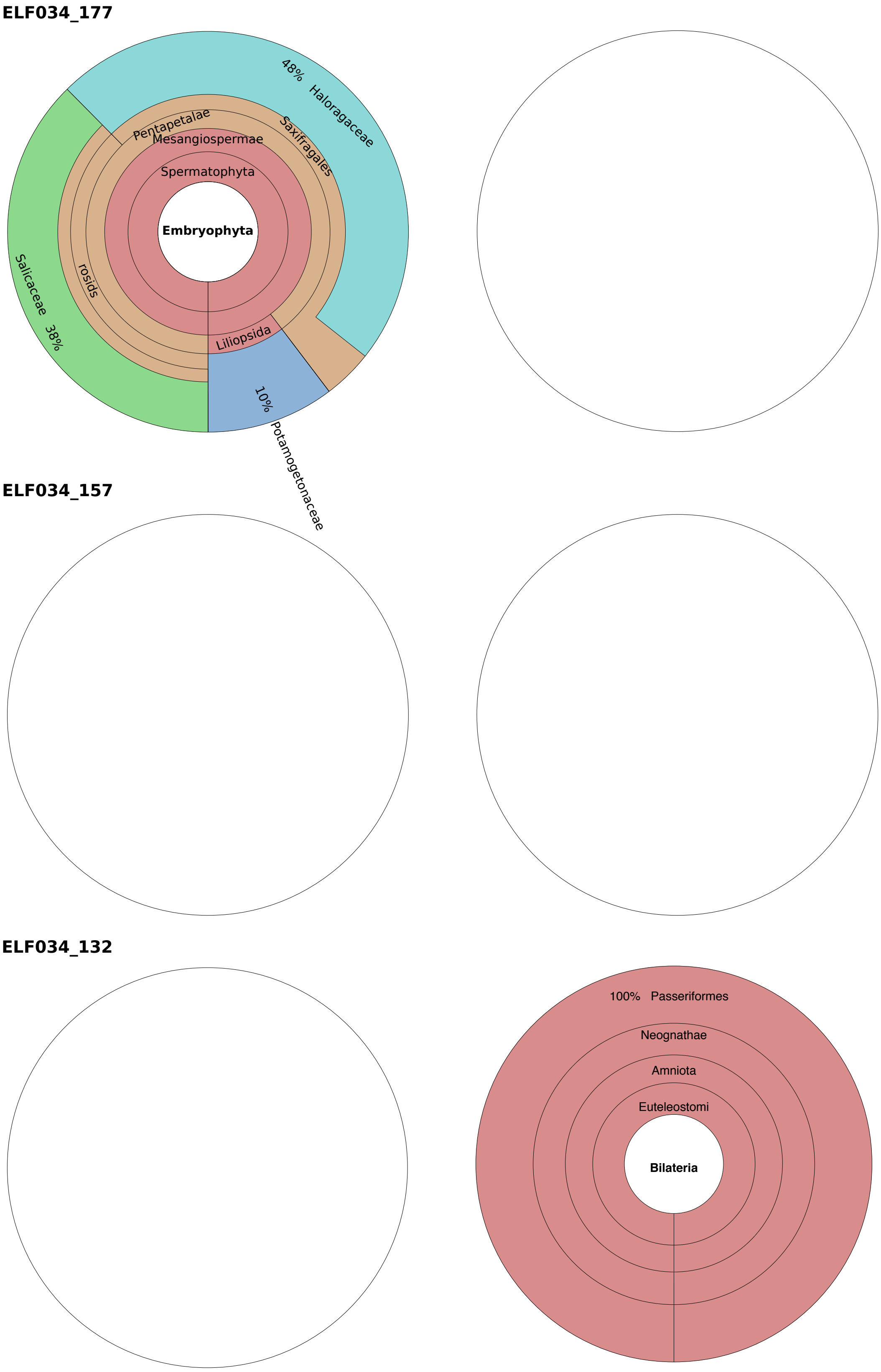
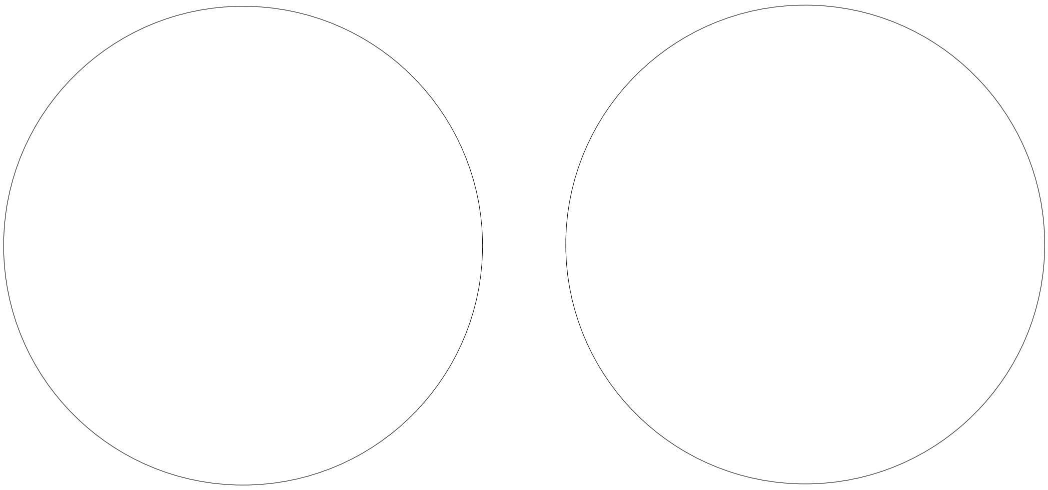
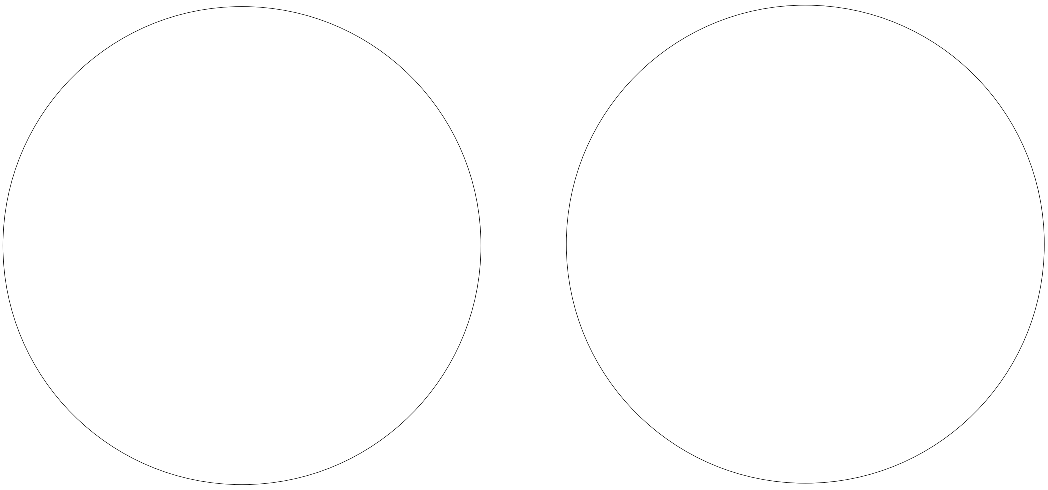


Figure 4.54: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF034, samples 177, 157, and 132. Continued in figure 4.55. See figure 4.4 for colour key.

ELF034_094



ELF034_079



ELF034_061

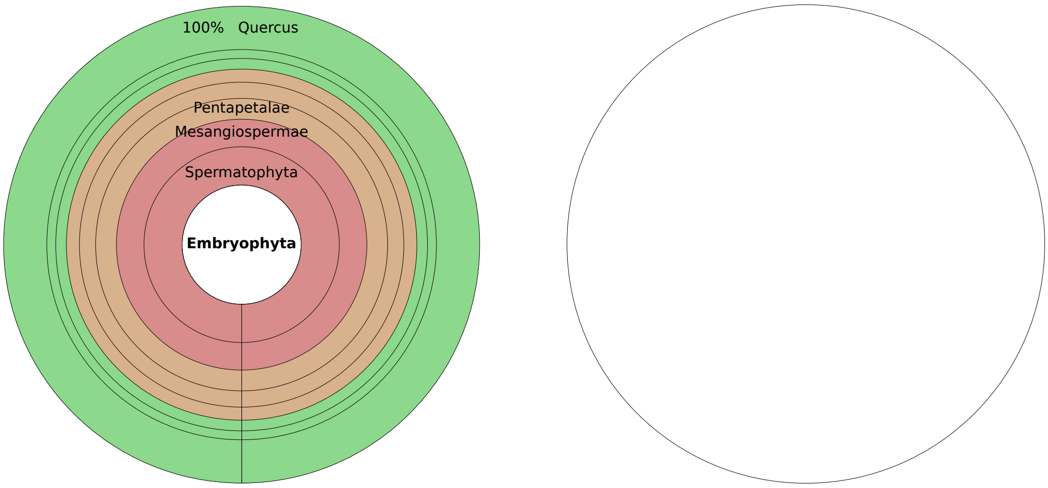


Figure 4.55: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF034, samples 94, 79, and 61. See figure 4.4 for colour key.

Table 4.23: Biogenomic masses of European taxa in Embryophyta from ELF034. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	219	202	185	177	157	132	94	79	61
		Ingroup	36.38	10.66	18.72	1.14	4.80	0.00	0.00	0.00	0.00	1.06
Freshwater aquatics	Water-milfoil family	Haloragaceae	4.62	0.00	2.31	0.00	2.31	0.00	0.00	0.00	0.00	0.00
Freshwater aquatics	Pondweeds	Potamogeton	0.54	0.00	0.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Pondweed family	Potamogetonaceae	7.93	0.00	7.44	0.00	0.50	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Yew family	Taxaceae	0.06	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Oaks	Quercus	1.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.06
Trees/shrubs	Willow family	Salicaceae	3.61	0.00	1.81	0.00	1.81	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Willow tribe	Saliceae	1.82	0.00	1.82	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Willows	Salix	3.46	0.00	3.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	**Cashew family**	**Anacardiaceae**	**1.83**	**1.83**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Trees/shrubs	Maple tribe	Acereae	1.08	0.00	0.00	1.08	0.00	0.00	0.00	0.00	0.00	0.00
Herbs	Cactus suborder	Cactineae	0.36	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	0.54	0.27	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Saxifragales	0.19	0.00	0.00	0.00	0.19	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.16	0.00	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grasses and relatives		Triticodae	0.22	0.00	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bryophytes	Liverworts	Marchantiophyta	0.53	0.00	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed		Tracheophyta	0.17	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.24: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF034. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	219	202	185	177	157	132	94	79	61
			Ingroup	7.64	2.58	3.19	0.00	0.00	0.37	0.76	0.75	0.00	0.00
Freshwater aquatics	Fishes	Carp order	Cypriniformes	0.71	0.71	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Mammals	Pig family	Suidae	0.64	0.64	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Mammals	Wild boar and domestic pig	Sus scrofa	0.32	0.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Birds	Landfowl	Galliformes	0.77	0.00	0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Birds	Perching birds	Passeriformes	0.76	0.00	0.00	0.00	0.00	0.00	0.76	0.00	0.00	0.00
Mixed	Ostracods	Ostracods	Ostracoda	0.91	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Hymenopterans	Wasps, ants and bees	Apocrita	2.42	0.00	2.42	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Embryophyta

The few Embryophyta reads in this core are in the four deepest samples with the exception of one *Quercus* in the very top. There is no clear pattern of any ecological category across the core, but trees and aquatics are most frequent. There are three rarely-seen woody taxa: Taxaceae (the yew family), Acereae (the maple tribe), and Anacardiaceae (the cashew family). There are two European genera in Anacardiaceae, *Pistacia* (pistacio genus) and *Rhus* (sumacs), both of which are native to southern regions and may indicate surprisingly warm and dry conditions, although *Rhus* has naturalised in Britain. However, both species are widely cultivated, suggesting possible contamination or mis-assignment, and there was not enough data for age-authentication. There is not enough evidence to accept Anacardiaceae as genuine.

Metazoa

There is very little Metazoa data, and the three terrestrial taxa are relatives of domesticates to which reads may have been over-assigned. Besides one freshwater read in the deepest sample, the animals are not very informative.

4.12.3 Pianka scores

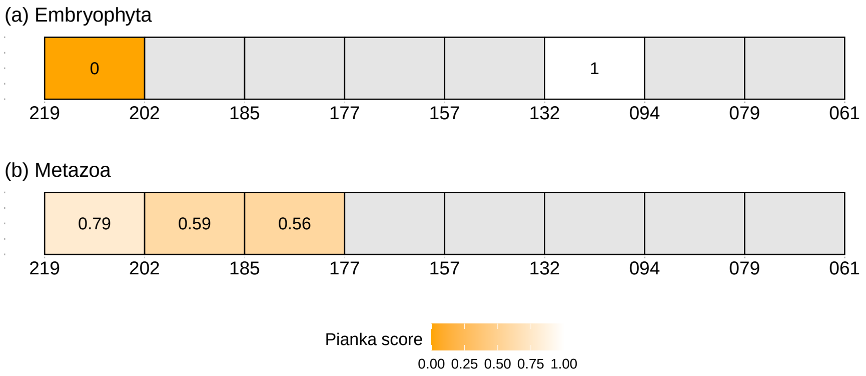


Figure 4.56: Pianka similarity scores between adjacent samples in ELF034 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

Pianka scores are limited with such sparse data, but it is clear that the three deepest samples share some overlap in Embryophyta.

4.12.4 Summary

The few informative reads indicate a woody environment at least towards the base of the core, with a small freshwater or ambiguous aquatic influence. Previous cores suggested a positive correlation between data yield and saltwater taxa, explained by the beneficial effect of salt on DNA preservation. ELF034 is a core with no saltwater taxa and low data yield. Absence is difficult to prove, but the prevalence of *Zostera* in other cores suggests that it may be genuine.

4.13 ELF054

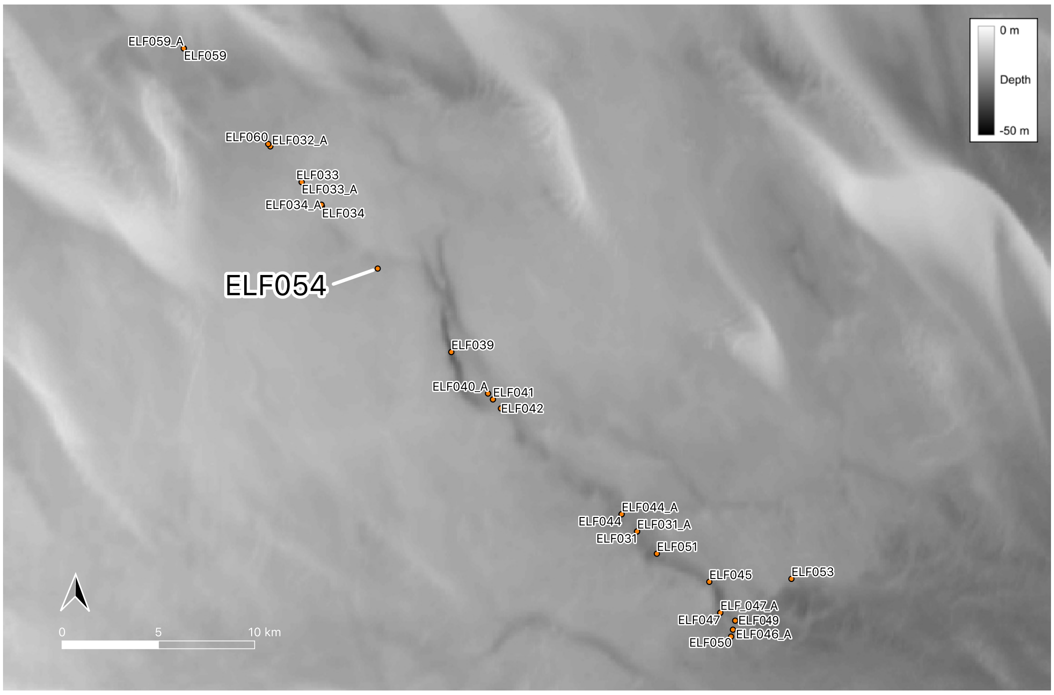


Figure 4.57: Bathymetry map of the palaeochannel transect highlighting ELF054. Darker grey indicates greater depth.

ELF054 is approximately 4 km southeast of ELF034 and the final core in the transect before the visible palaeochannel.

4.13.1 Read counts

Similar to ELF034A, read count in ELF054 broadly decreases with depth. The deepest samples have no more than 31 Embryophyta reads. Sample 182 has a reasonable 460, 140 dips to 81, and sample 58 has over 2,200. The Metazoa counts follow a similar trend on a much smaller scale, with the addition of a second peak at sample 291.

Note that the four deepest samples, 356-291 cm, only contain data from the initial sequencing run. This is because the negative control associated with the replicates of these samples was excessively large, so the batch of samples was excluded. In the top four samples, the initial and replicate runs contributed similar numbers of reads. If this is true for the deepest samples, the replicates are unlikely to have added a significant amount of data.

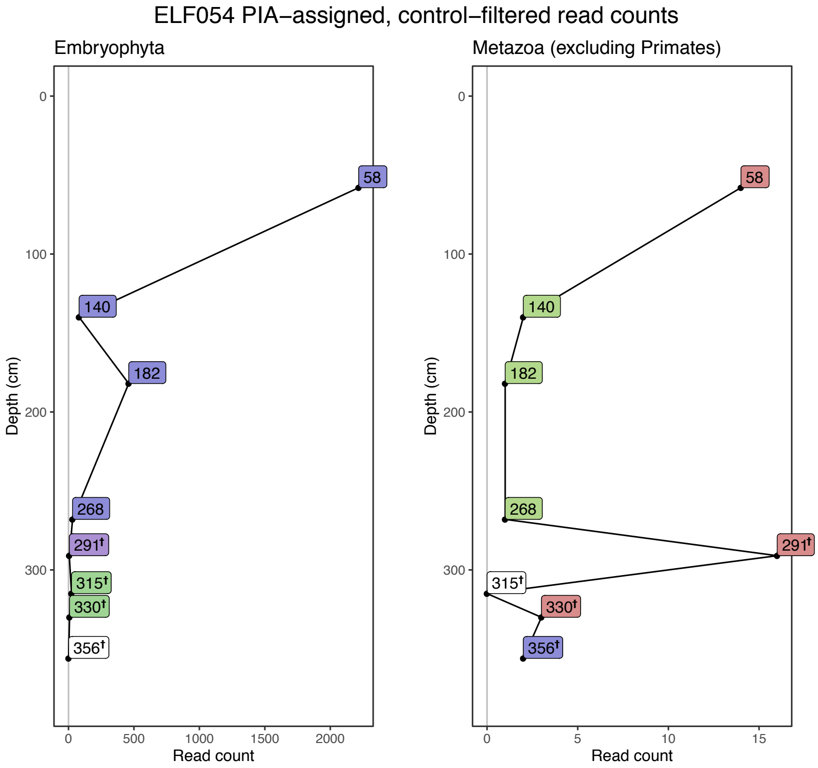
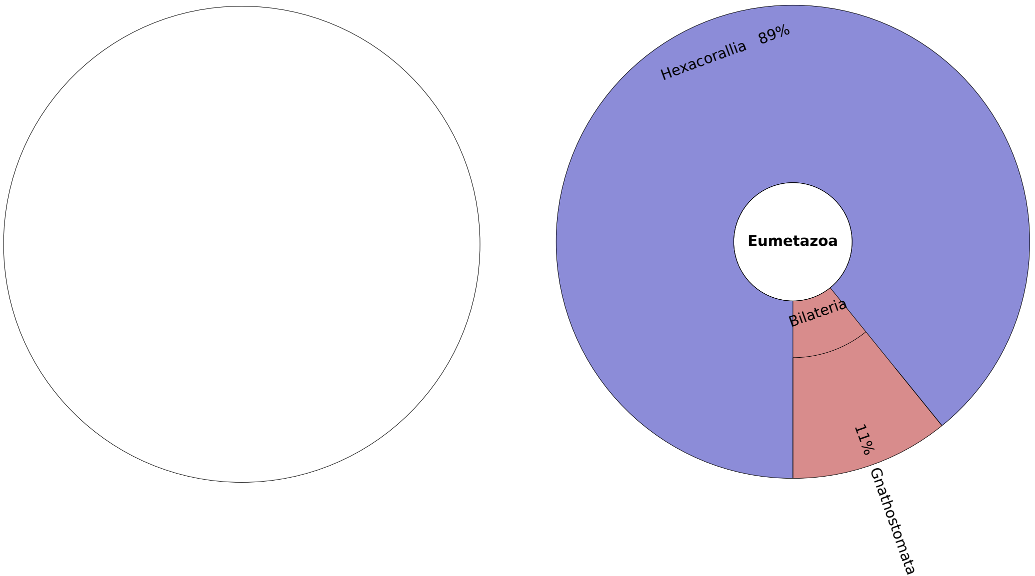


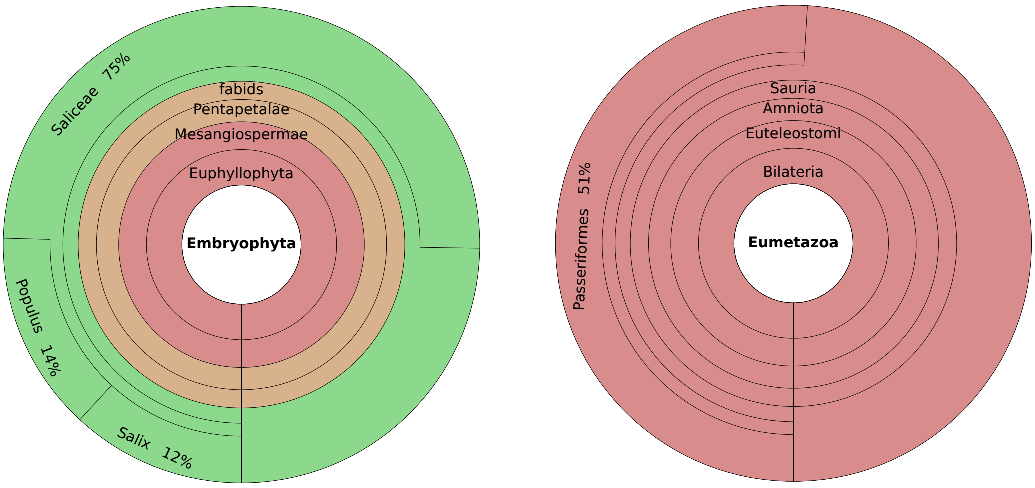
Figure 4.58: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF054. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Samples 291, 315, 330, and 356 (†) do not include data from replicates because they were associated with excessive contamination. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.13.2 Taxonomic profiles

ELF054_356†



ELF054_330†



ELF054_315†

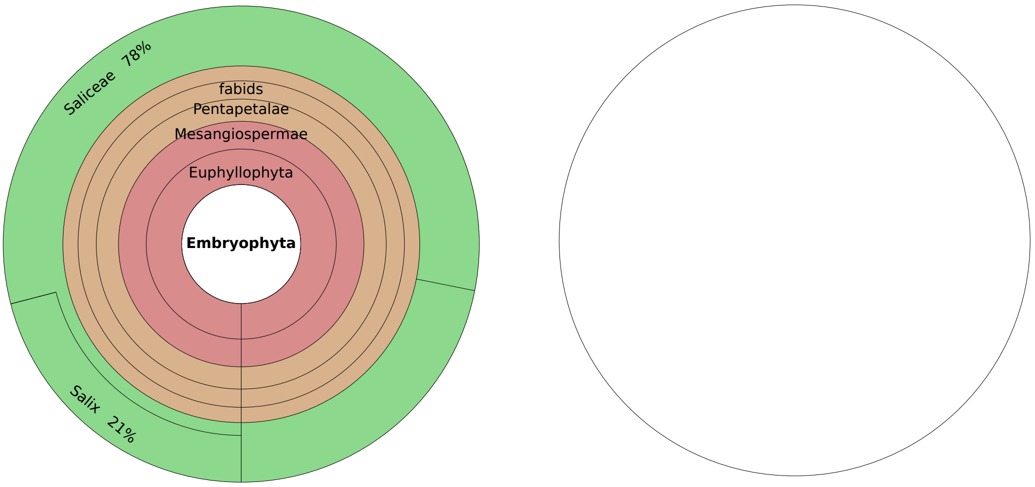
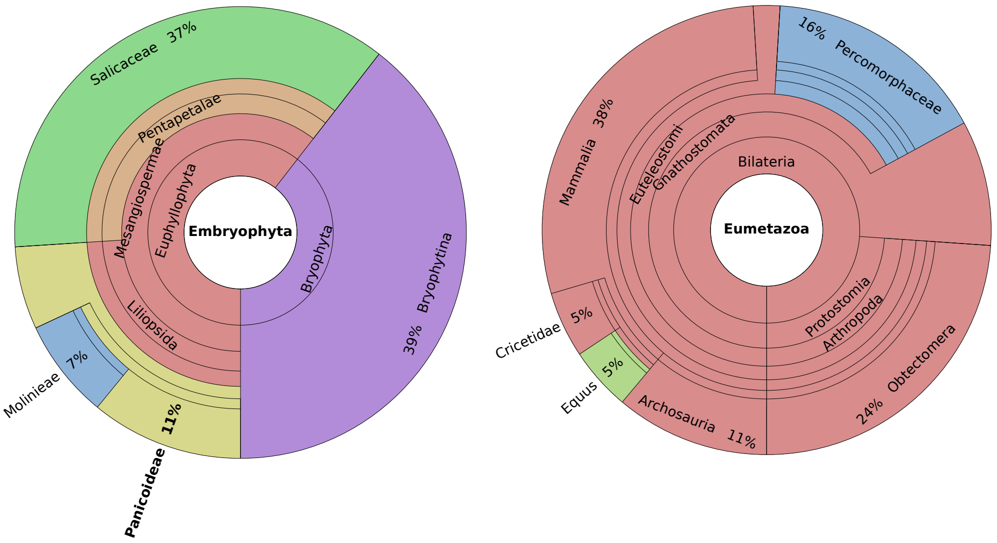
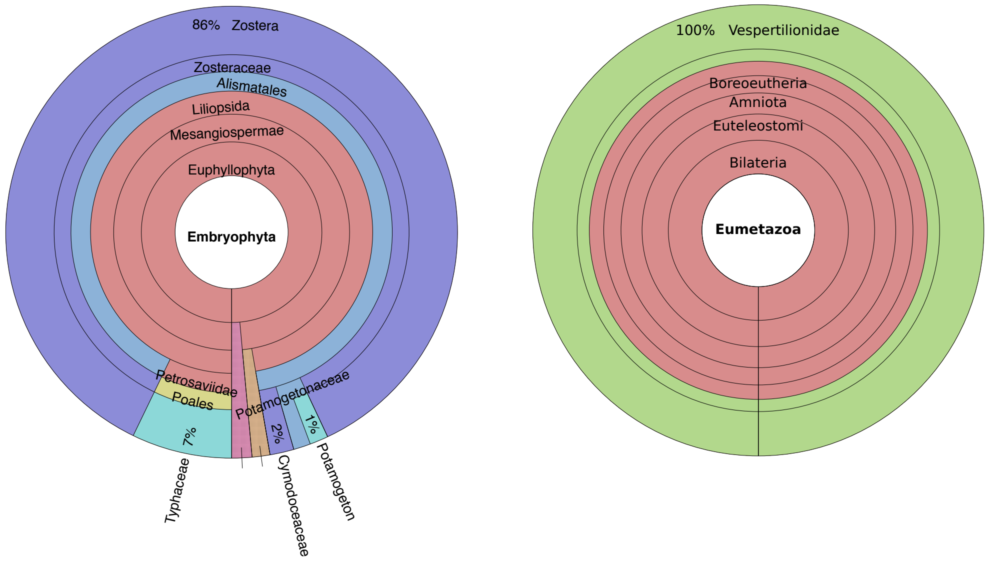


Figure 4.59: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF054, samples 356, 330, and 315. None include data from replicates because they were associated with excessive contamination. Continued in figure 4.60. See figure 4.4 for colour key.

ELF054_291†



ELF054_268



ELF054_182

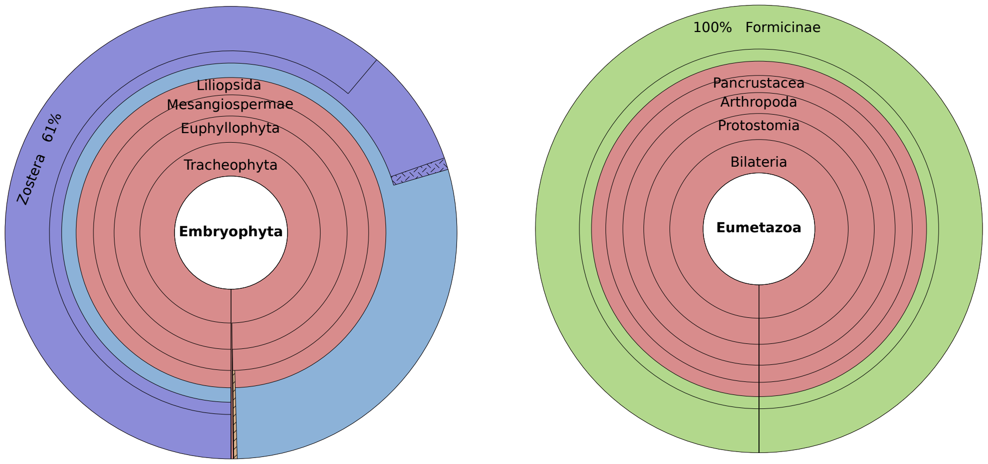
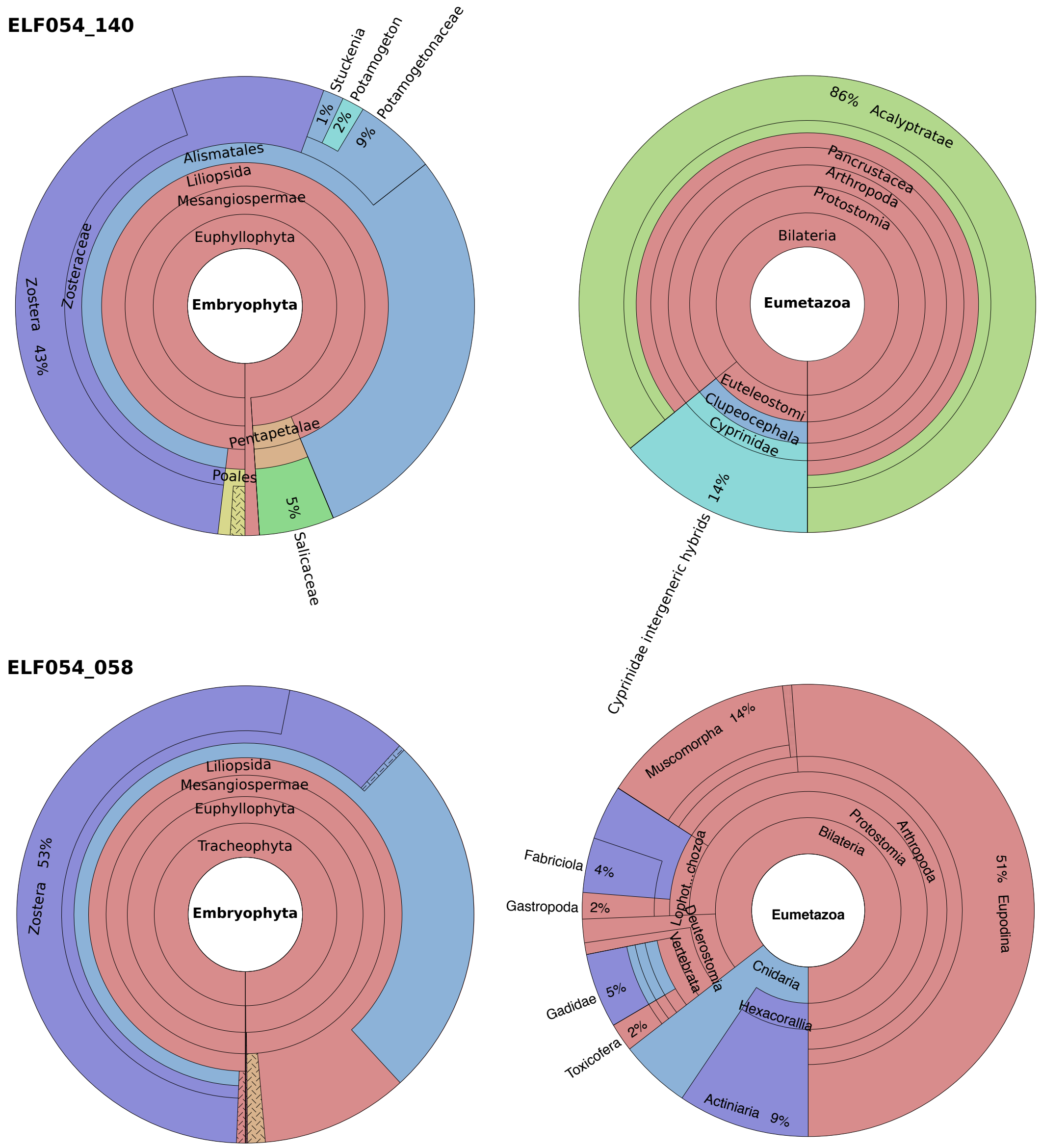


Figure 4.60: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF054, samples 291, 268, and 182. Sample 291 (†) does not include data from the replicate sequencing run because it was associated with excessive contamination. Continued in figure 4.61. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.


ELF054_140

ELF054_058

Figure 4.61: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF054, samples 140 and 58. See figure 4.4 for colour key.

Table 4.25: Biogenomic masses of European taxa in Embryophyta from ELF054. Samples are in cm. The deeper samples (†) do not include data from duplicates as they were associated with excessive contamination. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	356†	330†	315†	291†	268	182	140	58
		Ingroup	1117.14	0.57	14.58	41.28	5.76	40.61	216.03	34.77	763.53
Freshwater aquatics	Pondweeds	Potamogeton	1.62	0.00	0.00	0.00	0.00	0.54	0.00	0.54	0.54
Freshwater aquatics	Bulrush family	Typhaceae	2.91	0.00	0.00	0.00	0.00	2.91	0.00	0.00	0.00
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	2.79	0.00	0.00	0.00	0.00	0.70	1.39	0.00	0.70
Salt/brackish aquatics	Tasselweeds	Ruppia	1.72	0.00	0.00	0.00	0.00	0.00	0.43	0.00	1.29
Salt/brackish aquatics	Eelgrass family	Zosteraceae	88.07	0.00	0.00	0.00	0.00	0.00	18.35	3.67	66.06
Salt/brackish aquatics	Eelgrasses	Zostera	581.65	0.00	0.00	0.00	0.00	34.86	132.11	14.68	400.00
Mixed aquatics	Alismatids	Alismatales	270.41	0.00	0.00	0.00	0.00	0.00	62.81	10.02	197.59
Mixed aquatics	Pondweed family	Potamogetonaceae	2.98	0.00	0.00	0.00	0.00	0.50	0.00	1.98	0.50
Mixed aquatics	Pondweeds	Stuckenia	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00
Mixed aquatics	Common reed tribe	Molinieae	0.71	0.00	0.00	0.00	0.35	0.00	0.00	0.35	0.00
Trees/shrubs	Willow family	Salicaceae	19.86	0.00	3.61	9.03	1.81	0.00	0.00	1.81	3.61
Trees/shrubs	Willow tribe	Saliceae	32.68	0.00	7.26	23.60	0.00	0.00	0.00	0.00	1.82
Trees/shrubs	Poplars	Populus	1.98	0.00	1.98	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Willows	Salix	12.11	0.00	1.73	8.65	0.00	0.00	0.00	0.00	1.73
Trees/shrubs	Elm family	Ulmaceae	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50
Herbs	Plantains	Plantago	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.46
Herbs	Arum subfamily	Aroideae	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.27
Herbs	Orchids	Orchidaceae	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17
Trees/shrubs and herbs		campanulids	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31
Trees/shrubs and herbs	Daisy family	Asteraceae	0.81	0.00	0.00	0.00	0.00	0.00	0.27	0.00	0.54
Trees/shrubs and herbs		Asteroideae	0.25	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00
Trees/shrubs and herbs	Aster tribe	Astereae	1.04	0.00	0.00	0.00	0.00	0.52	0.00	0.00	0.52
Trees/shrubs and herbs		Asparagales	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08
Grasses and relatives	Grass order	Poales	1.46	0.00	0.00	0.00	0.29	0.00	0.00	0.29	0.88
Grasses and relatives	Sedges	Cyperaceae	1.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.54
Grasses and relatives	Rushes	Juncus	1.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.52
Grasses and relatives		**Panicoideae**	**1.08**	**0.00**	**0.00**	**0.00**	**0.54**	**0.00**	**0.00**	**0.00**	**0.54**
Ferns	Leptosporangiate ferns	Polypodiidae	0.08	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00
Ferns	Polypod ferns	Polypodiales	0.56	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.40
Ferns		Thelypteridaceae	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20
Ferns	Bracken family	Dennstaedtiaceae	0.21	0.00	0.00	0.00	0.00	0.21	0.00	0.00	0.00
Ferns		Polypodiineae	0.06	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.00
Ferns		Polypodiaceae	0.23	0.00	0.00	0.00	0.00	0.08	0.08	0.00	0.08
Bryophytes		Bryophytina	1.95	0.00	0.00	0.00	1.95	0.00	0.00	0.00	0.00
Mixed		Tracheophyta	0.86	0.00	0.00	0.00	0.00	0.00	0.35	0.00	0.52
Mixed		Euphyllophyta	0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.00
Mixed	Monocots	Liliopsida	79.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	79.54

Table 4.26: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF054. Samples are in cm. The deeper samples (†) do not include data from duplicates as they were associated with excessive contamination. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	356†	330†	315†	291†	268	182	140	58
			Ingroup	45.69	1.81	4.93	0.00	6.79	0.42	3.42	5.07	23.24
Freshwater aquatics	Fishes	Carp family hybrids	Cyprinidae intergeneric hybrids	0.72	0.00	0.00	0.00	0.00	0.00	0.00	0.72	0.00
Salt/brackish aquatics	Fishes	Cod family	Gadidae	1.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.23
Salt/brackish aquatics	Annelids		Fabriciidae	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.90
Salt/brackish aquatics	Annelids		Fabriciola	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.90
Salt/brackish aquatics	Cnidarians	Hexacorals	Hexacorallia	1.62	1.62	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Sea anemones	Sea anemones	Actiniaria	2.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.17
Mixed aquatics	Fishes		Percomorphaceae	1.06	0.00	0.00	0.00	1.06	0.00	0.00	0.00	0.00
Mixed aquatics	Cnidarians	Cnidarians	Cnidaria	1.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.17
Terrestrial	Mammals	Vespertilionid bats	Vespertilionidae	0.42	0.00	0.00	0.00	0.00	0.42	0.00	0.00	0.00
Terrestrial	Mammals	Horses, zebras and asses	Equus	0.30	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.00
Terrestrial	Dipterans		Acalyptratae	4.35	0.00	0.00	0.00	0.00	0.00	0.00	4.35	0.00
Terrestrial	Hymenopterans		Formicinae	3.42	0.00	0.00	0.00	0.00	0.00	3.42	0.00	0.00
Mixed	Animals	Animals	Metazoa	0.50	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.25
Mixed	Animals	Deuterostomes	Deuterostomia	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.39
Mixed	Vertebrates	Vertebrates	Vertebrata	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19
Mixed	Vertebrates	Jawed vertebrates	Gnathostomata	0.79	0.20	0.00	0.00	0.59	0.00	0.00	0.00	0.00
Mixed	Vertebrates	Lobe-finned fish	Sarcopterygii	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00
Mixed	Mammals	Mammals	Mammalia	1.87	0.00	0.00	0.00	1.87	0.00	0.00	0.00	0.00
Mixed	Mammals	Hamster family	Cricetidae	0.31	0.00	0.00	0.00	0.31	0.00	0.00	0.00	0.00
Mixed	Vertebrates	Archosaurs	Archosauria	1.46	0.00	0.73	0.00	0.73	0.00	0.00	0.00	0.00
Mixed	Birds	Perching birds	Passeriformes	0.76	0.00	0.76	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Squamates	Snakes and related lizards	Toxicofera	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.47
Mixed	Mites		Eupodina	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	11.76
Mixed	Crustaceans		Pleocyemata	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15
Mixed	Lepidopterans		Obtectomera	1.56	0.00	0.00	0.00	1.56	0.00	0.00	0.00	0.00
Mixed	Dipterans		Muscomorpha	3.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.24
Mixed	Gastropods	Snails and slugs	Gastropoda	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.43

Embryophyta

Again similar to ELF034A, there appears to be a move from woody taxa to aquatics, although the low read counts of deeper samples makes it difficult to be sure. Embryophyta data is absent from the deepest sample. 330, 315, and 291 contain the willow group, which reappears as a minor component of the top two samples alongside a small signal from Ulmaceae in 58. Reeds and Typhaceae (the bulrush family) appear in samples 291 and 268. The remaining four samples contain mostly *Zostera*. This sequence could represent floodplain woodland succeeding to reed bed and then open estuary. The top sample, probably by virtue of its high data yield, also contains a number of other terrestrial plants at low frequency, so the terrestrial influence is not lost.

Metazoa

Ecologically informative taxa from Metazoa do not show a clear pattern. However, the Hexacorallia in the deepest sample suggests some saltwater influence even here. The apparently terrestrial strata in ELF054 may already have been near to the coast.

4.13.3 Pianka scores

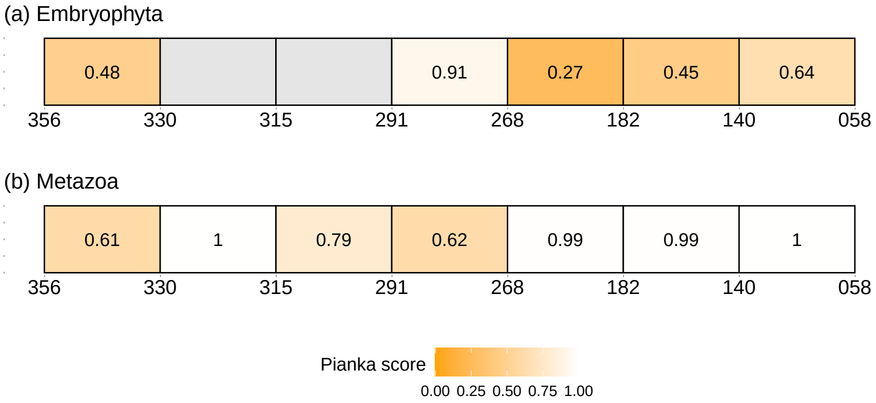


Figure 4.62: Pianka similarity scores between adjacent samples in ELF054 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

Note that the partial overlap between samples 356 and 330 is due to reads assigned to non-European taxa. 356 is devoid of more plausible taxa. The other Pianka scores correspond well to the taxonomic profiles. For Embryophyta, samples 330-291 are relatively similar, there is a clear change in 268 as the aquatics appear, and the remaining samples are nearly identical. The Metazoa data is much more varied, so Pianka scores are generally low. Sample 268 contains only one read, but it overlaps with several higher taxa in the diverse 291. Overall, Embryophyta shows some pattern while Metazoa is chaotic.

4.13.4 Summary

ELF054 resembles ELF034A in the apparent transition from floodplain woodland, through reeds, to a *Zostera*-dominated marine environment, and in the lack of data in apparently terrestrial samples (even allowing for the missing replicate data). This is consistent with salt improving DNA preservation, but limits interpretation of the deeper samples. The Metazoa data, despite relatively high read counts in some samples, is not particularly informative.

4.14 ELF039

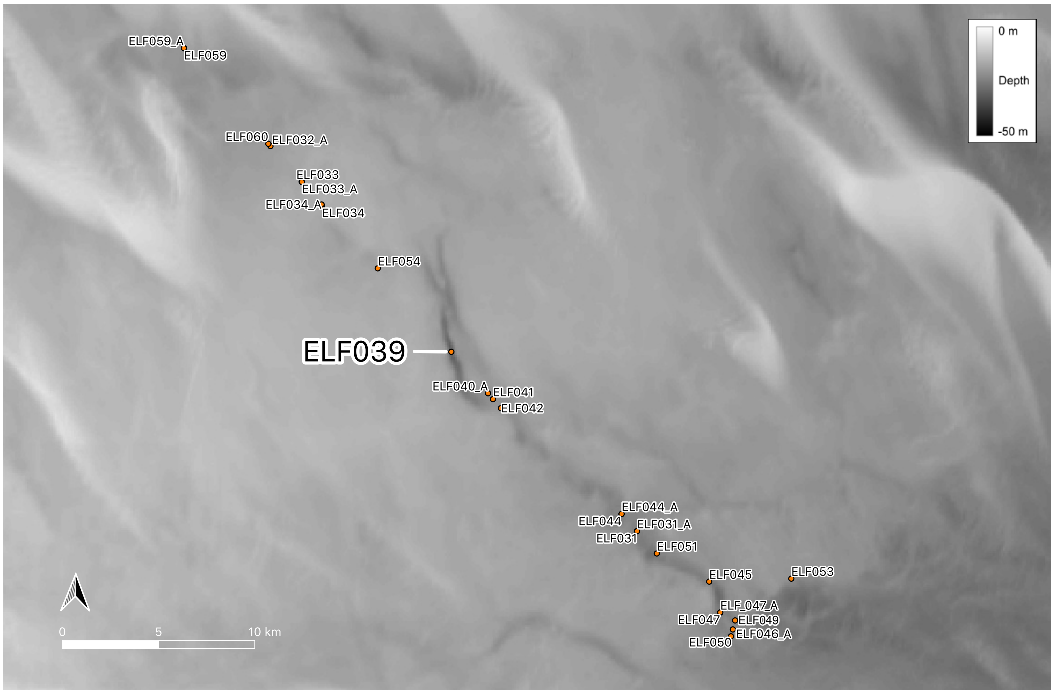


Figure 4.63: Bathymetry map of the palaeochannel transect highlighting ELF039. Darker grey indicates greater depth.

ELF039 is the first core inside the visible palaeochannel. It is approximately half way along the transect, ∼20 km from ELF059.

4.14.1 Read counts

The Embryophyta read count is very variable. The first and final samples have no or nearly no reads, samples 460-341 have between ∼200 and 5,700, and the penultimate two samples have extraordinarily high counts. At nearly 157,000 reads, 321 cm has the largest data yield across all cores. This results in a cramped *x*-axis; many of the deeper samples do have useful amounts of data, despite how they appear on the plot. One of these is sample 341 (*), which was sequenced in greater depth alongside other samples suspected to relate to the Storegga tsunami. The pattern of Metazoa reads is similar, if on a much reduced scale, with the lowest counts at the very start and end and relatively large numbers of reads (for Metazoa) through most of the core.

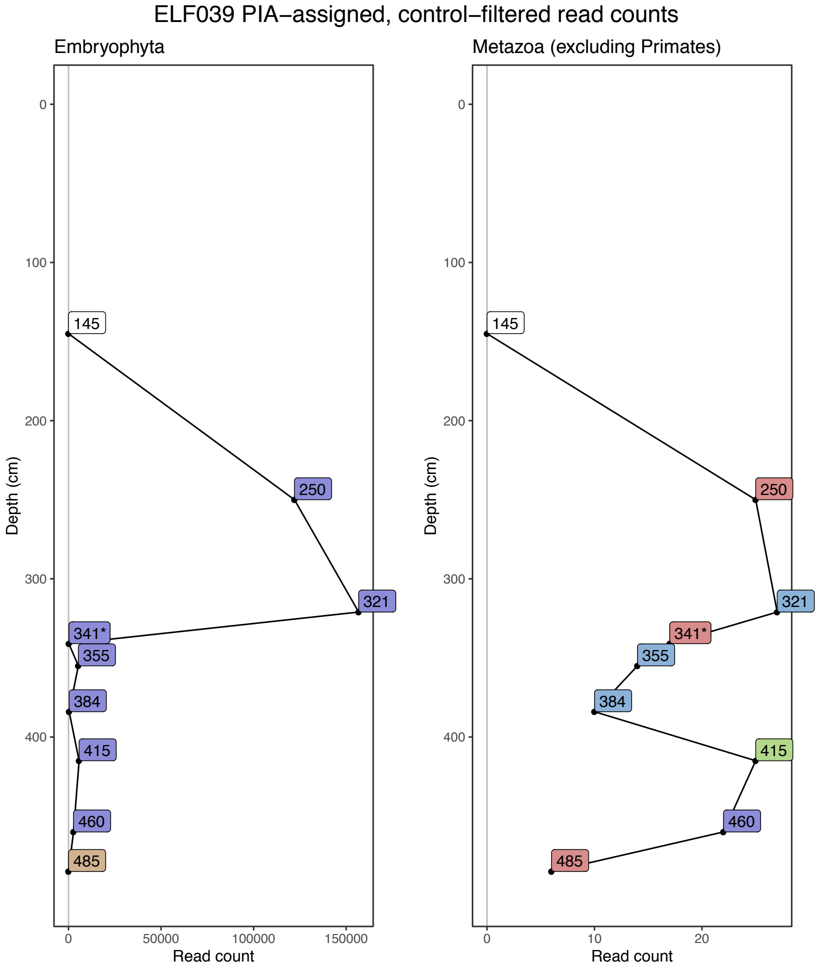
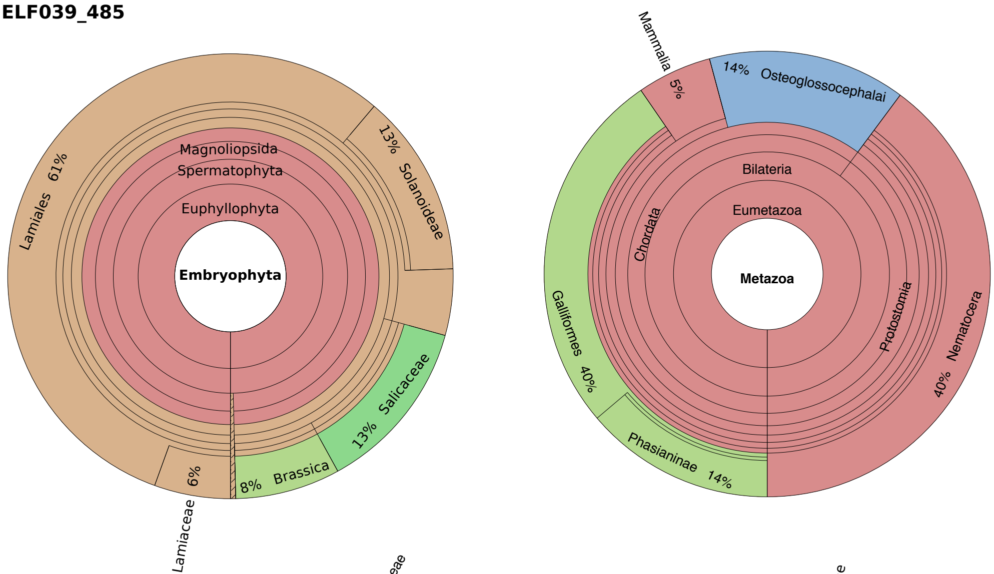


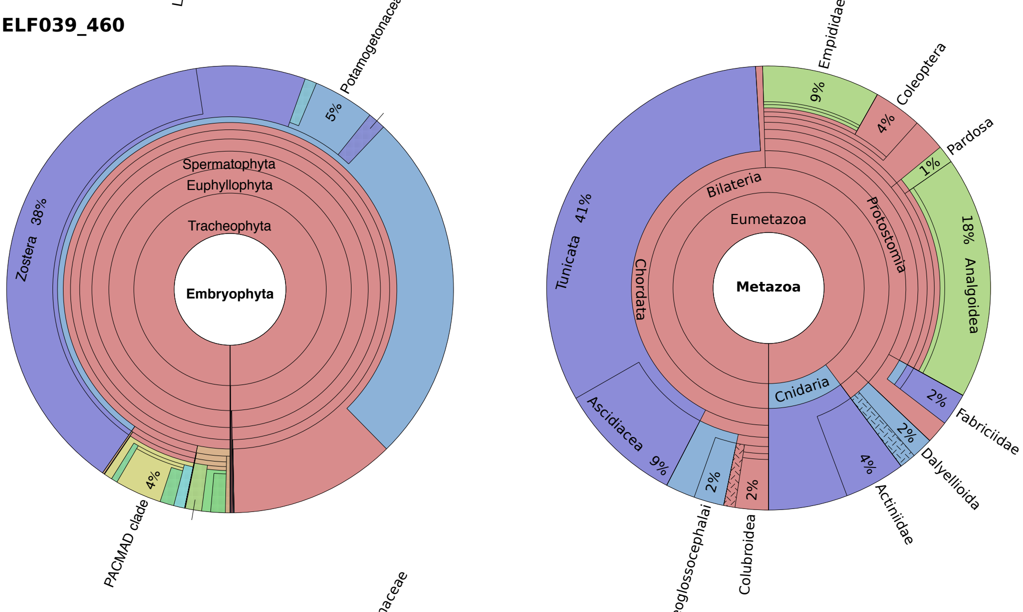
Figure 4.64: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF033. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample 341(*) was sequenced in greater depth. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.14.2 Taxonomic profiles

ELF039_485



ELF039_460



ELF039_415

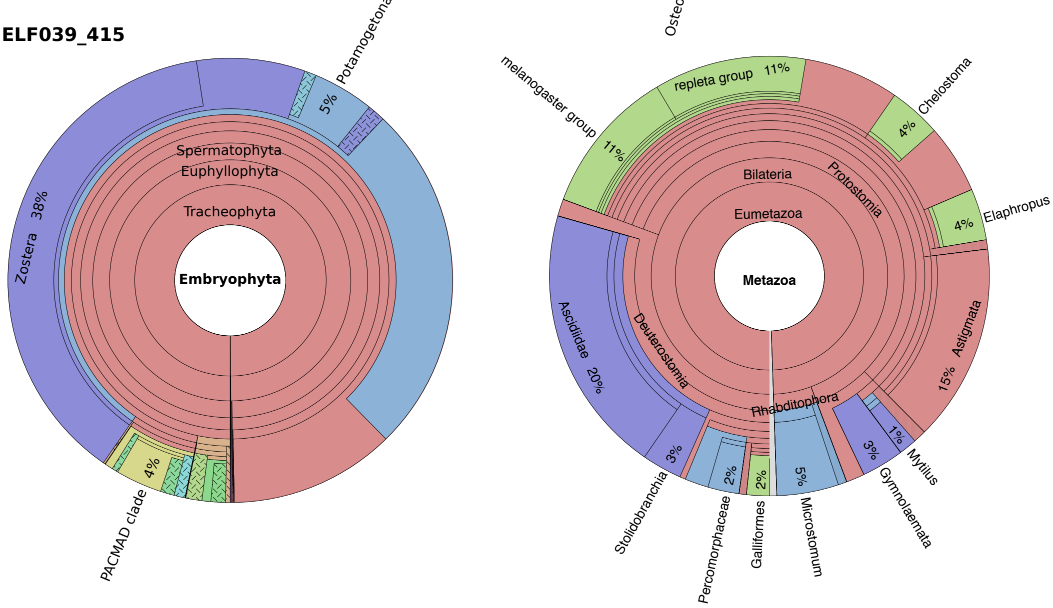
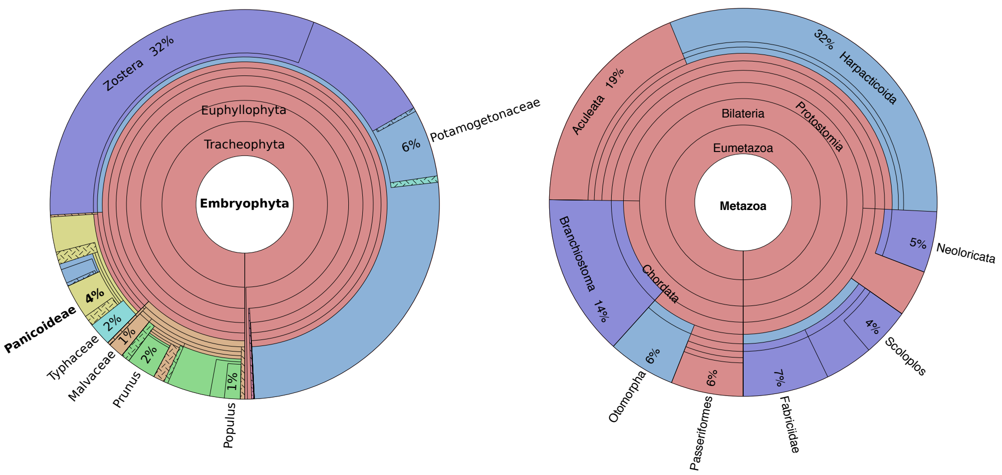
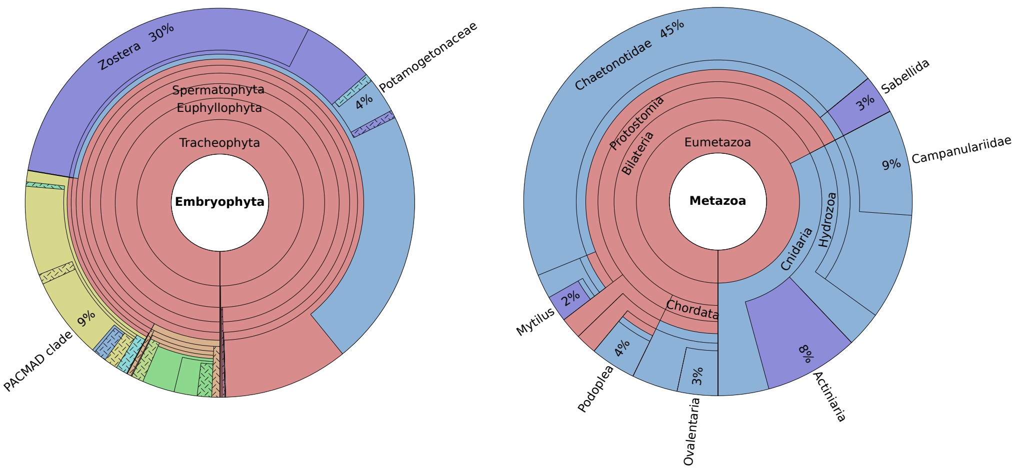


Figure 4.65: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF039, samples 485, 460, and 415. Continued in figure 4.66. See figure 4.4 for colour key.

ELF039_384



ELF039_355



ELF039_341*

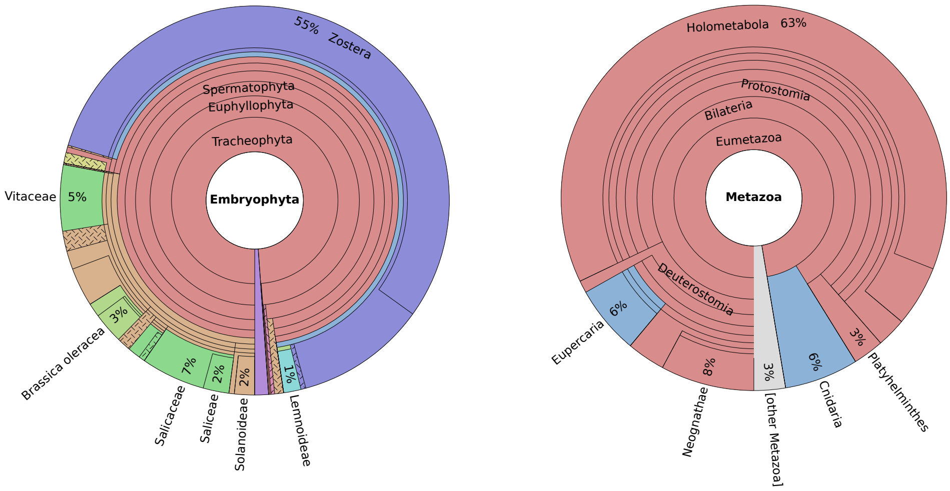
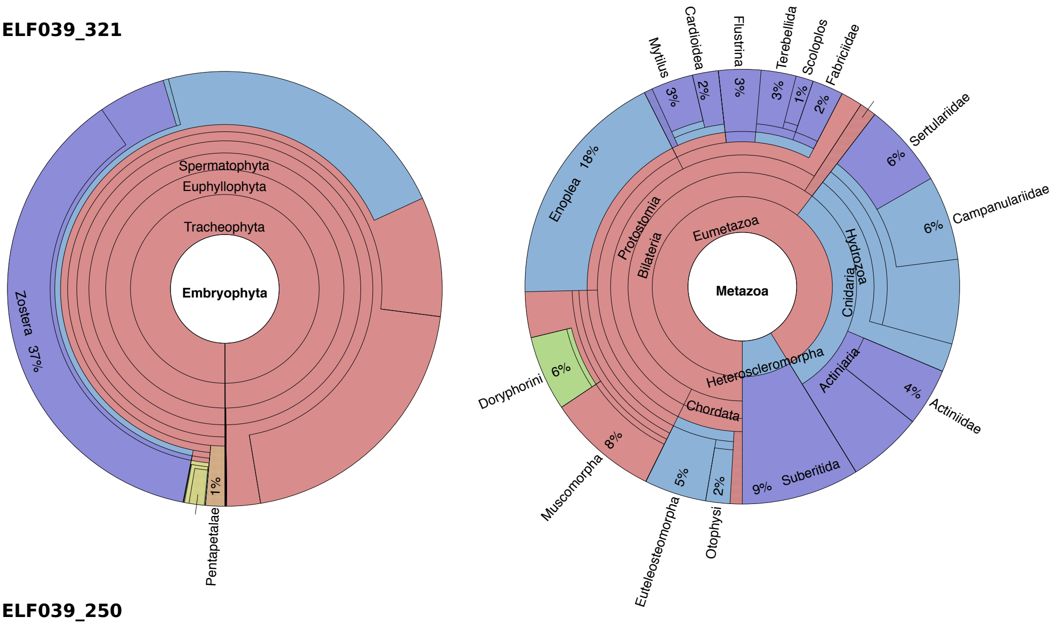
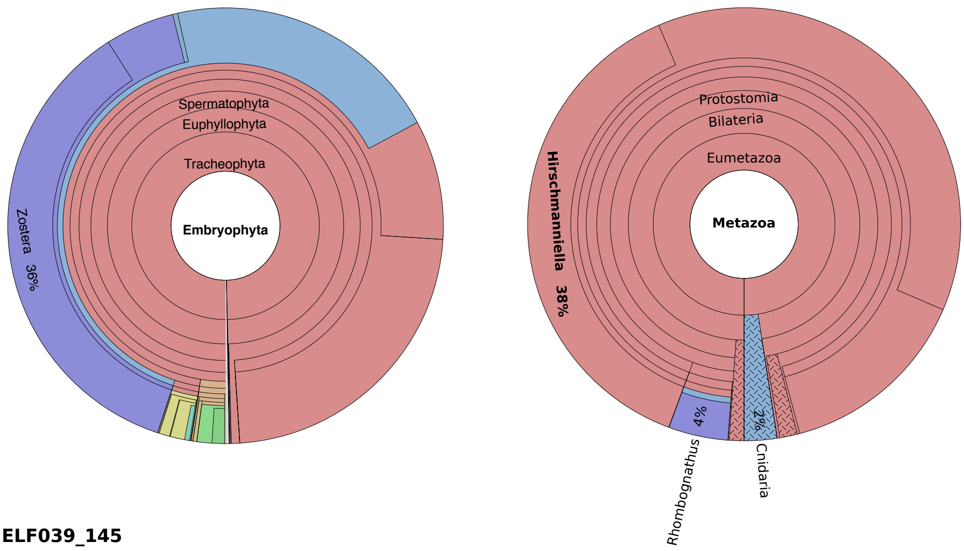


Figure 4.66: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF039, samples 384, 355, and 341. 341* was sequenced in greater depth. Continued in figure 4.67. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

ELF039_321



ELF039_250



ELF039_145

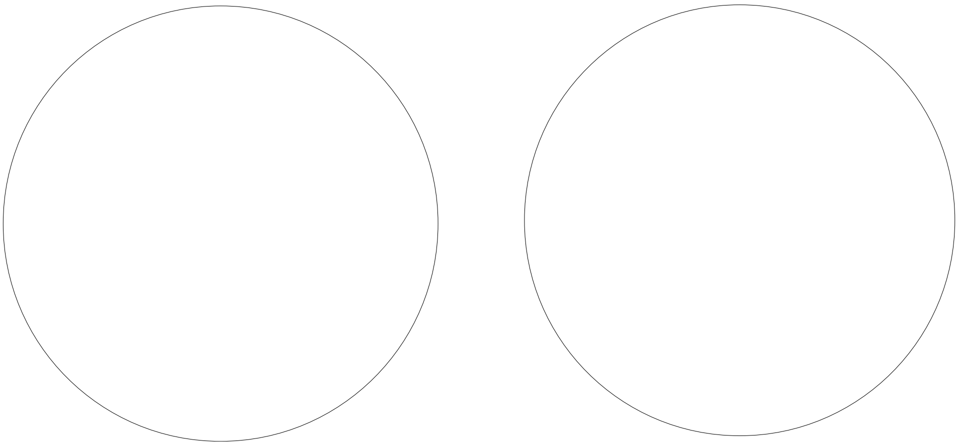


Figure 4.67: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF039, samples 321, 450, and 145. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

Table 4.27: Biogenomic masses of European taxa in Embryophyta from ELF039. Samples are in cm. Sample 341 (*) was resequenced in greater depth. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	485	460	415	384	355	341*	321	250	145
		Ingroup	82430.15	14.20	839.30	1817.28	156.44	1623.35	284.93	43623.45	34071.19	0.00
Freshwater aquatics	Coontails	Ceratophyllum	4.35	0.00	0.00	0.00	0.00	0.00	0.00	4.35	0.00	0.00
Freshwater aquatics	Bladderworts	Utricularia	8.59	0.00	0.00	0.00	0.00	0.00	0.00	8.59	0.00	0.00
Freshwater aquatics	Water-milfoil family	Haloragaceae	6.92	0.00	0.00	2.31	0.00	0.00	0.00	4.62	0.00	0.00
Freshwater aquatics	Water-milfoils	Myriophyllum	18.46	0.00	0.00	4.62	0.00	2.31	0.00	9.23	2.31	0.00
Freshwater aquatics	Water-plantain family	Alismataceae	0.90	0.00	0.00	0.42	0.00	0.07	0.00	0.28	0.14	0.00
Freshwater aquatics	Arrowheads	Sagittaria	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.12	0.00
Freshwater aquatics	Duckweed family	Lemnoideae	19.52	0.00	0.00	1.22	0.00	0.00	2.44	8.54	7.32	0.00
Freshwater aquatics	Elodea family	Hydrocharitaceae	2.37	0.00	0.59	0.20	0.00	0.20	0.00	0.59	0.79	0.00
Freshwater aquatics	Pondweeds	Potamogeton	58.22	0.00	3.77	13.48	0.00	5.39	0.00	19.95	15.63	0.00
Freshwater aquatics	Bulrush family	Typhaceae	61.17	0.00	2.91	5.83	2.91	11.65	0.00	17.48	20.39	0.00
Freshwater aquatics	Bulrushes	Typha	33.96	0.00	3.77	3.77	0.00	3.77	0.00	15.09	7.55	0.00
Freshwater aquatics	Waterlily order	Nymphaeales	2.26	0.00	0.00	0.00	0.00	0.00	0.00	2.26	0.00	0.00
Freshwater aquatics	Waterlily family	Nymphaeaceae	3.19	0.00	0.53	0.53	0.00	0.53	0.00	1.06	0.53	0.00
Freshwater aquatics	Waterlilies	Nymphaea	0.77	0.00	0.00	0.00	0.00	0.77	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	31.36	0.00	4.18	4.18	0.00	5.57	0.70	11.15	5.57	0.00
Salt/brackish aquatics	Tasselweeds	Ruppia	30.90	0.00	5.58	7.73	0.43	4.29	0.00	5.58	7.30	0.00
Salt/brackish aquatics	**Neptune-grasses**	**Posidonia**	**1.60**	**0.00**	**0.32**	**0.00**	**0.00**	**0.32**	**0.00**	**0.64**	**0.32**	**0.00**
Salt/brackish aquatics	Eelgrass family	Zosteraceae	4192.66	0.00	66.06	67.89	16.51	97.25	18.35	2176.15	1750.46	0.00
Salt/brackish aquatics	Eelgrasses	Zostera	29979.82	0.00	317.43	554.13	47.71	480.73	93.58	16269.72	12216.51	0.00
Mixed aquatics	Alismatids	Alismatales	17742.34	0.00	214.39	380.40	39.21	344.25	0.00	9731.28	7032.81	0.00
Mixed aquatics	Pondweed family	Potamogetonaceae	387.77	0.00	37.19	102.64	8.43	47.11	0.00	111.57	80.83	0.00
Mixed aquatics	Pondweeds	Stuckenia	47.60	0.00	3.47	14.88	0.50	5.45	0.00	13.39	9.92	0.00
Mixed aquatics		Fuireneae	1.79	0.00	0.00	0.00	0.00	0.00	0.00	1.79	0.00	0.00
Mixed aquatics	Club-rushes	Bolboschoenus	1.79	0.00	0.00	0.00	0.00	1.79	0.00	0.00	0.00	0.00
Mixed aquatics	Reeds	Arundinoideae	43.85	0.00	0.71	5.31	0.71	3.89	0.35	18.04	14.85	0.00
Mixed aquatics	Common reed tribe	Molinieae	194.95	0.00	2.47	19.74	1.76	14.10	0.00	88.13	68.74	0.00
Mixed aquatics	Common reeds	Phragmites	10.92	0.00	0.84	0.42	0.42	2.94	0.00	2.52	3.78	0.00
Halophytes	Goosefoots	Chenopodium	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.80	0.00
Halophytes	Glasswort subfamily	Salicornioideae	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.46	0.00	0.00
Halophytes		Salicornia subg. Salicornia	1.45	0.00	0.00	0.00	0.00	0.72	0.00	0.00	0.72	0.00
Halophytes	Leadwort family	Plumbaginaceae	5.58	0.00	0.00	0.00	0.00	0.00	0.21	5.17	0.21	0.00
Halophytes	Sea-lavenders	Limonium	2.13	0.00	0.00	0.00	0.00	0.00	0.00	1.28	0.85	0.00
Halophytes	**Desert thumb**	**Cynomorium**	**0.19**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.19**	**0.00**	**0.00**
Halophytes	Saltmarsh grass	Puccinellia	0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.28	0.00	0.00
Xerophytes	Mugworts	Artemisia	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.19	0.00
Xerophytes	Goosefoot family	Chenopodiaceae	3.98	0.00	0.00	0.00	0.00	0.00	0.00	3.98	0.00	0.00
Xerophytes	Atriplex tribe	Atripliceae	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.80	0.00
Trees/shrubs		Cupressales	0.15	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.07	0.00
Trees/shrubs	Cypress family	Cupressaceae	1.25	0.00	0.00	0.00	0.00	0.00	1.25	0.00	0.00	0.00
Trees/shrubs	Elder family	Adoxaceae	0.40	0.00	0.00	0.00	0.00	0.13	0.00	0.26	0.00	0.00
Trees/shrubs		Viburnum	1.58	0.00	0.00	0.00	0.00	0.26	0.00	1.05	0.26	0.00
Trees/shrubs		Cornales	1.23	0.00	0.00	0.61	0.00	0.00	0.00	0.61	0.00	0.00
Trees/shrubs	Dogwoods	Cornus	1.19	0.00	0.00	0.59	0.00	0.00	0.00	0.00	0.59	0.00

Continued on next page

Table 4.27 continued

Ecological category	Common name	Taxon	Total	485	460	415	384	355	341*	321	250	145
		Ingroup	82430.15	14.20	839.30	1817.28	156.44	1623.35	284.93	43623.45	34071.19	0.00
Trees/shrubs	Heather family	Ericaceae	0.67	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.00	0.00
Trees/shrubs	Birch family	Betulaceae	41.04	0.00	3.24	1.08	1.08	4.32	0.00	24.84	6.48	0.00
Trees/shrubs	Alders	Alnus	9.68	0.00	0.00	0.00	0.00	3.23	0.00	0.00	6.45	0.00
Trees/shrubs	Birches	Betula	1.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.94	0.00
Trees/shrubs	Hazels	Corylus	9.09	0.00	2.27	4.55	0.00	2.27	0.00	0.00	0.00	0.00
Trees/shrubs	Oaks	Quercus	6.35	0.00	1.06	0.00	0.00	0.00	0.00	1.06	4.23	0.00
Trees/shrubs	English oak	Quercus robur	1.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.11	0.00
Trees/shrubs	Willow family	Salicaceae	682.31	1.81	5.42	57.76	5.42	43.32	9.03	182.31	377.26	0.00
Trees/shrubs	Willow tribe	Saliceae	326.78	0.00	3.63	30.86	1.82	30.86	3.63	99.85	156.13	0.00
Trees/shrubs	Poplars	Populus	77.14	0.00	0.00	29.67	1.98	1.98	0.00	19.78	23.74	0.00
Trees/shrubs	Willows	Salix	283.74	0.00	5.19	38.06	0.00	17.30	0.00	88.24	134.95	0.00
Trees/shrubs		Amygdaleae	4.42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.42	0.00
Trees/shrubs	Stone fruit trees	Prunus	104.09	0.00	1.83	7.30	3.65	1.83	0.00	34.70	54.79	0.00
Trees/shrubs	Apple tribe	Maleae	10.70	0.00	0.89	0.89	0.00	2.67	1.78	1.78	2.67	0.00
Trees/shrubs	Hawthorns	Crataegus	0.86	0.00	0.00	0.00	0.00	0.00	0.86	0.00	0.00	0.00
Trees/shrubs	Apples	Malus	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00
Trees/shrubs	Blackberry and relatives	Rubus	4.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.44	0.00
Trees/shrubs	Elm family	Ulmaceae	3.51	0.00	0.50	0.50	0.50	0.50	0.00	1.00	0.50	0.00
Trees/shrubs	Elms	Ulmus	0.86	0.00	0.00	0.00	0.00	0.00	0.00	0.86	0.00	0.00
Trees/shrubs	Lime subfamily	Tilioideae	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.80	0.00	0.00
Trees/shrubs	Limes	Tilia	4.80	0.00	0.00	0.80	0.00	0.00	0.00	2.40	1.60	0.00
Trees/shrubs	**Cashew family**	**Anacardiaceae**	**1.83**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**1.83**	**0.00**	**0.00**	**0.00**
Trees/shrubs	Soapberry family	Sapindaceae	0.82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.82	0.00
Trees/shrubs	**Grape family**	**Vitaceae**	**9.29**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**9.29**	**0.00**	**0.00**	**0.00**
Trees/shrubs	Barberry subfamily	Berberidoideae	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00
Herbs	Celeries	Apieae	0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.51	0.00
Herbs	Water-parsnip tribe	Oenantheae	0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.79	0.00
Herbs	Thistle tribe	Cardueae	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.48	0.00	0.00
Herbs	Primrose family	Primulaceae	0.35	0.00	0.00	0.00	0.00	0.35	0.00	0.00	0.00	0.00
Herbs	Madder tribe	Rubieae	0.71	0.00	0.00	0.71	0.00	0.00	0.00	0.00	0.00	0.00
Herbs		**Trichosporeae**	**1.05**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**1.05**	**0.00**	**0.00**
Herbs	Dead-nettle subfamily	Lamioideae	0.63	0.00	0.00	0.00	0.00	0.63	0.00	0.00	0.00	0.00
Herbs	Catmint subfamily	Nepetoideae	11.29	0.00	0.00	0.00	0.00	0.00	0.00	7.91	3.39	0.00
Herbs	Mint tribe	Mentheae	3.68	0.00	1.23	0.00	0.00	1.23	0.00	0.00	1.23	0.00
Herbs	Mints	Mentha	1.52	0.00	0.00	0.00	0.00	0.00	0.00	1.52	0.00	0.00
Herbs	Sages	Salvia	2.63	0.00	0.00	0.00	0.00	0.00	0.00	2.63	0.00	0.00
Herbs	Skullcaps	Scutellaria	1.72	0.00	0.00	0.00	0.00	0.00	0.00	1.72	0.00	0.00
Herbs	Plantains	Plantago	0.92	0.00	0.46	0.00	0.00	0.46	0.00	0.00	0.00	0.00
Herbs	Nightshades	Solanum	0.71	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.71	0.00
Herbs	Pink family	Caryophyllaceae	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.00	0.00
Herbs	Sundew family	Droseraceae	1.54	0.00	0.00	0.00	0.00	0.00	0.00	1.54	0.00	0.00
Herbs		NPAAA clade	0.36	0.00	0.00	0.00	0.00	0.00	0.36	0.00	0.00	0.00
Herbs		Hologalegina	0.56	0.00	0.00	0.28	0.00	0.00	0.00	0.00	0.28	0.00
Herbs	**Liquorice genus**	**Glycyrrhiza**	**0.67**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.67**	**0.00**	**0.00**	**0.00**
Herbs	Spurge family	Euphorbiaceae	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31	0.00
Herbs	Hemp family	Cannabaceae	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.44	0.00	0.00
Herbs	Avens	Geum	0.63	0.00	0.00	0.00	0.00	0.00	0.00	0.63	0.00	0.00
Herbs	Cinquefoil tribe	Potentilleae	1.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.27	0.00
Herbs	Strawberry subtribe	Fragariinae	1.47	0.00	0.00	0.00	0.00	0.00	0.00	1.47	0.00	0.00
Herbs	Mustard tribe	Brassiceae	1.93	0.00	0.00	0.00	0.00	0.00	1.93	0.00	0.00	0.00

Continued on next page

Table 4.27 continued

Ecological category	Common name	Taxon	Total	485	460	415	384	355	341*	321	250	145
		Ingroup	82430.15	14.20	839.30	1817.28	156.44	1623.35	284.93	43623.45	34071.19	0.00
Herbs	Mustards	Brassica	1.08	1.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Herbs	Wild and cultivated cabbages	Brassica oleracea	4.33	0.00	0.00	0.00	0.00	0.00	4.33	0.00	0.00	0.00
Herbs		Camelineae	3.14	0.00	0.00	0.00	0.00	0.00	0.00	3.14	0.00	0.00
Herbs	Thale cresses	Arabidopsis	3.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.07	0.00
Herbs	Willowherb subfamily	Onagroideae	0.95	0.00	0.00	0.95	0.00	0.00	0.00	0.00	0.00	0.00
Herbs	Arum family	Araceae	2.03	0.00	0.18	1.11	0.37	0.18	0.00	0.00	0.18	0.00
Herbs	Onions, leeks, garlics	Allium	0.24	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00	0.00
Herbs	Asparagus family	Asparagaceae	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.31	0.00	0.00
Herbs	Bluebell family	Hyacinthaceae	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.00
Herbs	Irises	Iris	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.11	0.00
Herbs	Orchids	Orchidaceae	1.70	0.00	0.51	0.00	0.17	0.51	0.00	0.17	0.34	0.00
Herbs		Epidendroideae	0.96	0.00	0.32	0.00	0.00	0.00	0.00	0.32	0.32	0.00
Herbs	Coralroot orchids	Corallorhiza	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.23	0.00
Herbs		Dioscoreales	0.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.93	0.00
Herbs		Dioscoreaceae	0.91	0.00	0.00	0.00	0.00	0.91	0.00	0.00	0.00	0.00
Herbs		Dioscorea	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.90	0.00
Herbs	Bunchflower family	Melanthiaceae	0.03	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00
Herbs		**Leontice**	**0.25**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.25**	**0.00**
Herbs	Poppy subfamily	Papaveroideae	0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.00	0.00
Herbs	Buttercup subfamily	Ranunculoideae	0.32	0.00	0.00	0.11	0.00	0.00	0.00	0.22	0.00	0.00
Herbs	Anemone tribe	Anemoneae	0.16	0.00	0.08	0.00	0.00	0.00	0.00	0.08	0.00	0.00
Herbs	Baneberries	Actaea	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00
Herbs	Buttercup tribe	Ranunculeae	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.31	0.00	0.00
Herbs	Buttercups	Ranunculus	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.00	0.00
Trees/shrubs and herbs	Gymnosperms	Acrogymnospermae	0.16	0.05	0.00	0.00	0.00	0.00	0.00	0.05	0.05	0.00
Trees/shrubs and herbs		Gunneridae	1.41	0.00	0.00	0.47	0.00	0.00	0.00	0.00	0.94	0.00
Trees/shrubs and herbs		campanulids	6.16	0.00	0.00	0.92	0.00	1.54	0.00	0.31	3.39	0.00
Trees/shrubs and herbs	Carrot order	Apiales	0.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.93	0.00
Trees/shrubs and herbs	Ivy family	Araliaceae	2.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.02	0.00
Trees/shrubs and herbs		Asterales	1.14	0.00	0.00	0.00	0.00	0.00	0.00	0.57	0.57	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	4.89	0.00	0.00	0.00	0.54	0.00	0.00	2.44	1.90	0.00
Trees/shrubs and herbs		Asteroideae	2.76	0.00	0.25	0.00	0.00	0.00	0.00	1.00	1.50	0.00
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	1.46	0.00	0.00	0.00	0.00	0.00	0.00	0.37	1.10	0.00
Trees/shrubs and herbs	Aster tribe	Astereae	4.15	0.00	0.00	0.52	0.00	0.52	0.00	1.04	2.07	0.00
Trees/shrubs and herbs		Dipsacales	0.87	0.00	0.00	0.44	0.00	0.00	0.00	0.00	0.44	0.00
Trees/shrubs and herbs		Ericales	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.36	0.00
Trees/shrubs and herbs		lamiids	6.16	0.68	0.00	1.37	0.00	2.05	0.00	0.68	1.37	0.00
Trees/shrubs and herbs		Gentianales	0.90	0.00	0.00	0.00	0.00	0.90	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	12.29	7.90	0.88	0.00	0.00	1.76	0.00	1.76	0.00	0.00
Trees/shrubs and herbs	Dead-nettle family	Lamiaceae	1.57	0.78	0.00	0.00	0.00	0.00	0.00	0.78	0.00	0.00
Trees/shrubs and herbs		Solanales	1.88	0.00	0.00	0.00	0.00	0.38	0.75	0.38	0.38	0.00
Trees/shrubs and herbs	Bindweed family	Convolvulaceae	0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.00
Trees/shrubs and herbs	Nightshade subfamily	Solanoideae	4.73	1.89	0.00	0.00	0.00	0.00	2.84	0.00	0.00	0.00
Trees/shrubs and herbs	Amaranth family	Amaranthaceae	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00
Trees/shrubs and herbs	Knotweed family	Polygonaceae	3.31	0.00	0.00	1.10	0.00	0.55	0.00	1.66	0.00	0.00
Trees/shrubs and herbs		robinioid clade	1.30	0.00	0.00	1.30	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	**Bean tribe**	**Phaseoleae**	**0.95**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.95**	**0.00**	**0.00**	**0.00**
Trees/shrubs and herbs	St. John's wort family	Hypericaceae	2.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.36	0.00
Trees/shrubs and herbs	Rose subfamily	Rosoideae	12.52	0.00	0.00	0.00	1.39	1.39	0.00	1.39	8.35	0.00
Trees/shrubs and herbs		Rosoideae incertae sedis	1.88	0.00	0.00	1.88	0.00	0.00	0.00	0.00	0.00	0.00

Continued on next page

Table 4.27 continued

Ecological category	Common name	Taxon	Total	485	460	415	384	355	341*	321	250	145
		Ingroup	82430.15	14.20	839.30	1817.28	156.44	1623.35	284.93	43623.45	34071.19	0.00
Trees/shrubs and herbs	Mustard order	Brassicales	5.19	0.00	0.00	0.00	0.00	0.00	2.60	0.00	2.60	0.00
Trees/shrubs and herbs	Mustard family	Brassicaceae	5.44	0.00	0.00	0.00	0.00	0.00	5.44	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow order	Malvales	4.68	0.00	0.00	0.00	0.00	0.00	0.00	3.51	1.17	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	120.73	0.00	0.00	1.96	1.96	1.31	0.00	80.27	35.24	0.00
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	8.33	0.00	0.00	1.11	0.00	0.00	0.00	6.66	0.56	0.00
Trees/shrubs and herbs		Myrtales	1.98	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.99	0.00
Trees/shrubs and herbs	Mistletoe family	Viscaceae	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
Trees/shrubs and herbs		Saxifragales	5.01	0.00	0.39	0.77	0.39	1.73	0.00	1.35	0.39	0.00
Trees/shrubs and herbs		Asparagales	1.44	0.00	0.32	0.40	0.08	0.16	0.00	0.32	0.16	0.00
Trees/shrubs and herbs	**Palms**	**Arecaceae**	**1.73**	**0.00**	**0.29**	**0.29**	**0.00**	**0.00**	**0.00**	**0.00**	**1.15**	**0.00**
Trees/shrubs and herbs		Liliales	0.17	0.00	0.00	0.03	0.00	0.00	0.00	0.07	0.07	0.00
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.78	0.00	0.00	0.16	0.00	0.16	0.00	0.31	0.16	0.00
Trees/shrubs and herbs	Buttercup family	Ranunculaceae	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00
Grasses and relatives	Grass order	Poales	85.20	0.00	4.98	12.30	4.39	14.93	1.17	16.10	31.33	0.00
Grasses and relatives	Sedges	Cyperaceae	20.05	0.00	0.00	1.54	1.54	3.08	0.00	13.88	0.00	0.00
Grasses and relatives		Cyperoideae	1.69	0.00	0.00	0.00	0.00	0.00	0.00	1.69	0.00	0.00
Grasses and relatives	True sedges	Carex	15.94	0.00	2.28	2.28	0.00	0.00	0.00	9.11	2.28	0.00
Grasses and relatives	Rushes	Juncaceae	1.04	0.00	0.00	0.00	0.00	1.04	0.00	0.00	0.00	0.00
Grasses and relatives	Rushes	Juncus	1.52	0.00	1.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grasses and relatives	Grass family	Poaceae	642.52	0.00	0.00	104.33	0.00	118.93	0.00	147.35	271.92	0.00
Grasses and relatives		BOP clade	60.67	0.00	0.00	12.07	0.00	13.24	0.00	17.76	17.60	0.00
Grasses and relatives		Poeae	0.18	0.00	0.00	0.00	0.00	0.18	0.00	0.00	0.00	0.00
Grasses and relatives		Agrostidinae	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.00
Grasses and relatives		Coleanthinae	0.42	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.00	0.00
Grasses and relatives	Meadow grass subtribe	Poinae	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00
Grasses and relatives	Barley subtribe	Hordeinae	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00
Grasses and relatives		PACMAD clade	938.60	0.00	28.28	91.78	0.00	109.64	0.00	345.77	363.14	0.00
Grasses and relatives		Chloridoideae	2.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.88	0.00
Grasses and relatives	Zoysia tribe	Zoysieae	12.12	0.00	0.00	0.00	0.00	0.00	0.00	3.03	9.09	0.00
Grasses and relatives	Dropseeds	Sporobolus	4.76	0.00	0.00	0.00	0.00	0.00	0.00	0.95	3.81	0.00
Grasses and relatives		**Panicoideae**	**54.52**	**0.00**	**3.78**	**9.18**	**4.32**	**13.49**	**0.00**	**9.18**	**14.57**	**0.00**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**8.78**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**4.39**	**4.39**	**0.00**
Grasses and relatives		**Paniceae**	**19.44**	**0.00**	**0.67**	**6.70**	**1.34**	**4.69**	**0.00**	**2.68**	**3.35**	**0.00**
Grasses and relatives		**Cenchrinae**	**1.15**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.58**	**0.58**	**0.00**
Ferns	Ferns	Polypodiopsida	0.54	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.20	0.00
Ferns	Horsetails	Equisetum	0.23	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.14	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	1.15	0.00	0.00	0.08	0.08	0.25	0.00	0.49	0.25	0.00
Ferns	Polypod ferns	Polypodiales	4.50	0.00	0.24	0.08	0.00	0.88	0.32	1.61	1.37	0.00
Ferns	Spleenwort suborder	Aspleniineae	6.84	0.00	0.18	0.27	0.18	0.90	0.18	3.42	1.71	0.00
Ferns		Athyriaceae	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.37	0.00	0.00
Ferns		Athyrium	0.14	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.00	0.00
Ferns		Thelypteridaceae	1.08	0.00	0.10	0.10	0.00	0.10	0.00	0.59	0.20	0.00
Ferns		Phegopteridoideae	0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.28	0.00	0.00
Ferns		Thelypteridoideae	0.42	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.17	0.00
Ferns	Marsh fern	Thelypteris palustris	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00
Ferns	Bracken family	Dennstaedtiaceae	0.21	0.00	0.00	0.00	0.00	0.10	0.00	0.10	0.00	0.00
Ferns		Polypodiineae	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00
Ferns	Buckler and male ferns	Dryopteris	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.15	0.08	0.00

Continued on next page

Table 4.27 continued

Ecological category	Common name	Taxon	Total	485	460	415	384	355	341*	321	250	145
		Ingroup	82430.15	14.20	839.30	1817.28	156.44	1623.35	284.93	43623.45	34071.19	0.00
Ferns		Polypodiaceae	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.46	0.00	0.00
Bryophytes		Bryopsida	11.90	0.00	0.00	1.98	0.00	0.00	0.00	7.94	1.98	0.00
Bryophytes		Bryidae	9.63	0.00	0.00	0.00	0.00	0.00	1.93	3.85	3.85	0.00
Bryophytes	Feather mosses	Hypnales	4.48	0.00	0.00	0.00	0.00	0.00	0.00	4.48	0.00	0.00
Bryophytes	Liverworts	Marchantiophyta	0.53	0.00	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Land plants	Embryophyta	88.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	88.77	0.00
Mixed		Tracheophyta	37.02	0.00	0.17	0.52	0.35	1.90	0.17	16.09	17.82	0.00
Mixed		Euphyllophyta	51.12	0.00	0.69	0.86	0.52	1.89	0.17	20.31	26.68	0.00
Mixed	Seed plants	Spermatophyta	211.84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	211.84	0.00
Mixed	Flowering plants	Magnoliopsida	1085.34	0.00	0.00	0.00	0.00	0.00	0.00	1085.34	0.00	0.00
Mixed	Core angiosperms	Mesangiospermae	16659.22	0.00	0.00	0.00	0.00	0.00	0.00	8849.15	7810.07	0.00
Mixed	Monocots	Liliopsida	7333.93	0.00	100.95	181.65	0.00	167.27	0.00	3868.13	3015.93	0.00
Mixed		commelinids	0.79	0.00	0.00	0.00	0.00	0.00	0.79	0.00	0.00	0.00

Table 4.28: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF039. Samples are in cm. Sample 341 (*) was resequenced in greater depth. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	485	460	415	384	355	341*	321	250	145
			Ingroup	461.84	5.77	37.72	45.92	12.52	27.63	18.79	44.35	269.13	0.00
Salt/brackish aquatics	Invertebrates		Branchiostoma	1.69	0.00	0.00	0.00	1.69	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tunicates	Tunicates	Tunicata	12.20	0.00	12.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tunicates	Sea squirts	Ascidiacea	3.45	0.00	3.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tunicates		Ascidiidae	9.09	0.00	0.00	9.09	0.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tunicates	Lightbulb sea squirts	Stolidobranchia	1.35	0.00	0.00	1.35	0.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Flatworms		Provorticidae	0.29	0.00	0.00	0.29	0.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Mites		Rhombognathus	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	11.76	0.00
Salt/brackish aquatics	Annelids	Tube worms and relatives	Sabellida	0.90	0.00	0.00	0.00	0.00	0.90	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Annelids		Fabriciidae	3.60	0.00	0.90	0.00	0.90	0.00	0.00	0.90	0.90	0.00
Salt/brackish aquatics	Annelids		Orbiniidae	0.51	0.00	0.00	0.00	0.51	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Annelids		Scoloplos	1.01	0.00	0.00	0.00	0.51	0.00	0.00	0.51	0.00	0.00
Salt/brackish aquatics	Annelids		Terebellida	1.01	0.00	0.00	0.00	0.00	0.00	0.00	1.01	0.00	0.00
Salt/brackish aquatics	Bryozoans		Gymnolaemata	1.43	0.00	0.00	1.43	0.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bryozoans		Flustrina	1.22	0.00	0.00	0.00	0.00	0.00	0.00	1.22	0.00	0.00
Salt/brackish aquatics	Bivalves	Cockle superfamily	Cardioidea	1.50	0.00	0.00	0.00	0.00	0.00	0.00	0.75	0.75	0.00
Salt/brackish aquatics	Bivalves	Cockles	Cardiidae	1.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.32	0.00
Salt/brackish aquatics	Bivalves		Mytilus	2.43	0.00	0.00	0.61	0.00	0.61	0.00	1.22	0.00	0.00
Salt/brackish aquatics	Cephalopods		Octopus	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00
Salt/brackish aquatics	Polyplacophorans		Neoloricata	0.64	0.00	0.00	0.00	0.64	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Sea anemones	Sea anemones	Actiniaria	6.50	0.00	2.17	0.00	0.00	2.17	0.00	2.17	0.00	0.00
Salt/brackish aquatics	Sea anemones		Actiniidae	5.02	0.00	1.67	0.00	0.00	0.00	0.00	1.67	1.67	0.00
Salt/brackish aquatics	Sea anemones	Plumose anemones	Metridium	2.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.17	0.00
Salt/brackish aquatics	Hydrozoans		Sertulariidae	2.44	0.00	0.00	0.00	0.00	0.00	0.00	2.44	0.00	0.00
Salt/brackish aquatics	Poriferans		Suberitida	3.41	0.00	0.00	0.00	0.00	0.00	0.00	3.41	0.00	0.00
Mixed aquatics	Fishes		Actinopteri	0.80	0.00	0.00	0.80	0.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Neopterygii	0.81	0.00	0.81	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Osteoglossocephalai	1.65	0.83	0.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Euteleosteomorpha	2.69	0.00	0.00	0.00	0.00	0.00	0.00	1.79	0.90	0.00
Mixed aquatics	Fishes		Percomorphaceae	3.17	0.00	0.00	1.06	0.00	1.06	0.00	0.00	1.06	0.00
Mixed aquatics	Fishes	Perch series	Eupercaria	1.09	0.00	0.00	0.00	0.00	0.00	1.09	0.00	0.00	0.00
Mixed aquatics	Fishes		Ovalentaria	0.93	0.00	0.00	0.00	0.00	0.93	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Otomorpha	0.71	0.00	0.00	0.00	0.71	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Otophysi	0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.00	0.00
Mixed aquatics	Flatworms		Microstomum	2.08	0.00	0.00	2.08	0.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Flatworms		Acotylea	0.45	0.00	0.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Flatworms		Dalyellioida	0.59	0.00	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Nematodes		Enoplea	6.98	0.00	0.00	0.00	0.00	0.00	0.00	6.98	0.00	0.00
Mixed aquatics	Copepods		Neocopepoda	0.31	0.00	0.00	0.31	0.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Copepods		Calanoida	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.00
Mixed aquatics	Copepods		Podoplea	1.05	0.00	0.00	0.00	0.00	1.05	0.00	0.00	0.00	0.00
Mixed aquatics	Copepods		Harpacticoida	4.00	0.00	0.00	0.00	4.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Gastrotrichs		Chaetonotidae	12.50	0.00	0.00	0.00	0.00	12.50	0.00	0.00	0.00	0.00
Mixed aquatics	Bivalves	Mussels	Mytilidae	0.56	0.00	0.00	0.00	0.00	0.56	0.00	0.00	0.00	0.00
Mixed aquatics	Cnidarians	Cnidarians	Cnidaria	2.33	0.00	0.00	0.00	0.00	1.17	1.17	0.00	0.00	0.00
Mixed aquatics	Hydrozoans	Hydrozoans	Hydrozoa	1.64	0.00	0.00	0.00	0.00	0.82	0.00	0.82	0.00	0.00
Mixed aquatics	Hydrozoans		Hydroidolina	0.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.87	0.00

Continued on next page

Table 4.28 continued

Ecological category	Group	Common name	Taxon	Total	485	460	415	384	355	341*	321	250	145
			Ingroup	461.84	5.77	37.72	45.92	12.52	27.63	18.79	44.35	269.13	0.00
Mixed aquatics	Hydrozoans		Filifera	1.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.75	0.00
Mixed aquatics	Hydrozoans	Thecate hydroids	Leptothecata	4.88	0.00	0.00	0.00	0.00	2.44	0.00	2.44	0.00	0.00
Mixed aquatics	Hydrozoans		Campanulariidae	4.88	0.00	0.00	0.00	0.00	2.44	0.00	2.44	0.00	0.00
Terrestrial	**Mammals**	**Guinea pig suborder**	**Hystricomorpha**	**0.25**	**0.00**	**0.00**	**0.25**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Terrestrial	Birds	Landfowl	Galliformes	3.07	1.54	0.00	0.77	0.00	0.00	0.00	0.00	0.77	0.00
Terrestrial	Birds	Pheasant subfamily	Phasianinae	0.80	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Mites		Analgoidea	6.67	0.00	6.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Spiders	Thin-legged wolf spiders	Pardosa	0.49	0.00	0.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Coleopterans		Elaphropus	1.74	0.00	0.00	1.74	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Coleopterans		Doryphorini	2.17	0.00	0.00	0.00	0.00	0.00	0.00	2.17	0.00	0.00
Terrestrial	Dipterans		repleta group	5.14	0.00	0.00	5.14	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Dipterans		melanogaster group	4.95	0.00	0.00	4.95	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Dipterans	Dagger flies	Empididae	3.24	0.00	3.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Hymenopterans		Chelostoma	1.77	0.00	0.00	1.77	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Animals	Animals	Metazoa	0.74	0.00	0.00	0.25	0.00	0.00	0.50	0.00	0.00	0.00
Mixed	Animals	Deuterostomes	Deuterostomia	0.77	0.00	0.00	0.58	0.00	0.00	0.19	0.00	0.00	0.00
Mixed	Animals	Chordates	Chordata	0.19	0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Vertebrates	Vertebrates	Vertebrata	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00
Mixed	Vertebrates	Jawed vertebrates	Gnathostomata	0.20	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Mammals	Mammals	Mammalia	0.62	0.31	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Vertebrates	Diapsids	Sauria	0.60	0.00	0.00	0.00	0.00	0.00	0.60	0.00	0.00	0.00
Mixed	Birds		Neognathae	1.48	0.00	0.00	0.00	0.00	0.00	1.48	0.00	0.00	0.00
Mixed	Birds	Perching birds	Passeriformes	0.76	0.00	0.00	0.00	0.76	0.00	0.00	0.00	0.00	0.00
Mixed	Testudines		Durocryptodira	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.00	0.00
Mixed	Squamates	Colubrid snake superfamily	Colubroidea	0.91	0.00	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Flatworms	Flatworms	Platyhelminthes	0.97	0.00	0.00	0.00	0.00	0.00	0.48	0.48	0.00	0.00
Mixed	Invertebrates		Protostomia	0.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.52	0.00
Mixed	Invertebrates		Ecdysozoa	1.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.50	0.00
Mixed	Nematodes		Tylenchoidea	38.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	38.00	0.00
Mixed	Nematodes		Pratylenchinae	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00
Mixed	**Nematodes**		**Hirschmanniella**	**100.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**100.00**	**0.00**
Mixed	Arthropods	Arthropods	Arthropoda	0.49	0.00	0.00	0.00	0.00	0.49	0.00	0.00	0.00	0.00
Mixed	Mites		Astigmata	6.67	0.00	0.00	6.67	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Arthropods		Pancrustacea	0.47	0.00	0.00	0.00	0.00	0.00	0.47	0.00	0.00	0.00
Mixed	Arthropods	Insects and relatives	Hexapoda	0.91	0.00	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Insects		Neoptera	0.96	0.00	0.00	0.00	0.00	0.00	0.96	0.00	0.00	0.00
Mixed	Insects	Metamorphosing insects	Holometabola	11.86	0.00	0.00	0.00	0.00	0.00	11.86	0.00	0.00	0.00
Mixed	Coleopterans	Beetles	Coleoptera	2.65	0.00	1.32	0.00	0.00	0.00	0.00	1.32	0.00	0.00
Mixed	Dipterans		Brachycera	3.20	0.00	0.00	3.20	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Dipterans		Muscomorpha	3.24	0.00	0.00	0.00	0.00	0.00	0.00	3.24	0.00	0.00
Mixed	Dipterans	Midge suborder	Nematocera	2.30	2.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Hymenopterans	Stinging wasps, ants and bees	Aculeata	4.66	0.00	0.00	2.33	2.33	0.00	0.00	0.00	0.00	0.00
Mixed	Invertebrates	Lophotrochozoans	Lophotrochozoa	1.86	0.00	0.62	0.62	0.00	0.00	0.00	0.62	0.00	0.00
Mixed	Molluscs	Molluscs	Mollusca	0.97	0.00	0.00	0.00	0.49	0.00	0.00	0.00	0.49	0.00
Mixed	Gastropods	Snails and slugs	Gastropoda	0.43	0.00	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.00

Embryophyta

Aquatics

Allowing for differences in read count, and excluding the first and final samples, ELF039 appears broadly homogeneous. Each of the main samples mostly comprises the saltwater *Zostera* and the mixed aquatic Alismatales. A diverse range of other aquatics occur at lower frequency, including *Ruppia*, *Typha*, and the reed group in nearly every sample. Many of the freshwater taxa are rarely seen and are probably a product of the high data yield in this core.

Halophytes and xerophytes

These groups appear in sample 355 but are mostly found in 321 and 250, which have the highest read counts. Halophytes and xerophytes typically grow in challenging conditions, so might be expected to have low biomass, which would explain their rarity. The lack of these taxa in samples with fewer reads could easily be due to lower sampling effort instead of genuine absence. In any case, their presence supports the saltwater signal towards the top of the core.

Trees/shrubs

All samples but the empty 145 have a clear but proportionally small woody signal. The most frequent taxa are the willow group, Betulaceae and child taxa, *Prunus*, *Crataegus* (hawthorn) and *Malus* (apple) with their parent taxon Maleae, and *Ulmus* (elms) and Ulmaceae. A wide variety of other woody taxa are found at lower frequency. These include *Tilia* and *Tiliodeae*, which when taken with the Malvales taxa listed under trees/shrubs and herbs, suggest lime trees throughout the core. Despite it being overshadowed by the saltwater aquatics, there is clearly a strong woodland influence in ELF039.

Trees/shrubs and herbs

Sample 485 does not fit the *Zostera*-dominated pattern through most of ELF039. It has only 18 reads, so its unusual composition could simply be due to very limited sampling of the environment, but most biogenomic mass is assigned to the dead-nettle family Lamiaceae and its parent order Lamiales, which are morphologically varied but terrestrial. There are no aquatics. 485 may represent a terrestrial environment before inundation.

Grasses and relatives

The main samples all contain at least Poales, usually accompanied by other higher grass taxa. As with previous samples, it may be possible to interpret many of the higher taxa reads as more reeds and rushes due to their prevalence. The Panicoideae group is also frequent; it may also actually have derived from reeds but could suggest a warm climate if accepted. The grass signal is approximately equal to that of woody taxa in most samples, although reduced in sample 341.

Metazoa

The Metazoa results are broadly consistent with Embryophyta. Aquatics are most frequent and are present in every sample with data, although saltwater aquatics are absent from the deepest sample, 485. There are enough reads that some taxa are present in multiple samples, including Fabriciidae (a tube worm family), Actiniaria (sea anemones), Percomorphaceae (a group of fishes), and *Mytilus* (a mussel genus). Terrestrial taxa are fewer, and many may be over-assigned, but they show some bias towards the three deepest samples. This would be consistent with inundation, although the aquatic signal is clear throughout.

4.14.3 Pianka scores

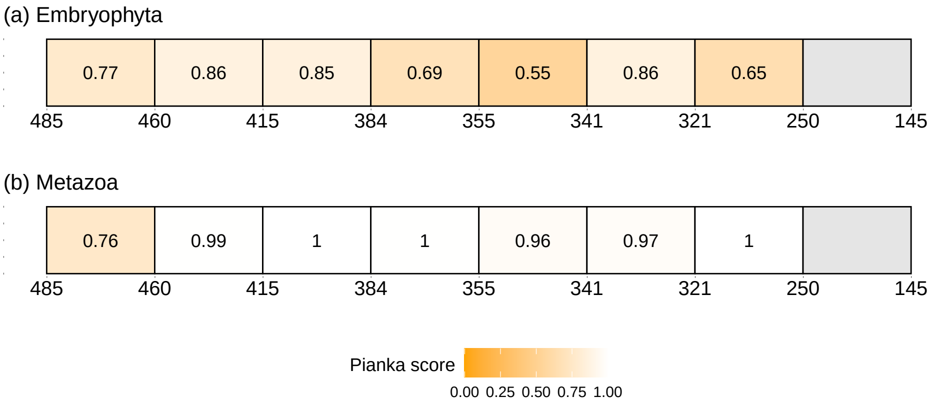


Figure 4.68: Pianka similarity scores between adjacent samples in ELF039 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

The Embryophyta Pianka scores closely reflect the taxonomic profiles: the deepest sample is distinct and the remainder highly similar. The Metazoa samples all appear moderately dissimilar to their neighbours with no apparent stability. The Metazoa read counts were relatively high for several samples, but 27 reads is still a very limited sample of the environment, so a lack of overlap is to be expected.

4.14.4 Summary

ELF039 is a mostly *Zostera*-dominated core. All but the deepest sample, and the curiously empty top sample, returned mostly *Zostera* and its parent taxa alongside smaller tree and grass signals. The deepest sample returned only 18 Embryophyta reads, all from terrestrial taxa. It may represent a terrestrial environment, without the additional benefits for DNA preservation from seawater, before marine inundation. The scant Metazoa data is consistent with this, showing a consistent aquatic signal but with terrestrial taxa clustered in deeper samples. However, the low data yields from Metazoa and from Embryophyta in 485 make this conclusion uncertain. The main message from ELF039 is of a coastal saltwater environment retaining some influence from floodplain woodland and reed beds.

4.15 ELF040A

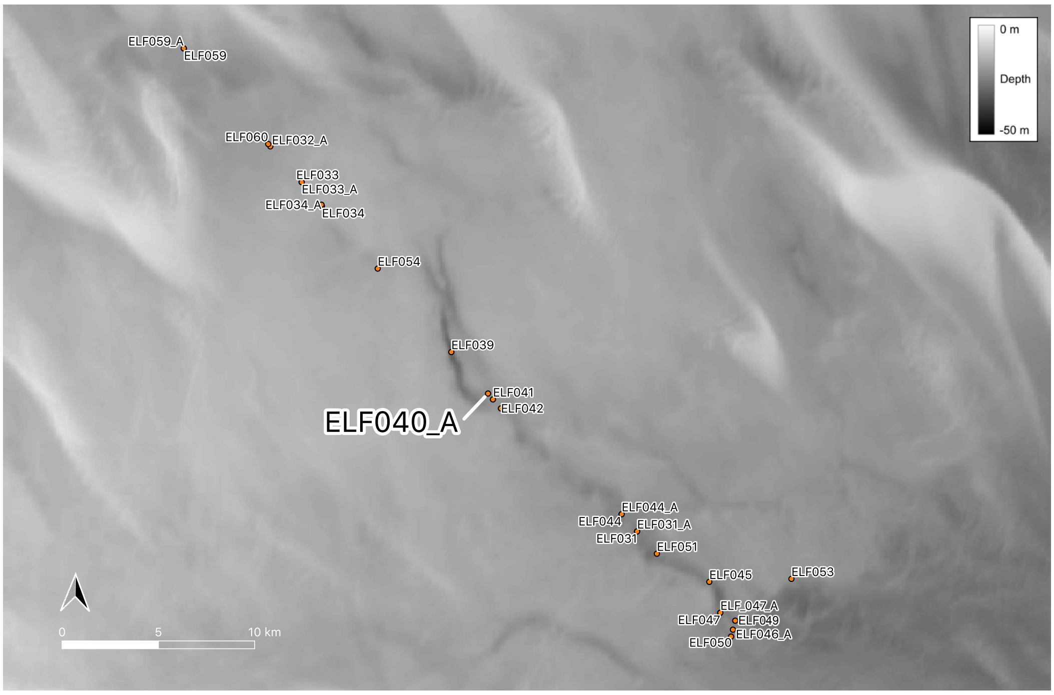


Figure 4.69: Bathymetry map of the palaeochannel transect highlighting ELF040A. Darker grey indicates greater depth.

ELF040A is the first in a cluster of three cores approximately 10 km from the start of the visible palaeochannel. The neighbouring cores are ELF041 and ELF042; there is no sedaDNA data from core ELF040.

4.15.1 Read counts

Data yield varies considerably with no clear pattern for Embryophyta; there is a decrease with depth for Metazoa, but with a peak of 12 reads, there is not much data to go on. Samples 487, 208, and 112 are similarly sparse for Embryophyta reads, but samples 350, 298, 192, and 95 have high counts. They should be enough for an impression of the core.

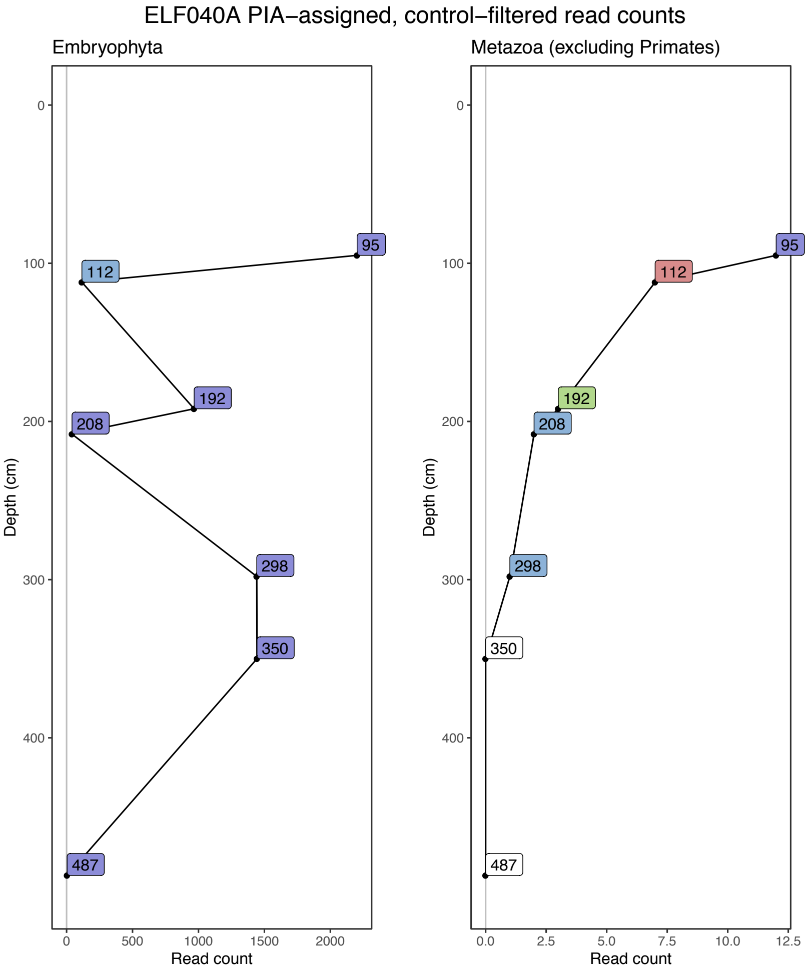
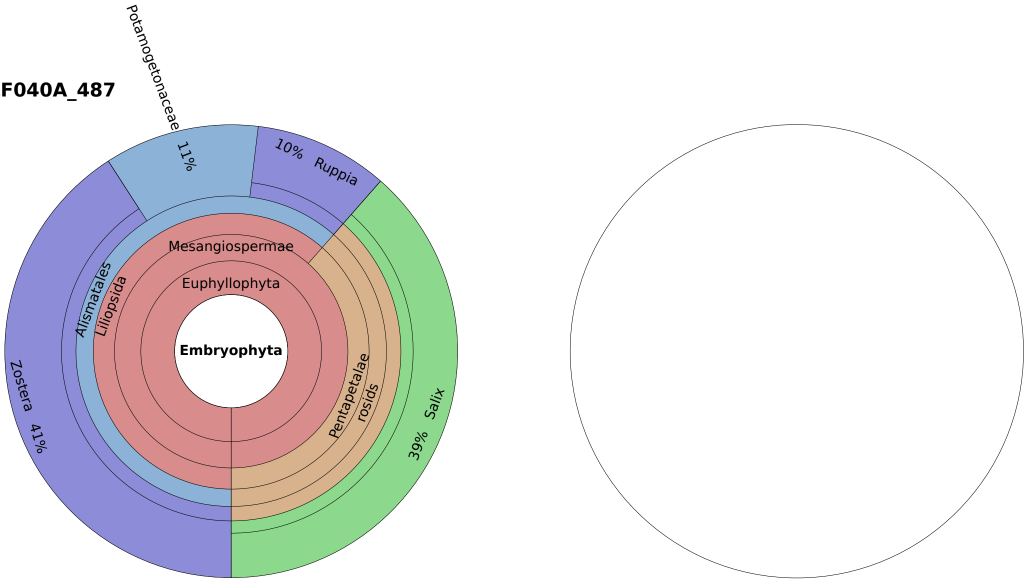


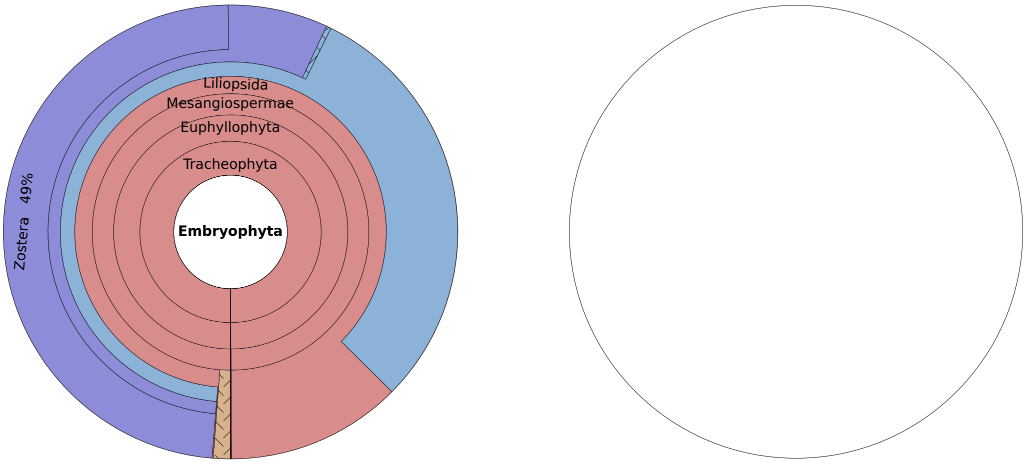
Figure 4.70: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF040A. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.15.2 Taxonomic profiles

ELF040A_487



ELF040A_350



ELF040A_298

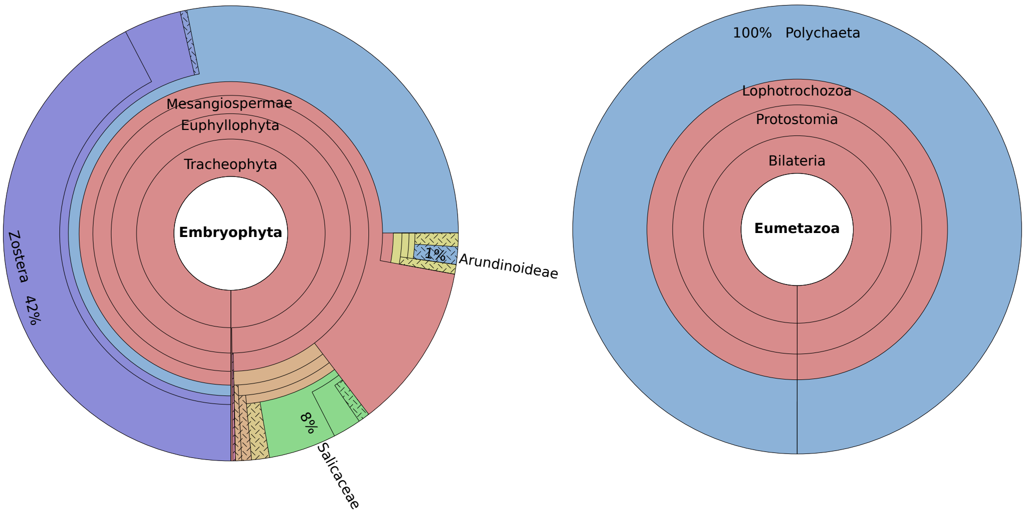


Figure 4.71: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF040A, samples 487, 350, and 298. Continued in figure 4.72. See figure 4.4 for colour key.

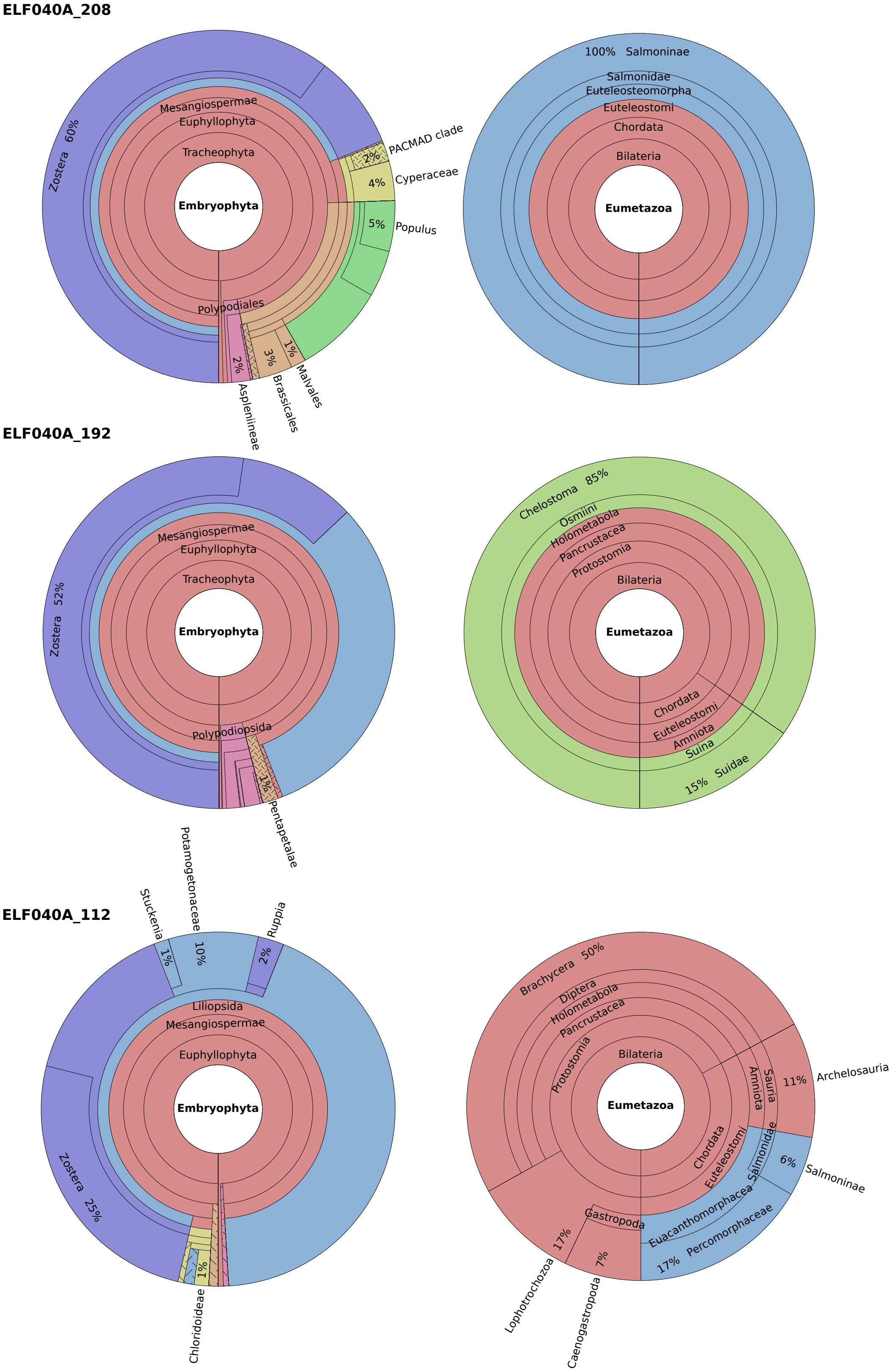


Figure 4.72: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF040A, samples 208, 192, and 112. Continued in figure 4.73. See figure 4.4 for colour key.

ELF040A_095

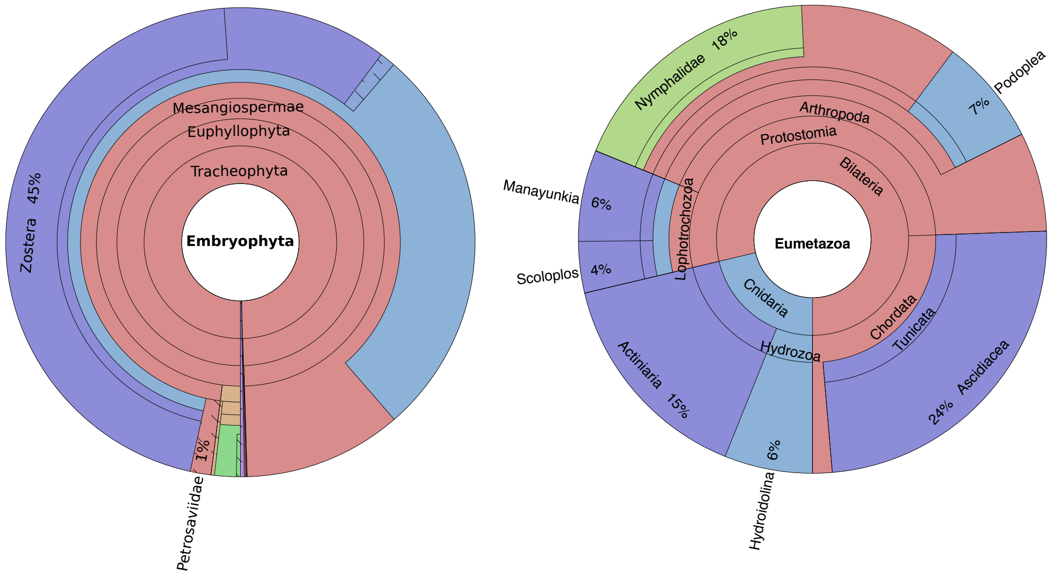


Figure 4.73: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF040A, sample 95. See figure 4.4 for colour key.

Table 4.29: Biogenomic masses of European taxa in Embryophyta from ELF040A. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	487	350	298	208	192	112	95
		Ingroup	2066.34	4.49	438.60	455.89	42.60	359.43	36.30	729.03
Freshwater aquatics	Arrowheads	Sagittaria	0.06	0.00	0.00	0.06	0.00	0.00	0.00	0.00
Freshwater aquatics	Pondweeds	Potamogeton	1.08	0.00	0.00	0.00	0.00	0.00	0.00	1.08
Freshwater aquatics	Bulrush family	Typhaceae	2.91	0.00	0.00	0.00	0.00	0.00	0.00	2.91
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	2.09	0.00	0.00	0.70	0.00	0.00	0.00	1.39
Salt/brackish aquatics	Tasselweeds	Ruppia	3.43	0.43	0.00	0.43	0.00	0.00	0.86	1.72
Salt/brackish aquatics	Eelgrass family	Zosteraceae	179.82	0.00	31.19	18.35	3.67	38.53	5.50	82.57
Salt/brackish aquatics	Eelgrasses	Zostera	957.80	1.83	212.84	190.83	25.69	187.16	9.17	330.28
Mixed aquatics	Alismatids	Alismatales	583.93	0.00	132.23	126.46	0.00	111.35	15.62	198.27
Mixed aquatics	Pondweed family	Potamogetonaceae	8.93	0.50	1.49	0.99	0.00	0.00	2.98	2.98
Mixed aquatics	Pondweeds	Stuckenia	0.99	0.00	0.00	0.00	0.00	0.00	0.50	0.50
Mixed aquatics	Reeds	Arundinoideae	1.06	0.00	0.00	0.71	0.00	0.00	0.00	0.35
Mixed aquatics	Common reed tribe	Molinieae	7.40	0.00	0.35	4.94	0.35	0.35	0.35	1.06
Halophytes	Glassworts	Salicornia	0.46	0.00	0.00	0.46	0.00	0.00	0.00	0.00
Halophytes	Leadwort family	Plumbaginaceae	0.41	0.00	0.00	0.21	0.00	0.21	0.00	0.00
Xerophytes	Goosefoot family	Chenopodiaceae	0.80	0.00	0.00	0.00	0.00	0.80	0.00	0.00
Trees/shrubs		Viburnum	0.26	0.00	0.00	0.26	0.00	0.00	0.00	0.00
Trees/shrubs	Birch family	Betulaceae	1.08	0.00	0.00	1.08	0.00	0.00	0.00	0.00
Trees/shrubs	Willow family	Salicaceae	37.91	0.00	1.81	21.66	3.61	0.00	0.00	10.83
Trees/shrubs	Willow tribe	Saliceae	12.71	0.00	1.82	9.08	1.82	0.00	0.00	0.00
Trees/shrubs	Poplars	Populus	5.93	0.00	0.00	1.98	1.98	0.00	0.00	1.98
Trees/shrubs	Willows	Salix	8.65	1.73	1.73	1.73	0.00	3.46	0.00	0.00
Trees/shrubs	Stone fruit trees	Prunus	3.65	0.00	0.00	3.65	0.00	0.00	0.00	0.00
Trees/shrubs	Apples	Malus	0.99	0.00	0.00	0.99	0.00	0.00	0.00	0.00
Trees/shrubs	**Cashew family**	**Anacardiaceae**	**1.83**	**0.00**	**0.00**	**1.83**	**0.00**	**0.00**	**0.00**	**0.00**
Herbs	Arum family	Araceae	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.18
Herbs		Dioscoreales	1.87	0.00	0.00	0.00	0.00	0.93	0.00	0.93
Herbs	Lily family	Liliaceae	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.03
Trees/shrubs and herbs		campanulids	0.31	0.00	0.00	0.00	0.00	0.00	0.31	0.00
Trees/shrubs and herbs	Carrot order	Apiales	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.47
Trees/shrubs and herbs	Daisy family	Asteraceae	0.81	0.00	0.00	0.54	0.27	0.00	0.00	0.00
Trees/shrubs and herbs	Aster tribe	Astereae	1.04	0.00	0.00	0.52	0.00	0.00	0.00	0.52
Trees/shrubs and herbs		Ericales	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.36
Trees/shrubs and herbs	Mustard order	Brassicales	2.60	0.00	0.00	1.30	1.30	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow order	Malvales	0.58	0.00	0.00	0.00	0.58	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	0.65	0.00	0.00	0.00	0.00	0.65	0.00	0.00
Trees/shrubs and herbs		Saxifragales	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.39
Trees/shrubs and herbs		Asparagales	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.08
Trees/shrubs and herbs		Liliales	0.07	0.00	0.00	0.00	0.07	0.00	0.00	0.00
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.16	0.00	0.00	0.00	0.00	0.16	0.00	0.00
Grasses and relatives	Grass order	Poales	1.46	0.00	0.00	0.00	0.00	0.29	0.00	1.17
Grasses and relatives	Sedges	Cyperaceae	4.63	0.00	0.00	3.08	1.54	0.00	0.00	0.00
Grasses and relatives	True sedges	Carex	2.28	0.00	0.00	0.00	0.00	0.00	0.00	2.28

Continued on next page

Table 4.29 continued

Ecological category	Common name	Taxon	Total	487	350	298	208	192	112	95
		Ingroup	2066.34	4.49	438.60	455.89	42.60	359.43	36.30	729.03
Grasses and relatives	Rushes	Juncaceae	1.04	0.00	0.00	0.00	0.00	0.00	0.00	1.04
Grasses and relatives		Poeae	0.18	0.00	0.00	0.00	0.00	0.00	0.18	0.00
Grasses and relatives		Chloridoideae	0.96	0.00	0.00	0.48	0.00	0.00	0.48	0.00
Grasses and relatives	Zoysia tribe	Zoysieae	1.52	0.00	0.00	1.52	0.00	0.00	0.00	0.00
Grasses and relatives	Dropseeds	Sporobolus	0.95	0.00	0.00	0.95	0.00	0.00	0.00	0.00
Grasses and relatives		**Panicoideae**	**0.54**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.54**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**1.20**	**0.00**	**0.00**	**0.80**	**0.40**	**0.00**	**0.00**	**0.00**
Grasses and relatives	**Panicgrass**	**Panicum**	**0.64**	**0.00**	**0.00**	**0.64**	**0.00**	**0.00**	**0.00**	**0.00**
Ferns	Ferns	Polypodiopsida	0.14	0.00	0.00	0.00	0.00	0.14	0.00	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	1.48	0.00	0.00	0.08	0.00	1.31	0.00	0.08
Ferns	Polypod ferns	Polypodiales	5.31	0.00	0.08	0.48	0.16	4.42	0.00	0.16
Ferns	Spleenwort suborder	Aspleniineae	0.99	0.00	0.00	0.18	0.72	0.00	0.09	0.00
Ferns		Athyriaceae	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.07
Ferns	Chain fern family	Blechnaceae	0.08	0.00	0.00	0.00	0.00	0.08	0.00	0.00
Ferns		Thelypteridaceae	0.20	0.00	0.00	0.00	0.10	0.10	0.00	0.00
Ferns		Thelypteridoideae	0.08	0.00	0.00	0.00	0.00	0.00	0.08	0.00
Ferns	Bracken	Pteridium aquilinum	0.13	0.00	0.00	0.00	0.00	0.13	0.00	0.00
Ferns		Polypodiineae	1.12	0.00	0.00	0.00	0.00	1.12	0.00	0.00
Ferns	Buckler and male ferns	Dryopteris	0.08	0.00	0.00	0.00	0.00	0.08	0.00	0.00
Ferns		Polypodiaceae	5.16	0.00	0.00	0.00	0.00	5.16	0.00	0.00
Ferns		Polypodioideae	0.24	0.00	0.00	0.00	0.00	0.24	0.00	0.00
Ferns	Polypodies	Polypodium	0.54	0.00	0.00	0.00	0.00	0.54	0.00	0.00
Bryophytes		Bryophytina	1.95	0.00	0.00	0.00	0.00	0.00	0.00	1.95
Bryophytes	Liverworts	Marchantiophyta	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.53
Mixed		Tracheophyta	1.73	0.00	0.17	0.69	0.17	0.17	0.00	0.52
Mixed		Euphyllophyta	1.20	0.00	0.00	0.00	0.17	0.86	0.17	0.00
Mixed	Monocots	Liliopsida	187.21	0.00	54.88	52.99	0.00	0.00	0.00	79.33

Table 4.30: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF040A. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	487	350	298	208	192	112	95
			Ingroup	27.12	0.00	0.00	0.74	0.79	2.40	8.70	14.49
Salt/brackish aquatics	Tunicates	Sea squirts	Ascidiacea	3.45	0.00	0.00	0.00	0.00	0.00	0.00	3.45
Salt/brackish aquatics	Annelids		Manayunkia	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.90
Salt/brackish aquatics	Annelids		Scoloplos	0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.51
Salt/brackish aquatics	Sea anemones	Sea anemones	Actiniaria	2.17	0.00	0.00	0.00	0.00	0.00	0.00	2.17
Mixed aquatics	Fishes		Percomorphaceae	1.06	0.00	0.00	0.00	0.00	0.00	1.06	0.00
Mixed aquatics	Fishes	Salmon subfamily	Salmoninae	0.71	0.00	0.00	0.00	0.35	0.00	0.35	0.00
Mixed aquatics	Copepods		Podoplea	1.05	0.00	0.00	0.00	0.00	0.00	0.00	1.05
Mixed aquatics	Annelids	Bristle worms	Polychaeta	0.74	0.00	0.00	0.74	0.00	0.00	0.00	0.00
Mixed aquatics	Hydrozoans		Hydroidolina	0.87	0.00	0.00	0.00	0.00	0.00	0.00	0.87
Terrestrial	Mammals	Pig family	Suidae	0.32	0.00	0.00	0.00	0.00	0.32	0.00	0.00
Terrestrial	Butterflies	Brushfoots	Nymphalidae	2.58	0.00	0.00	0.00	0.00	0.00	0.00	2.58
Terrestrial	Hymenopterans		Chelostoma	1.77	0.00	0.00	0.00	0.00	1.77	0.00	0.00
Mixed	Animals	Animals	Metazoa	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.25
Mixed	Animals	Chordates	Chordata	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.19
Mixed	Vertebrates		Archelosauria	0.68	0.00	0.00	0.00	0.00	0.00	0.68	0.00
Mixed	Arthropods	Arthropods	Arthropoda	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.98
Mixed	Lepidopterans		Ditrysia	1.56	0.00	0.00	0.00	0.00	0.00	0.00	1.56
Mixed	Dipterans		Brachycera	3.20	0.00	0.00	0.00	0.00	0.00	3.20	0.00
Mixed	Invertebrates	Lophotrochozoans	Lophotrochozoa	0.62	0.00	0.00	0.00	0.00	0.00	0.62	0.00
Mixed	Gastropods		Caenogastropoda	0.46	0.00	0.00	0.00	0.00	0.00	0.46	0.00

Embryophyta

Aquatics

ELF040A is a mostly aquatic core, and mostly *Zostera* at that. All but the deepest sample (which has only four reads) also contain the reed group, although often at lower frequency than in previous *Zostera*-dominated cores. Specifically freshwater taxa are nearly absent.

Terrestrial

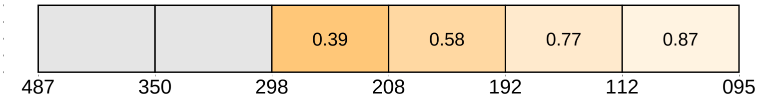
A small terrestrial signal exists in all samples, although not as diverse as might be expected for the amount of data in the core. Only the willow group are noticeably consistent. However, sample 192 is notable for its unusually diverse collection of ferns. Though their proportional biogenomic mass is low, the ferns in 192 include eight more taxa than other samples in ELF040A, and are frequent enough to make a rare appearance on the Krona chart (figure 4.72). The general scarcity of fern reads suggests an unusual situation in 192, although as ferns live in such varied habitats, it is difficult to suggest what.

Metazoa

Metazoa reads in ELF040A are too few to infer any ecological change. There is a mixture of aquatic and terrestrial taxa throughout the core, supporting the mixed signal from Embryophyta. Similarly, the top sample, 95, has the most informative profile and suggests a mostly saltwater environment.

4.15.3 Pianka scores

(a) Embryophyta



(b) Metazoa

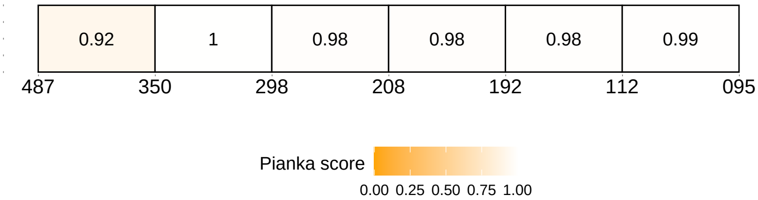


Figure 4.74: Pianka similarity scores between adjacent samples in ELF040A for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

The very high Embryophyta Pianka scores reflect the very similar taxonomic profiles, despite the likely influence of variable read counts. The unexpected fern signal in 192 was proportionally small and appears to have gone unnoticed. Unfortunately, there is not enough data to say the same for Metazoa.

4.15.4 Summary

ELF040A is another predominantly saltwater core. Like in many samples seen previously, the most frequent taxon is *Zostera*, although the reed and bulrush signal typically seen with *Zostera* may be reduced, perhaps suggesting a reduction in marsh or reed beds or a greater distance from shore. The small but consistent terrestrial signal is more typical, with the exception of an unusually large diversity of ferns in sample 192. The Metazoa data is sparse but consistent with a coastal environment.

4.16 ELF041

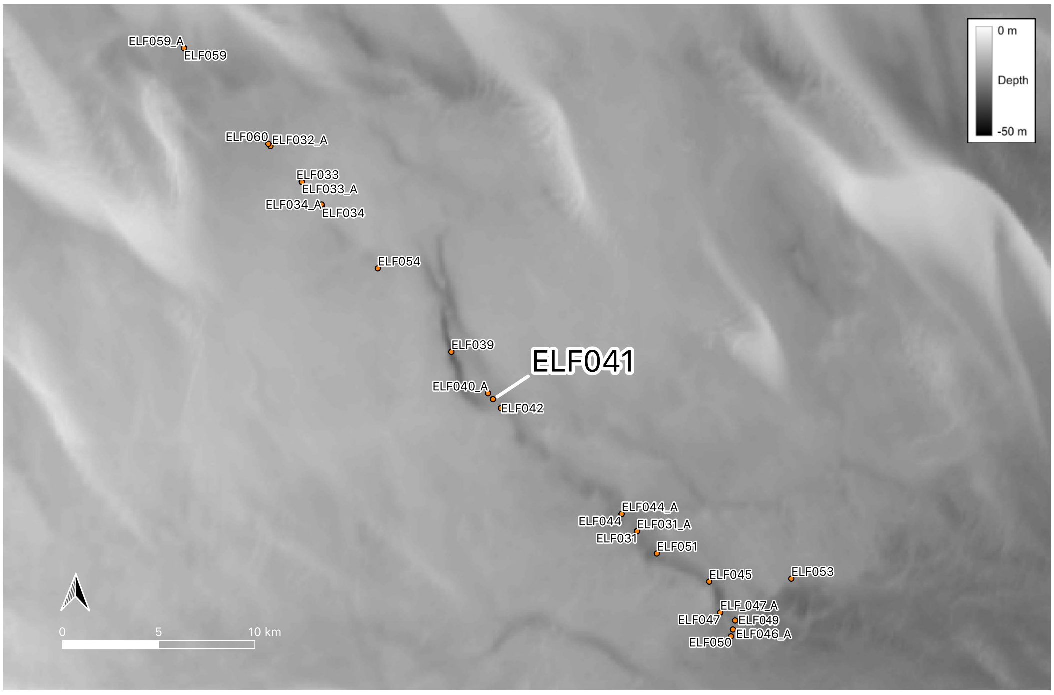


Figure 4.75: Bathymetry map of the palaeochannel transect highlighting ELF01. Darker grey indicates greater depth.

ELF041 is the second in a small cluster of cores approximately a third the way down the visible palaeochannel. It is close to ELF040A and ELF042.

4.16.1 Read counts

This is generally a low-yield core with the exception of sample 180, which returned around 1500 Embryophyta reads.

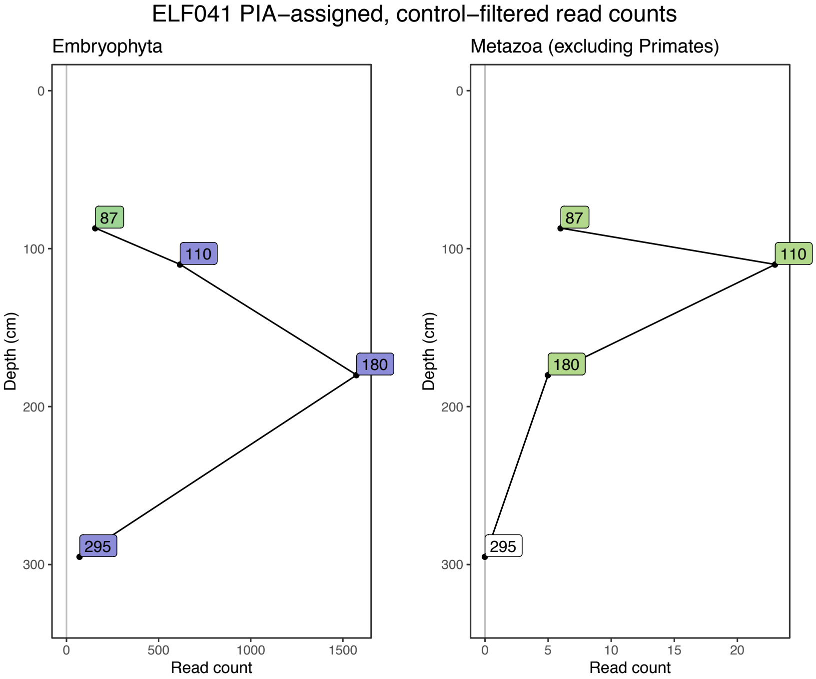
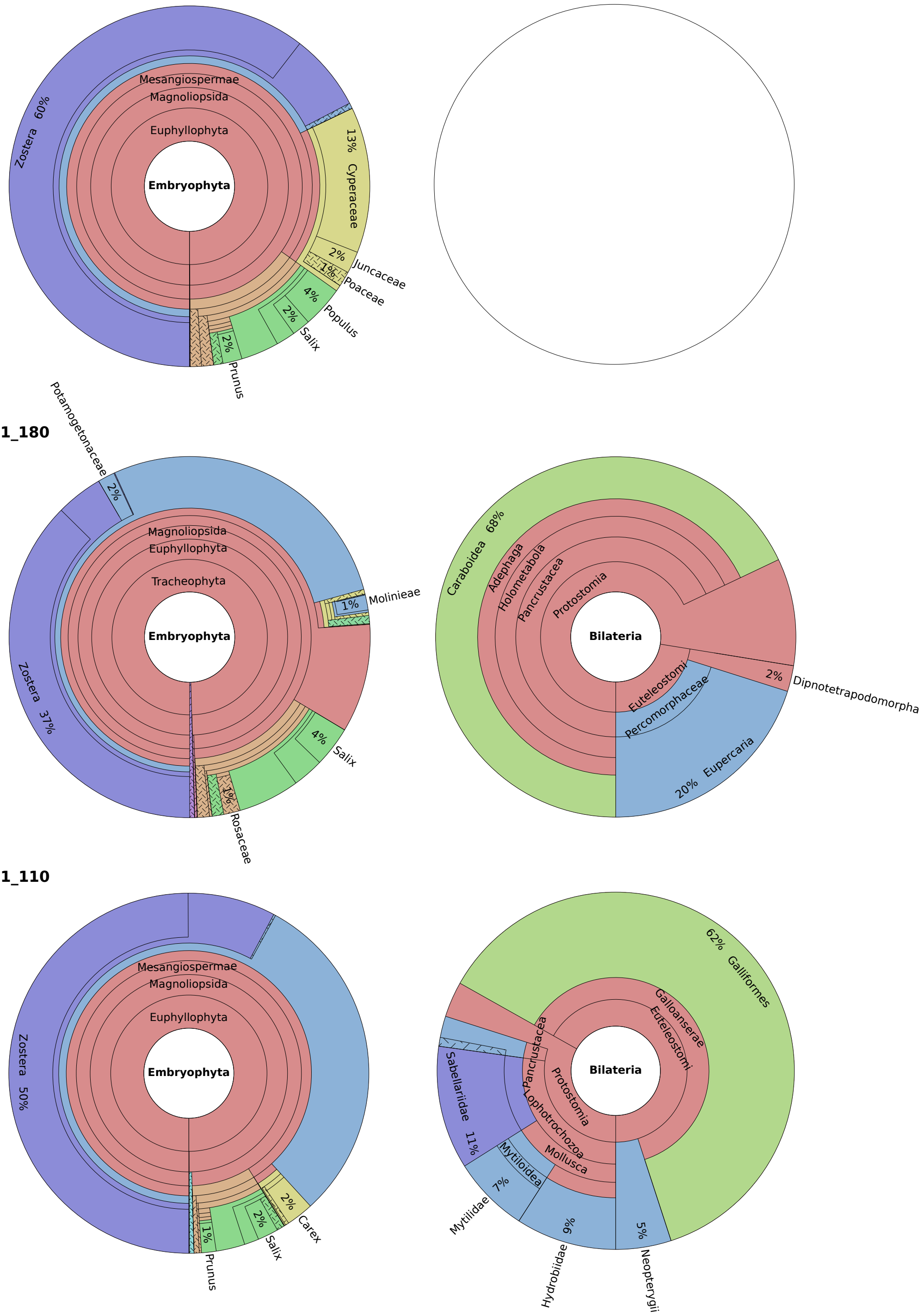


Figure 4.76: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF041. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.16.2 Taxonomic profiles



ELF041_295

ELF041_180

ELF041_110

Figure 4.77: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF040A, samples 295, 180, and 110. Continued in figure 4.72. See figure 4.4 for colour key.

ELF041_087

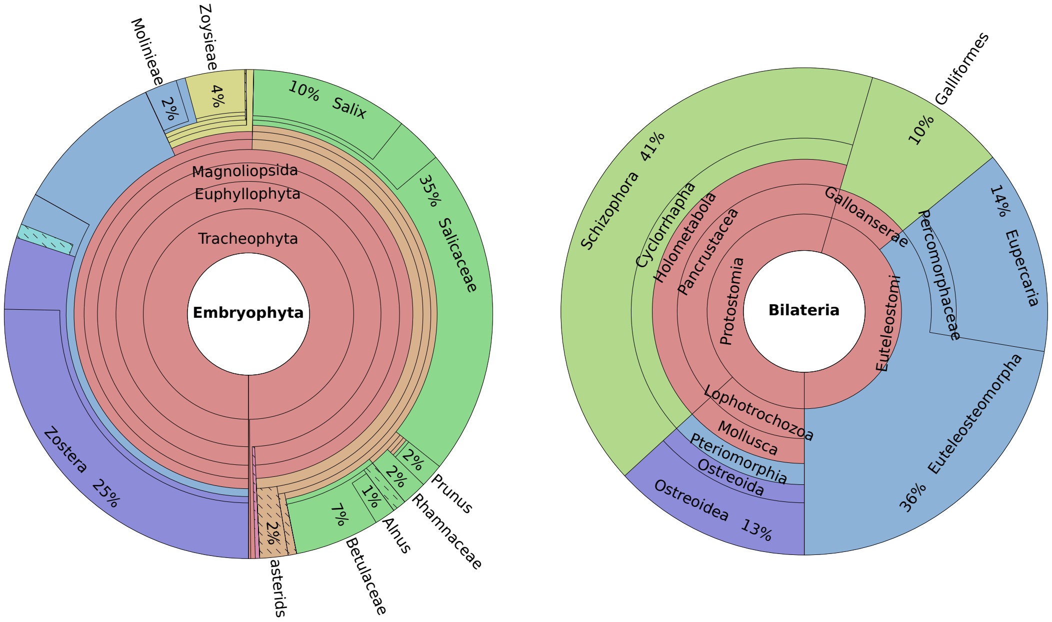


Figure 4.78: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF041, sample 87See figure 4.4 for colour key.

Table 4.31: Biogenomic masses of European taxa in Embryophyta from ELF041. Samples are in cm.

Ecological category	Common name	Taxon	Total	295	180	110	87
		Ingroup	1037.64	110.45	529.18	281.07	116.93
Freshwater aquatics	Pondweeds	Potamogeton	1.62	0.54	0.00	0.00	1.08
Freshwater aquatics	Bulrushes	Typha	3.77	0.00	3.77	0.00	0.00
Freshwater aquatics	Waterlily order	Nymphaeales	1.13	0.00	0.00	1.13	0.00
Salt/brackish aquatics	Tasselweeds	Ruppia	0.43	0.00	0.43	0.00	0.00
Salt/brackish aquatics	Eelgrass family	Zosteraceae	56.88	7.34	22.02	22.02	5.50
Salt/brackish aquatics	Eelgrasses	Zostera	429.36	64.22	196.33	139.45	29.36
Mixed aquatics	Alismatids	Alismatales	240.53	0.00	144.62	84.36	11.54
Mixed aquatics	Pondweed family	Potamogetonaceae	10.91	0.00	7.93	0.50	2.48
Mixed aquatics	Reeds	Arundinoideae	2.12	0.35	1.06	0.00	0.71
Mixed aquatics	Common reed tribe	Molinieae	9.87	0.00	7.05	0.35	2.47
Mixed aquatics	Common reeds	Phragmites	0.84	0.00	0.42	0.42	0.00
Halophytes	Leadwort family	Plumbaginaceae	0.21	0.00	0.21	0.00	0.00
Trees/shrubs	Dogwoods	Cornus	0.59	0.00	0.59	0.00	0.00
Trees/shrubs	Olive tribe	Oleeae	0.80	0.00	0.00	0.00	0.80
Trees/shrubs	Birch family	Betulaceae	6.48	0.00	0.00	0.00	6.48
Trees/shrubs	Alders	Alnus	4.84	0.00	3.23	0.00	1.61
Trees/shrubs	Hazels	Corylus	2.27	0.00	2.27	0.00	0.00
Trees/shrubs	Oaks	Quercus	1.06	0.00	0.00	0.00	1.06
Trees/shrubs	Willow family	Salicaceae	64.98	3.61	28.88	7.22	25.27
Trees/shrubs	Willow tribe	Saliceae	23.60	1.82	14.52	3.63	3.63
Trees/shrubs	Poplars	Populus	5.93	3.96	0.00	1.98	0.00
Trees/shrubs	Willows	Salix	38.06	1.73	19.03	5.19	12.11
Trees/shrubs	Buckthorn family	Rhamnaceae	1.92	0.00	0.00	0.00	1.92
Trees/shrubs	Stone fruit trees	Prunus	12.78	1.83	5.48	3.65	1.83
Trees/shrubs	Apple tribe	Maleae	1.78	0.89	0.89	0.00	0.00
Herbs		apioid superclade	0.35	0.00	0.00	0.00	0.35
Herbs	Chicory tribe	Cichorieae	0.28	0.28	0.00	0.00	0.00
Herbs	Mint tribe	Mentheae	2.46	0.00	2.46	0.00	0.00
Trees/shrubs and herbs		campanulids	0.31	0.00	0.31	0.00	0.00
Trees/shrubs and herbs	Carrot order	Apiales	0.47	0.00	0.47	0.00	0.00
Trees/shrubs and herbs		Apiineae	0.47	0.00	0.00	0.00	0.47
Trees/shrubs and herbs	Daisy family	Asteraceae	1.09	0.27	0.00	0.27	0.54
Trees/shrubs and herbs	Aster tribe	Astereae	1.55	0.52	0.00	1.04	0.00
Trees/shrubs and herbs		lamiids	0.68	0.00	0.68	0.00	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	0.88	0.00	0.88	0.00	0.00
Trees/shrubs and herbs	Rose subfamily	Rosoideae	1.39	0.00	1.39	0.00	0.00
Trees/shrubs and herbs	Mallow order	Malvales	0.58	0.58	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	1.31	0.00	0.65	0.00	0.65
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	1.67	0.56	0.56	0.56	0.00
Trees/shrubs and herbs		Asparagales	0.08	0.00	0.08	0.00	0.00
Trees/shrubs and herbs		Liliales	0.10	0.00	0.00	0.10	0.00
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.16	0.00	0.16	0.00	0.00
Grasses and relatives	Grass order	Poales	1.76	0.59	0.29	0.29	0.59
Grasses and relatives	Sedges	Cyperaceae	13.88	13.88	0.00	0.00	0.00
Grasses and relatives	True sedges	Carex	9.11	0.00	2.28	6.83	0.00
Grasses and relatives	Rushes	Juncaceae	2.09	2.09	0.00	0.00	0.00
Grasses and relatives		Coleanthinae	0.42	0.42	0.00	0.00	0.00
Grasses and relatives		Triticodae	0.22	0.00	0.00	0.11	0.11
Grasses and relatives	Zoysia tribe	Zoysieae	6.06	0.00	1.52	0.00	4.55
Grasses and relatives		Paniceae	0.67	0.67	0.00	0.00	0.00
Ferns	Ferns	Polypodiopsida	0.07	0.07	0.00	0.00	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	0.08	0.00	0.00	0.00	0.08
Ferns	Polypod ferns	Polypodiales	0.64	0.00	0.40	0.16	0.08
Ferns	Spleenwort suborder	Aspleniineae	0.63	0.00	0.54	0.00	0.09
Ferns		Phegopteridoideae	0.14	0.00	0.14	0.00	0.00
Ferns		Thelypteridoideae	0.08	0.00	0.00	0.00	0.08
Bryophytes	Feather mosses	Hypnales	2.24	0.00	2.24	0.00	0.00
Mixed		Tracheophyta	0.35	0.00	0.17	0.00	0.17
Mixed		Euphyllophyta	0.34	0.00	0.00	0.00	0.34
Mixed	Monocots	Liliopsida	51.00	0.00	51.00	0.00	0.00

Table 4.32: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF041. Samples are in cm.

Ecological category	Group	Common name	Taxon	Total	295	180	110	87
			Ingroup	29.81	0.00	5.40	16.38	8.02
Salt/brackish aquatics	Copepods		Calanidae	0.13	0.00	0.00	0.13	0.00
Salt/brackish aquatics	Annelids	Honeycomb worm family	Sabellariidae	1.80	0.00	0.00	1.80	0.00
Salt/brackish aquatics	Bivalves	Oyster superfamily	Ostreoidea	1.06	0.00	0.00	0.00	1.06
Mixed aquatics	Fishes		Neopterygii	0.81	0.00	0.00	0.81	0.00
Mixed aquatics	Fishes		Euteleosteomorpha	1.79	0.00	0.00	0.00	1.79
Mixed aquatics	Fishes	Perch series	Eupercaria	2.17	0.00	1.09	0.00	1.09
Mixed aquatics	Copepods		Neocopepoda	0.31	0.00	0.00	0.31	0.00
Mixed aquatics	Bivalves	Mussels	Mytilidae	1.12	0.00	0.00	1.12	0.00
Mixed aquatics	Gastropods	Mud snails	Hydrobiidae	1.47	0.00	0.00	1.47	0.00
Terrestrial	Birds	Landfowl	Galliformes	10.75	0.00	0.00	9.98	0.77
Terrestrial	Coleopterans	Ground beetle superfamily	Caraboidea	3.67	0.00	3.67	0.00	0.00
Terrestrial	Dipterans		Schizophora	3.31	0.00	0.00	0.00	3.31
Mixed	Animals		Eumetazoa	0.25	0.00	0.00	0.25	0.00
Mixed	Vertebrates	Tetropods and lungfish	Dipnotetrapodomorpha	0.13	0.00	0.13	0.00	0.00
Mixed	Invertebrates		Protostomia	1.04	0.00	0.52	0.52	0.00

Embryophyta

The four samples are relatively similar. The most frequent taxon in each sample is *Zostera*, although it is less dominant than in many other cores. Other aquatics are limited to the ambiguous Alismatales and Potamogetonaceae (pondweed family) and a small signal from the reed group. Woody taxa are found throughout the core, mostly consisting of the willow group and *Prunus*. Malvales, Malvaceae, and Malvoideae are also across all samples, possibly suggesting a *Tilia* signal, although no reads were assigned directly to *Tilia*. Woody taxa nearly match *Zostera* in the upper three samples.

Metazoa

Metazoa data is absent from the deepest sample. The remainder all show a fairly balanced split between aquatic and terrestrial, consistent with the mixed signal from Embryophyta.

4.16.3 Pianka scores

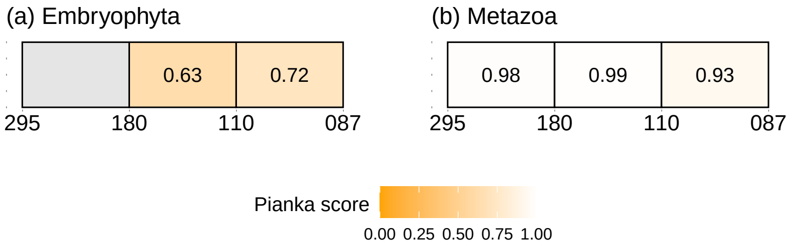


Figure 4.79: Pianka similarity scores between adjacent samples in ELF041 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

Pianka scores are usually unrealistically small in samples with low read counts, but the Embryophyta Pianka scores are very high, so the similarity is probably real. The Metazoa Pianka scores are less useful.

4.16.4 Summary

Unlike the more conventionally *Zostera*-dominated ELF040A, Embryophyta and Metazoa in ELF041 are both quite evenly divided between aquatic and terrestrial. It is unusual to have *Zostera* in moderation.

4.17 ELF042

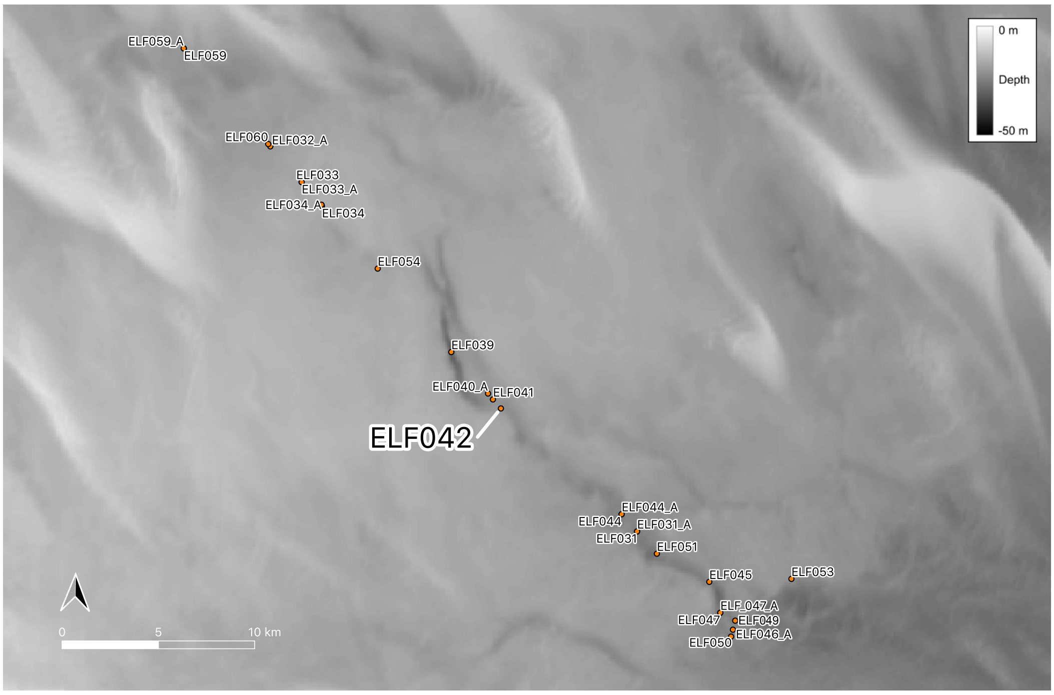


Figure 4.80: Bathymetry map of the palaeochannel transect highlighting ELF042. Darker grey indicates greater depth.

ELF042 is third in the small cluster of samples also containing ELF040A and ELF041.

4.17.1 Read counts

Embryophyta read counts are all high: around 1-2,000 in the deeper three samples and nearly 9,000 for sample 65. Metazoa counts have a similar uptick for 65 but peak at only 25 reads.

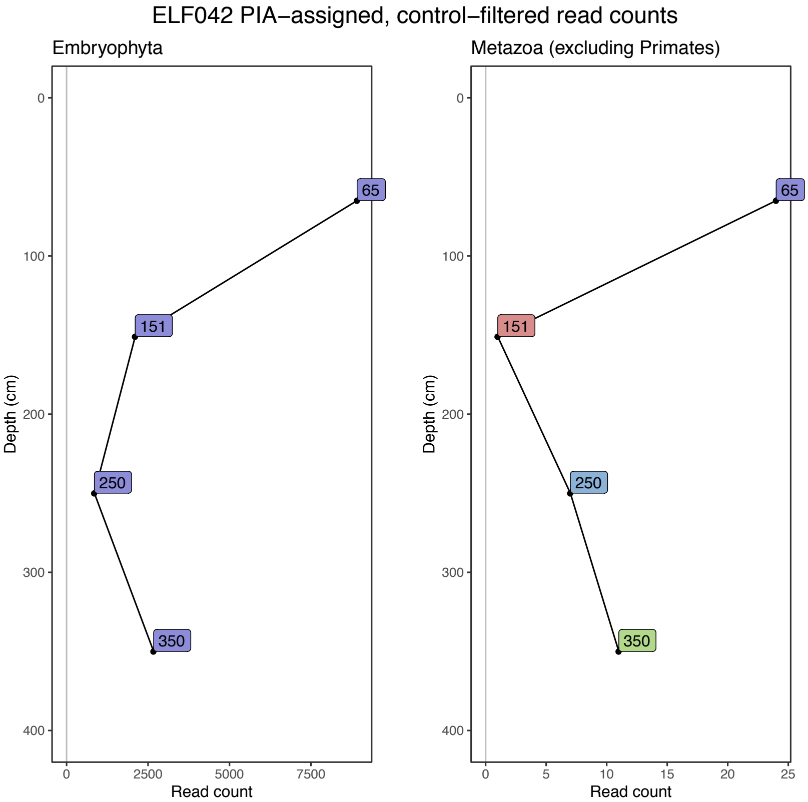
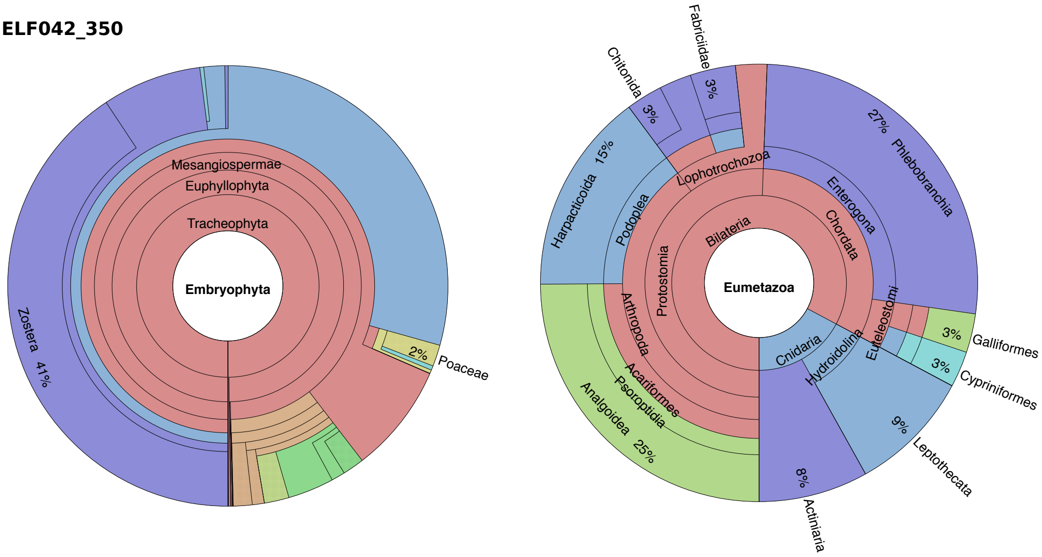


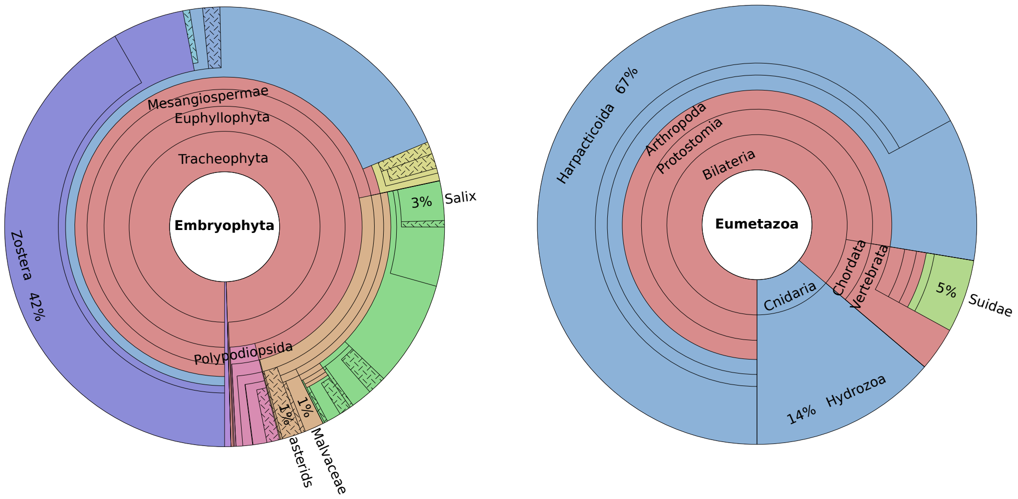
Figure 4.81: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF042. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.17.2 Taxonomic profiles

ELF042_350



ELF042_250



ELF042_151

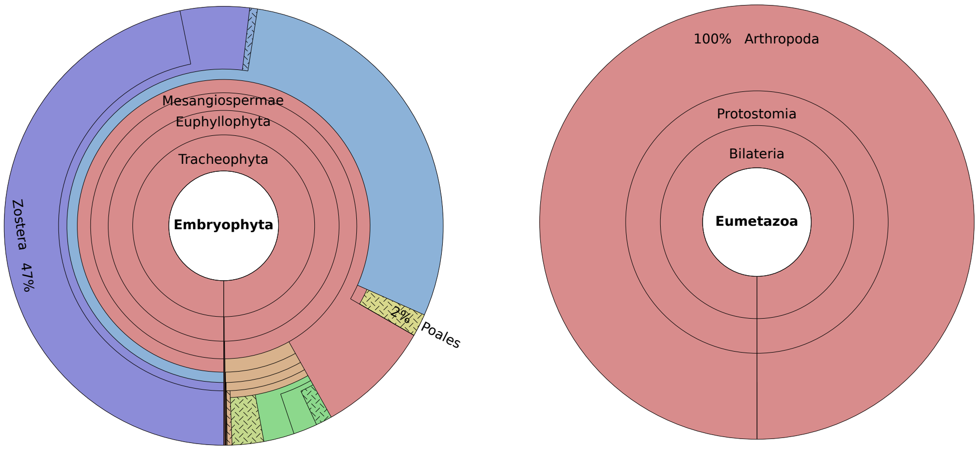


Figure 4.82: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF042, samples 350, 350, and 151. Continued in figure 4.83. See figure 4.4 for colour key.

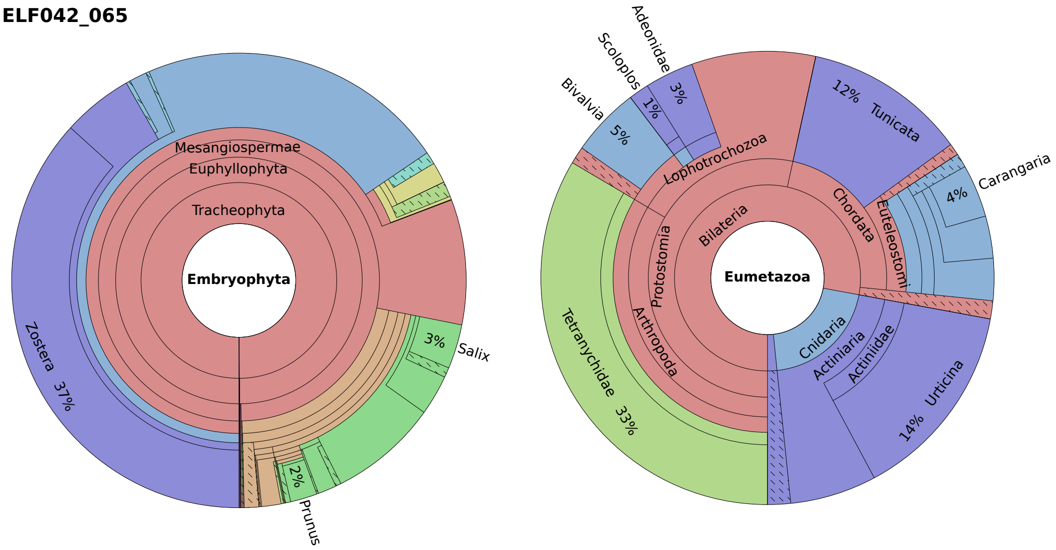


Figure 4.83: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF042, sample 65. See figure 4.4 for colour key.

Table 4.33: Biogenomic masses of European taxa in Embryophyta from ELF042. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	350	250	151	65
		Ingroup	5420.70	909.33	418.24	723.04	3370.09
Freshwater aquatics	Coontails	Ceratophyllum	1.45	0.00	0.00	0.00	1.45
Freshwater aquatics		Menyanthaceae	1.43	1.43	0.00	0.00	0.00
Freshwater aquatics	Water-milfoil family	Haloragaceae	6.92	2.31	0.00	0.00	4.62
Freshwater aquatics	Water-milfoils	Myriophyllum	2.31	0.00	0.00	0.00	2.31
Freshwater aquatics	Water-plantain family	Alismataceae	0.07	0.00	0.00	0.00	0.07
Freshwater aquatics	Duckweed family	Lemnoideae	2.44	0.00	0.00	0.00	2.44
Freshwater aquatics	Elodea family	Hydrocharitaceae	0.20	0.00	0.00	0.20	0.00
Freshwater aquatics	Pondweeds	Potamogeton	8.09	1.62	1.08	1.08	4.31
Freshwater aquatics	Bulrush family	Typhaceae	14.56	2.91	0.00	5.83	5.83
Freshwater aquatics	Bulrushes	Typha	3.77	0.00	0.00	0.00	3.77
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	4.88	1.39	1.39	0.00	2.09
Salt/brackish aquatics	Tasselweeds	Ruppia	7.73	0.86	3.00	0.86	3.00
Salt/brackish aquatics	Eelgrass family	Zosteraceae	293.58	66.06	22.02	36.70	168.81
Salt/brackish aquatics	Eelgrasses	Zostera	2113.76	368.81	174.31	337.61	1233.03
Mixed aquatics	Alismatids	Alismatales	1294.49	265.82	79.61	210.49	738.57
Mixed aquatics	Pondweed family	Potamogetonaceae	60.50	13.88	3.97	0.99	41.65
Mixed aquatics	Pondweeds	Stuckenia	8.93	0.99	0.99	0.99	5.95
Mixed aquatics	Club-rushes	Bolboschoenus	1.79	0.00	0.00	0.00	1.79
Mixed aquatics	Reeds	Arundinoideae	2.83	0.35	0.00	0.35	2.12
Mixed aquatics	Common reed tribe	Molinieae	29.61	7.76	1.76	1.41	18.68
Mixed aquatics	Common reeds	Phragmites	1.26	0.42	0.00	0.00	0.84
Halophytes	Beets	Beta	0.78	0.78	0.00	0.00	0.00
Halophytes		Salicornia subg. Salicornia	0.72	0.00	0.00	0.00	0.72
Halophytes	Sea-blites	Suaeda	0.94	0.94	0.00	0.00	0.00
Halophytes	Leadwort family	Plumbaginaceae	2.07	0.00	0.00	0.00	2.07
Halophytes	Sea-lavenders	Limonium	0.85	0.21	0.00	0.00	0.64
Xerophytes	Mugworts	Artemisia	0.19	0.00	0.00	0.00	0.19
Xerophytes	Goosefoot family	Chenopodiaceae	2.39	0.00	0.00	0.00	2.39
Trees/shrubs		Viburnum	0.26	0.26	0.00	0.00	0.00
Trees/shrubs	Honeysuckles	Lonicera	3.42	0.85	2.56	0.00	0.00
Trees/shrubs		Cornales	0.61	0.00	0.00	0.00	0.61
Trees/shrubs	Dogwoods	Cornus	0.59	0.00	0.00	0.00	0.59
Trees/shrubs	Privet	Ligustrum	0.68	0.00	0.00	0.00	0.68
Trees/shrubs	Birch family	Betulaceae	65.87	7.56	6.48	7.56	44.27
Trees/shrubs	Alders	Alnus	12.90	1.61	4.84	0.00	6.45
Trees/shrubs	Birches	Betula	0.97	0.00	0.97	0.00	0.00
Trees/shrubs	Hazels	Corylus	6.82	0.00	0.00	0.00	6.82
Trees/shrubs	Oaks	Quercus	6.35	0.00	1.06	1.06	4.23
Trees/shrubs	Willow family	Salicaceae	339.35	30.69	32.49	16.25	259.93
Trees/shrubs	Willow tribe	Saliceae	136.16	9.08	18.15	12.71	96.22
Trees/shrubs	Poplars	Populus	29.67	3.96	1.98	0.00	23.74
Trees/shrubs	Willows	Salix	136.68	10.38	12.11	8.65	105.54
Trees/shrubs		Amygdaleae	4.42	0.00	4.42	0.00	0.00
Trees/shrubs	Stone fruit trees	Prunus	78.53	7.30	3.65	3.65	63.92
Trees/shrubs	Apple tribe	Maleae	11.59	0.00	0.00	0.00	11.59
Trees/shrubs	Apples	Malus	2.97	0.00	0.99	0.99	0.99
Trees/shrubs	Blackberry and relatives	Rubus	2.22	0.00	0.00	0.00	2.22
Trees/shrubs	Elm family	Ulmaceae	6.02	0.00	0.50	3.51	2.01
Trees/shrubs	Elms	Ulmus	0.43	0.00	0.00	0.00	0.43
Trees/shrubs	Limes	Tilia	4.80	0.80	0.00	0.00	4.00
Trees/shrubs	Barberry subfamily	Berberidoideae	0.25	0.00	0.00	0.00	0.25
Herbs	Thistle tribe	Cardueae	0.48	0.48	0.00	0.00	0.00
Herbs		Carlininae	0.26	0.00	0.00	0.00	0.26
Herbs	Primrose family	Primulaceae	0.70	0.35	0.35	0.00	0.00
Herbs	Catmint subfamily	Nepetoideae	2.26	0.00	1.13	0.00	1.13
Herbs	Mint tribe	Mentheae	2.46	1.23	1.23	0.00	0.00
Herbs	Mints	Mentha	3.05	0.00	0.00	0.00	3.05
Herbs	Campions and catchflies	Silene	0.45	0.00	0.00	0.00	0.45
Herbs	Knotweed subfamily	Polygonoideae	0.56	0.00	0.00	0.00	0.56
Herbs	Rhubarb family	Rumiceae	0.41	0.00	0.00	0.00	0.41
Herbs		Camelineae	3.14	0.00	0.00	0.00	3.14
Herbs	Tofieldia family	Tofieldiaceae	0.87	0.00	0.87	0.00	0.00
Herbs	Asparagus family	Asparagaceae	0.21	0.00	0.00	0.10	0.10
Herbs		**Agavoideae**	**0.15**	**0.00**	**0.00**	**0.00**	**0.15**
Herbs	Iris family	Iridaceae	0.18	0.00	0.00	0.00	0.18
Herbs		Malaxidinae	0.53	0.00	0.00	0.00	0.53
Herbs	Buttercups	Ranunculus	0.15	0.00	0.00	0.00	0.15
Herbs	Meadow-rues	Thalictrum	0.92	0.00	0.00	0.00	0.92
Trees/shrubs and herbs		Gunneridae	1.41	0.47	0.00	0.47	0.47
Trees/shrubs and herbs		campanulids	2.77	0.62	0.00	0.00	2.16
Trees/shrubs and herbs	Carrot order	Apiales	0.47	0.47	0.00	0.00	0.00
Trees/shrubs and herbs	Ivy family	Araliaceae	1.01	0.00	0.50	0.00	0.50
Trees/shrubs and herbs		Asterales	0.86	0.00	0.00	0.00	0.86
Trees/shrubs and herbs	Daisy family	Asteraceae	2.44	0.00	0.27	0.00	2.17
Trees/shrubs and herbs		Asteroideae	0.75	0.00	0.00	0.00	0.75
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	0.37	0.00	0.00	0.18	0.18
Trees/shrubs and herbs	Aster tribe	Astereae	3.63	0.00	0.00	0.00	3.63
Trees/shrubs and herbs		lamiids	0.68	0.00	0.00	0.00	0.68
Trees/shrubs and herbs	Dead-nettle order	Lamiales	3.51	1.76	0.00	0.00	1.76
Trees/shrubs and herbs	Dead-nettle family	Lamiaceae	0.78	0.00	0.00	0.00	0.78
Trees/shrubs and herbs	Bindweed family	Convolvulaceae	0.35	0.35	0.00	0.00	0.00
Trees/shrubs and herbs	Nightshade subfamily	Solanoideae	0.47	0.00	0.00	0.00	0.47
Trees/shrubs and herbs	Knotweed family	Polygonaceae	1.10	0.55	0.55	0.00	0.00
Trees/shrubs and herbs	Mallow order	Malvales	2.92	0.00	0.00	0.00	2.92
Trees/shrubs and herbs	Mallow family	Malvaceae	61.99	7.18	5.87	2.61	46.33
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	1.67	0.00	0.00	0.00	1.67

Continued on next page

Table 4.33 continued

Ecological category	Common name	Taxon	Total	350	250	151	65
		Ingroup	5420.70	909.33	418.24	723.04	3370.09
Trees/shrubs and herbs	Loosestrife family	Lythraceae	1.26	0.00	1.26	0.00	0.00
Trees/shrubs and herbs		Saxifragales	0.96	0.00	0.00	0.19	0.77
Trees/shrubs and herbs		Asparagales	0.40	0.16	0.08	0.00	0.16
Trees/shrubs and herbs		**Proteales**	**0.32**	**0.00**	**0.32**	**0.00**	**0.00**
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.16	0.00	0.00	0.00	0.16
Grasses and relatives	Grass order	Poales	12.59	2.34	2.34	0.00	7.90
Grasses and relatives	Sedges	Cyperaceae	10.80	0.00	1.54	1.54	7.71
Grasses and relatives	True sedges	Carex	22.78	0.00	4.56	2.28	15.94
Grasses and relatives	Rushes	Juncus	1.52	0.00	0.00	0.00	1.52
Grasses and relatives	False brome grasses	Brachypodium	3.63	3.63	0.00	0.00	0.00
Grasses and relatives		Triticodae	0.11	0.00	0.00	0.00	0.11
Grasses and relatives		PACMAD clade	49.61	0.00	0.00	0.00	49.61
Grasses and relatives		Chloridoideae	0.48	0.48	0.00	0.00	0.00
Grasses and relatives	Zoysia tribe	Zoysieae	4.55	0.00	0.00	0.00	4.55
Grasses and relatives	Dropseeds	Sporobolus	1.90	0.95	0.95	0.00	0.00
Grasses and relatives		**Panicoideae**	**3.24**	**0.00**	**1.08**	**0.54**	**1.62**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**0.40**	**0.00**	**0.00**	**0.00**	**0.40**
Grasses and relatives		**Paniceae**	**2.01**	**0.67**	**0.00**	**0.00**	**1.34**
Ferns	Ferns	Polypodiopsida	0.54	0.00	0.47	0.00	0.07
Ferns	Horsetails	Equisetum	0.09	0.00	0.00	0.00	0.09
Ferns		Ophioglossidae	0.03	0.00	0.03	0.00	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	2.30	0.08	2.05	0.00	0.16
Ferns	Polypod ferns	Polypodiales	4.34	0.40	2.98	0.08	0.88
Ferns	Spleenwort suborder	Aspleniineae	5.76	0.27	4.23	0.18	1.08
Ferns		Athyriaceae	1.40	0.07	0.96	0.00	0.37
Ferns		Athyrium	2.92	0.00	2.92	0.00	0.00
Ferns		Thelypteridaceae	0.10	0.00	0.00	0.00	0.10
Ferns		Phegopteridoideae	0.14	0.00	0.00	0.00	0.14
Ferns		Thelypteridoideae	0.17	0.00	0.08	0.00	0.08
Ferns	Bracken family	Dennstaedtiaceae	0.41	0.00	0.00	0.00	0.41
Ferns	Buckler and male ferns	Dryopteris	0.31	0.00	0.15	0.00	0.15
Bryophytes	Mosses	Bryophyta	1.94	0.00	1.94	0.00	0.00
Bryophytes	Feather mosses	Hypnales	2.24	0.00	0.00	0.00	2.24
Bryophytes	Liverworts	Marchantiophyta	0.53	0.53	0.00	0.00	0.00
Mixed		Tracheophyta	3.98	1.21	0.69	0.69	1.38
Mixed		Euphyllophyta	2.58	0.52	0.34	0.00	1.72
Mixed	Monocots	Liliopsida	433.40	72.93	0.00	61.49	298.97

Table 4.34: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF042. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	350	250	151	65
			Ingroup	68.60	26.77	6.21	0.49	35.13
Freshwater aquatics	Fishes	Carp order	Cypriniformes	0.71	0.71	0.00	0.00	0.00
Salt/brackish aquatics	Tunicates	Tunicates	Tunicata	4.07	0.00	0.00	0.00	4.07
Salt/brackish aquatics	Tunicates		Phlebobranchia	7.14	7.14	0.00	0.00	0.00
Salt/brackish aquatics	Annelids		Fabriciidae	0.90	0.90	0.00	0.00	0.00
Salt/brackish aquatics	Annelids		Scoloplos	0.51	0.00	0.00	0.00	0.51
Salt/brackish aquatics	Bryozoans		Adeonidae	1.22	0.00	0.00	0.00	1.22
Salt/brackish aquatics	Polyplacophorans		Neoloricata	0.64	0.64	0.00	0.00	0.00
Salt/brackish aquatics	Polyplacophorans		Chitonida	0.72	0.72	0.00	0.00	0.00
Salt/brackish aquatics	Sea anemones	Sea anemones	Actiniaria	4.33	2.17	0.00	0.00	2.17
Salt/brackish aquatics	Sea anemones		Urticina	5.02	0.00	0.00	0.00	5.02
Salt/brackish aquatics	Ctenophores		Tentaculata	0.58	0.00	0.00	0.00	0.58
Mixed aquatics	Fishes		Euacanthomorphacea	1.06	0.00	0.00	0.00	1.06
Mixed aquatics	Fishes		Percomorphaceae	1.06	0.00	0.00	0.00	1.06
Mixed aquatics	Fishes		Carangaria	1.32	0.00	0.00	0.00	1.32
Mixed aquatics	Fishes	Salmon subfamily	Salmoninae	0.35	0.00	0.00	0.00	0.35
Mixed aquatics	Mammals	Toothed whales	Odontoceti	0.29	0.00	0.00	0.00	0.29
Mixed aquatics	Flatworms		Acotylea	0.45	0.00	0.00	0.00	0.45
Mixed aquatics	Copepods		Neocopepoda	0.62	0.00	0.62	0.00	0.00
Mixed aquatics	Copepods		Harpacticoida	8.00	4.00	4.00	0.00	0.00
Mixed aquatics	Bivalves	Bivalves	Bivalvia	1.74	0.00	0.00	0.00	1.74
Mixed aquatics	Hydrozoans	Hydrozoans	Hydrozoa	0.82	0.00	0.82	0.00	0.00
Mixed aquatics	Hydrozoans	Thecate hydroids	Leptothecata	2.44	2.44	0.00	0.00	0.00
Terrestrial	Mammals	Pig family	Suidae	0.32	0.00	0.32	0.00	0.00
Terrestrial	Birds	Landfowl	Galliformes	0.77	0.77	0.00	0.00	0.00
Terrestrial	Mites		Analgoidea	6.67	6.67	0.00	0.00	0.00
Terrestrial	Mites	Spider mites	Tetranychidae	11.76	0.00	0.00	0.00	11.76
Mixed	Animals	Animals	Metazoa	0.25	0.00	0.25	0.00	0.00
Mixed	Vertebrates	Vertebrates	Vertebrata	0.19	0.00	0.19	0.00	0.00
Mixed	Arthropods	Arthropods	Arthropoda	0.49	0.00	0.00	0.49	0.00
Mixed	Invertebrates	Lophotrochozoans	Lophotrochozoa	3.73	0.62	0.00	0.00	3.10
Mixed	Gastropods	Snails and slugs	Gastropoda	0.43	0.00	0.00	0.00	0.43

Embryophyta

Aquatics

The taxonomic profiles resemble ELF041 but with more data, including more low-frequency taxa. Besides *Zostera*, most or all samples contain *Ruppia*, the reed group, Typhaceae, and *Potamogeton*. Sample 65 also has a rare appearance from *Bolboschoenus* (club-rushes).

Halophytes and xerophytes

Halophytes, primarily Plumbaginaceae (the leadwort family), occur at either end of the core in samples 350 and 65. These are the samples with the most data, so the absence of halophytes from intermediate samples may not be genuine. The two xerophyte taxa also in 65, *Artemisia* and Chenopodiaceae, contain halophytes so could be interpreted as such. Together, these taxa suggest a coastal environment and support the saltwater signal.

Trees/shrubs

Woody taxa are the next most frequent group after salt and mixed aquatics. Every sample has a sizeable signal mostly from taxa in Betulaceae, the willow group, and *Prunus*. The small *Tilia* signal could be bolstered by the strong Malvaceae (trees/shrubs and herbs) found throughout the core.

Grasses and relatives

The grass signal is consistent but perhaps proportionally smaller than in similar *Zostera*-dominated cores. This may indicate a more open aquatic environment than reed bed or saltmarsh.

Metazoa

Most informative Metazoa taxa are saltwater or mixed aquatics, including a rare read from Odontoceti (toothed whales). The high biogenomic mass of the two mite groups (owing to their small genome sizes) makes terrestrial taxa appear significant in the first and top samples, but this seems unlikely.

4.17.3 Pianka scores

The Embryophyta Pianka scores are nearly 1 for all three comparisons, reflecting the highly similar taxonomic profiles, and the lower Metazoa scores acknowledge some similarity while being influenced by the scatter of low-frequency taxa characteristic of low read counts.

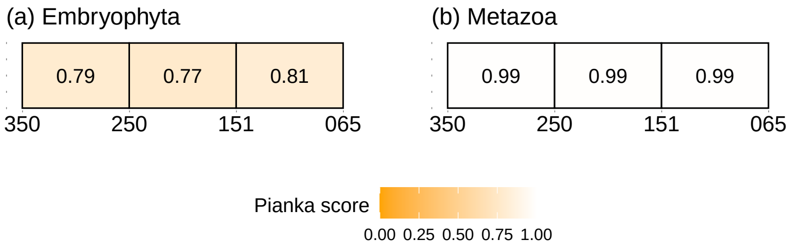


Figure 4.84: Pianka similarity scores between adjacent samples in ELF042 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

4.17.4 Summary

ELF042 is much like ELF041, although with more Embryophyta data and, correspondingly, more *Zostera* and Alismatales, with Liliopsida appearing in the highest-yielding sample. It shows a clearly coastal environment with trees and halophytes from terrestrial habitats, but the reed and more general grass signal may be smaller than in many previous coastal cores, perhaps suggesting a reduced reed or marsh transition zone between land and sea. The Metazoa results support a partially saltwater environment but do not add much detail.

4.18 ELF044

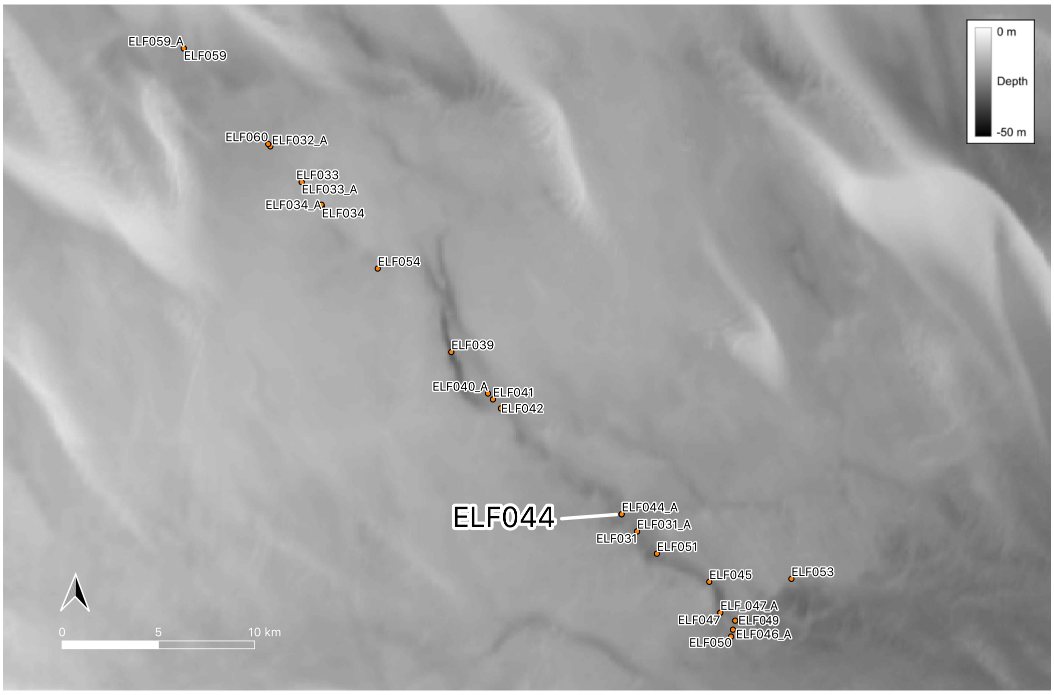


Figure 4.85: Bathymetry map of the palaeochannel transect highlighting ELF044. Darker grey indicates greater depth.

ELF044 is approximately 8 km southeast of ELF042 and two thirds along the visible palaeochannel. It shares this location with ELF044A.

4.18.1 Read counts

Only two samples were taken from ELF044. Both have relatively high Embryophyta read counts, including over 7,000 reads in sample 90, but there is hardly any Metazoa data. Note that neither sample contains data from the replicate extractions: these were excluded because they were associated with an unacceptable amount of contamination.

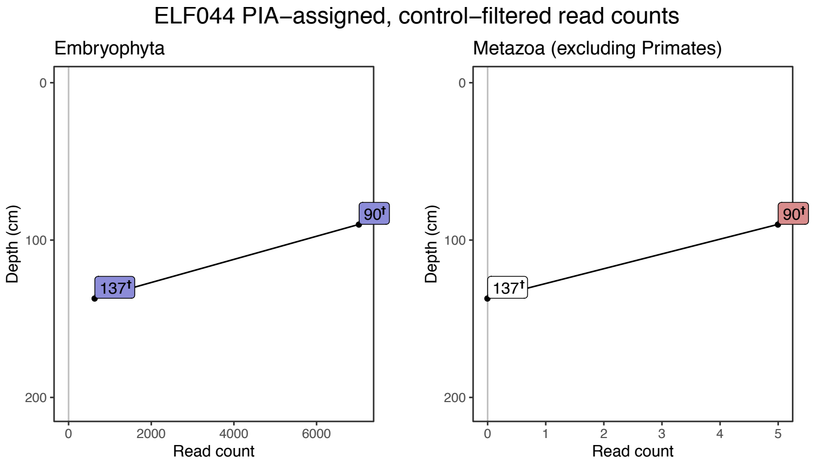


Figure 4.86: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF044. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Neither sample includes data from replicates because they were associated with excessive contamination. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.18.2 Taxonomic profiles

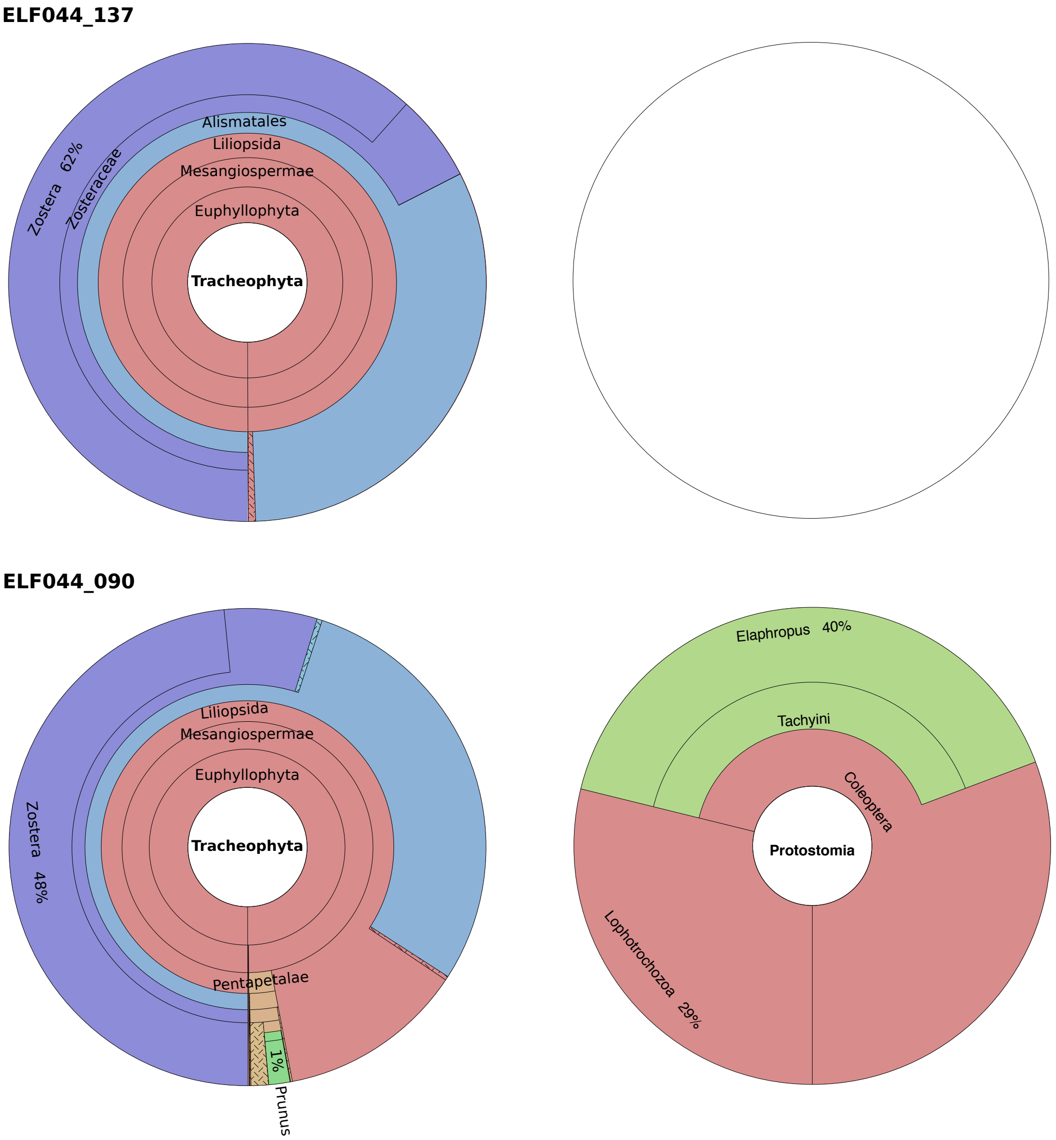


Figure 4.87: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF044. See figure 4.4 for colour key.

Table 4.35: Biogenomic masses of European taxa in Embryophyta from ELF044. Samples are in cm. Taxa in bold are not native to Great Britain. Neither sample includes data from duplicates as they were associated with excessive contamination (†).

Ecological category	Common name	Taxon	Total	137†	90†
		Ingroup	2442.34	281.42	2160.92
Freshwater aquatics	Duckweed family	Lemnoideae	1.22	0.00	1.22
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	1.39	0.00	1.39
Salt/brackish aquatics	Tasselweeds	Ruppia	0.86	0.00	0.86
Salt/brackish aquatics	Eelgrass family	Zosteraceae	152.29	16.51	135.78
Salt/brackish aquatics	Eelgrasses	Zostera	1218.35	172.48	1045.87
Mixed aquatics	Alismatids	Alismatales	717.01	89.46	627.56
Mixed aquatics	Pondweed family	Potamogetonaceae	3.97	0.00	3.97
Mixed aquatics	Reeds	Arundinoideae	0.35	0.00	0.35
Mixed aquatics	Common reed tribe	Molinieae	4.23	0.71	3.53
Mixed aquatics	Common reeds	Phragmites	0.84	0.00	0.84
Trees/shrubs	Birch family	Betulaceae	2.16	0.00	2.16
Trees/shrubs	Willow family	Salicaceae	18.05	0.00	18.05
Trees/shrubs	Willow tribe	Saliceae	3.63	0.00	3.63
Trees/shrubs	Poplars	Populus	1.98	0.00	1.98
Trees/shrubs	Stone fruit trees	Prunus	31.05	0.00	31.05
Trees/shrubs	Apple tribe	Maleae	0.89	0.00	0.89
Herbs	Arum family	Araceae	0.18	0.00	0.18
Herbs	Daffodil family	Amaryllidaceae	0.05	0.00	0.05

Continued on next page

Table 4.35 continued

Ecological category	Common name	Taxon	Total	137†	90†
		Ingroup	2442.34	281.42	2160.92
Freshwater aquatics	Duckweed family	Lemnoideae	1.22	0.00	1.22
Herbs	Orchids	Orchidaceae	0.17	0.00	0.17
Trees/shrubs and herbs		campanulids	0.31	0.00	0.31
Trees/shrubs and herbs	Daisy family	Asteraceae	0.54	0.00	0.54
Trees/shrubs and herbs		Asteroideae	0.25	0.00	0.25
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	0.18	0.00	0.18
Trees/shrubs and herbs	Aster tribe	Astereae	0.52	0.00	0.52
Trees/shrubs and herbs	Mallow family	Malvaceae	2.61	0.00	2.61
Grasses and relatives	Grass order	Poales	0.88	0.00	0.88
Grasses and relatives		**Paniceae**	**0.67**	**0.67**	**0.00**
Ferns	Polypod ferns	Polypodiales	0.16	0.00	0.16
Ferns	Spleenwort suborder	Aspleniineae	0.09	0.00	0.09
Ferns		Phegopteridoideae	0.14	0.00	0.14
Ferns	Buckler and male ferns	Dryopteris	0.08	0.00	0.08
Mixed		Tracheophyta	0.86	0.17	0.69
Mixed		Euphyllophyta	1.89	0.00	1.89
Mixed	Monocots	Liliopsida	271.79	0.00	271.79

Table 4.36: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF044. Samples are in cm. Neither sample includes data from duplicates as they were associated with excessive contamination (†).

Ecological category	Group	Common name	Taxon	Total	137†	90†
			Ingroup	9.76	0.00	9.76
Terrestrial	Coleopterans		Elaphropus	1.74	0.00	1.74
Mixed	Coleopterans	Beetles	Coleoptera	1.32	0.00	1.32
Mixed	Invertebrates	Lophotrochozoans	Lophotrochozoa	1.24	0.00	1.24

Embryophyta

Despite the large number of Embryophyta reads, the taxonomic profile of ELF044 lacks diversity. Both samples are dominated by *Zostera* and Alismatales, with the addition of Liliopsida for the higher-yielding 90. Of lower-frequency taxa, there is a minute reed and grass signal in sample 137 but a more typical range of taxa in 90. This includes Betulaceae and the willow group and also reeds, though other grasses are nearly absent. Evidence of plants other than *Zostera* is proportionally smaller than for most *Zostera*-dominated cores.

Metazoa

Only three taxa were returned, and only for sample 90. Two are higher and ambiguous. However, the third is the ground beetle *Elaphropus*, which provides a modicum of support to the small terrestrial signal from Embryophyta.

4.18.3 Pianka scores

Such a high Pianka score for the Embryophyta data is expected after the highly similar taxonomic profiles. No Metazoa score can be calculated.

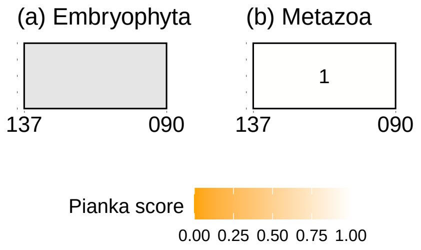


Figure 4.88: Pianka similarity scores between adjacent samples in ELF044 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores. Comparisons involving empty samples have no score.

4.18.4 Summary

The majority of samples in this project return a majority of *Zostera*, but ELF044 does so particularly thoroughly. In the deeper sample, only three of 635 reads are from unrelated plants. The top sample contains a greater diversity, but it is still proportionally small. This is a particularly saltwater core. This may suggest distance; perhaps this sediment was located further from the source of terrestrial DNA than in other cores, possibly further out into an estuary.

4.19 ELF044A

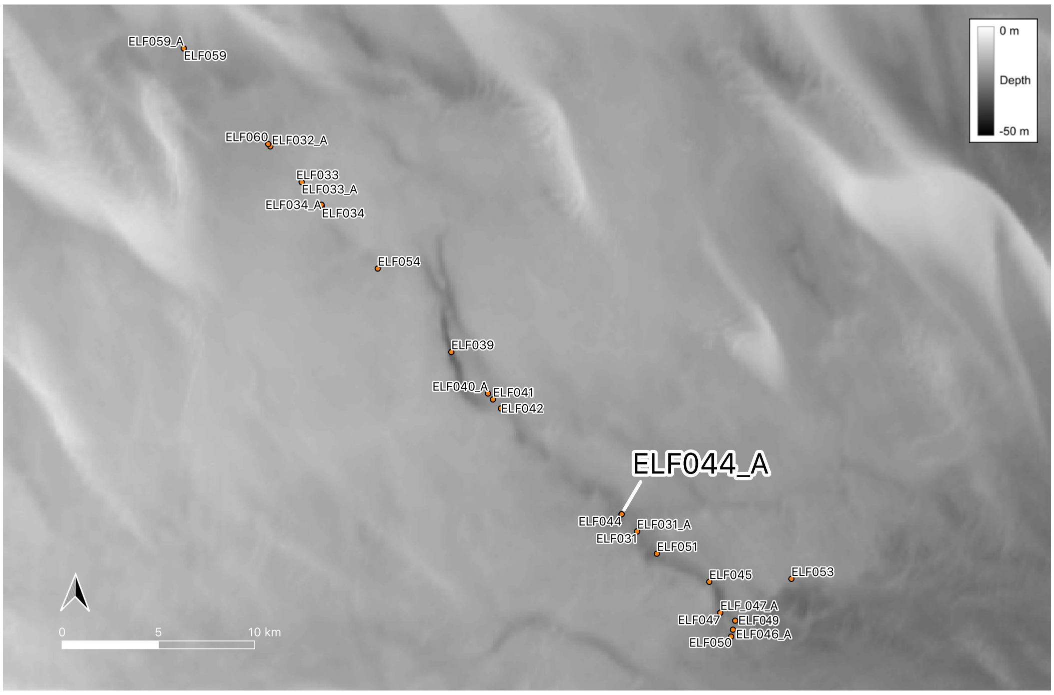


Figure 4.89: Bathymetry map of the palaeochannel transect highlighting ELF044A. Darker grey indicates greater depth.

ELF044A is from the same location as ELF044, approximately two thirds along the visible palaeochannel.

4.19.1 Read counts

Only one sample was taken from ELF044A. It returned only eight Metazoa reads, but over 18,000 from Embryophyta.

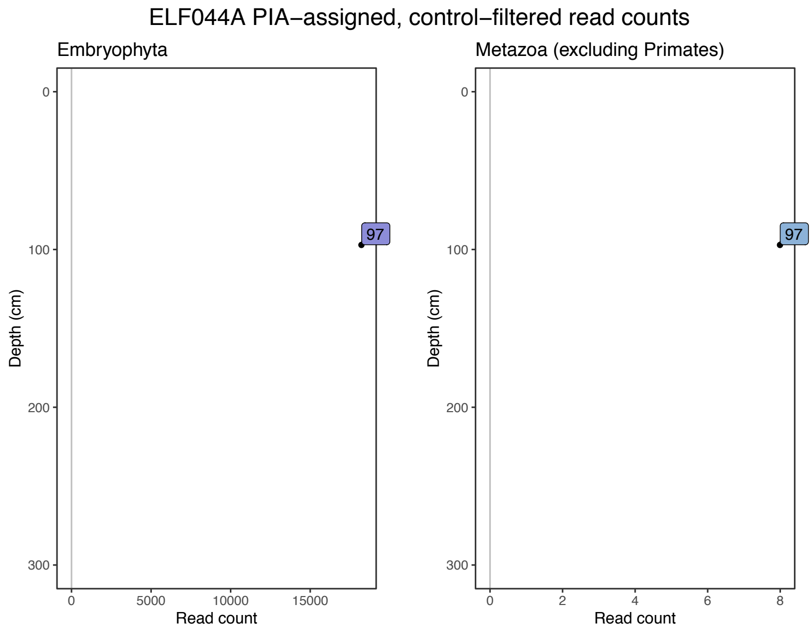


Figure 4.90: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF044A. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.19.2 Taxonomic profiles

ELF044A_097

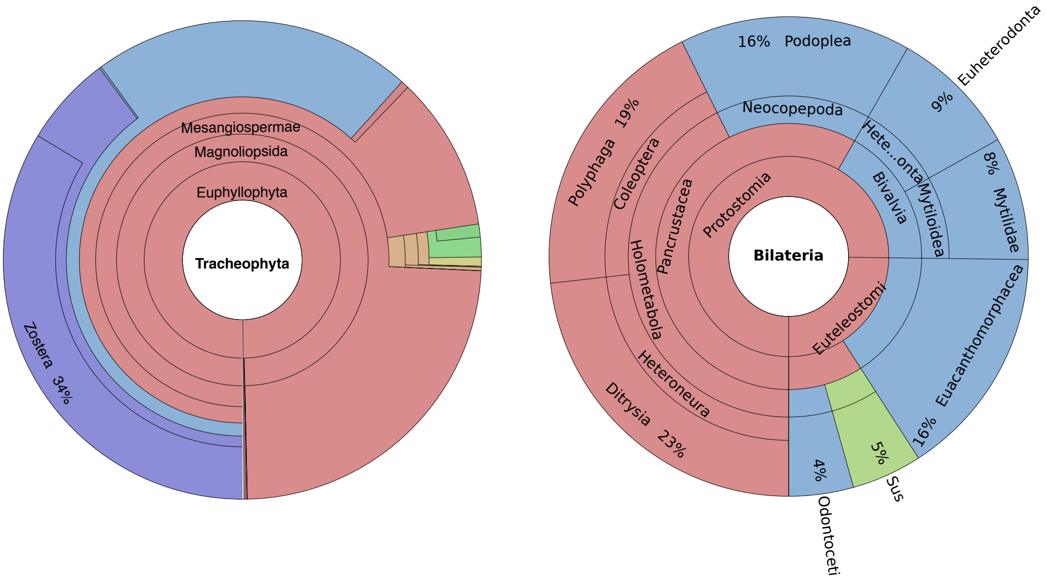


Figure 4.91: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF044A. See figure 4.4 for colour key.

Table 4.37: Biogenomic masses of European taxa in Embryophyta from ELF044A. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	97
		Ingroup	4880.84	4880.84
Freshwater aquatics	Elodea family	Hydrocharitaceae	0.40	0.40
Freshwater aquatics	Bulrushes	Typha	7.55	7.55
Freshwater aquatics	Waterlily family	Nymphaeaceae	0.53	0.53
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	0.70	0.70
Salt/brackish aquatics	Tasselweeds	Ruppia	1.72	1.72
Salt/brackish aquatics	Eelgrass family	Zosteraceae	299.08	299.08
Salt/brackish aquatics	Eelgrasses	Zostera	1644.04	1644.04
Mixed aquatics	Alismatids	Alismatales	1058.71	1058.71
Mixed aquatics	Pondweed family	Potamogetonaceae	2.98	2.98
Mixed aquatics	Reeds	Arundinoideae	1.06	1.06
Mixed aquatics	Common reed tribe	Molinieae	6.70	6.70
Mixed aquatics	Common reeds	Phragmites	0.42	0.42
Halophytes	Leadwort family	Plumbaginaceae	0.62	0.62
Halophytes	Sea-lavenders	Limonium	0.21	0.21
Xerophytes	Mugworts	Artemisia	0.19	0.19
Xerophytes	Goosefoot family	Chenopodiaceae	0.80	0.80
Trees/shrubs	Birch family	Betulaceae	3.24	3.24
Trees/shrubs	Alders	Alnus	1.61	1.61
Trees/shrubs	Oaks	Quercus	4.23	4.23
Trees/shrubs	Willow family	Salicaceae	64.98	64.98
Trees/shrubs	Willow tribe	Saliceae	25.42	25.42
Trees/shrubs	Poplars	Populus	1.98	1.98
Trees/shrubs	Willows	Salix	15.57	15.57
Trees/shrubs	Stone fruit trees	Prunus	14.61	14.61
Trees/shrubs	Apple tribe	Maleae	2.67	2.67
Trees/shrubs	Elm family	Ulmaceae	0.50	0.50
Trees/shrubs	Elms	Ulmus	0.43	0.43
Herbs	Bellflower family	Campanulaceae	0.47	0.47
Herbs	Primrose family	Primulaceae	0.70	0.70
Herbs	Mint tribe	Mentheae	2.46	2.46
Herbs	Broomrape family	Orobanchaceae	0.32	0.32
Herbs	Strawberry subtribe	Fragariinae	1.47	1.47
Herbs	Arum family	Araceae	0.37	0.37
Herbs	Daffodil family	Amaryllidaceae	0.05	0.05
Herbs	Buttercups	Ranunculus	0.15	0.15
Trees/shrubs and herbs		campanulids	0.31	0.31
Trees/shrubs and herbs	Daisy family	Asteraceae	0.27	0.27
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	0.37	0.37
Trees/shrubs and herbs	Dead-nettle order	Lamiales	2.63	2.63
Trees/shrubs and herbs		Solanales	0.38	0.38
Trees/shrubs and herbs	Nightshade subfamily	Solanoideae	0.95	0.95
Trees/shrubs and herbs	Knotweed family	Polygonaceae	0.55	0.55
Trees/shrubs and herbs	Rose subfamily	Rosoideae	1.39	1.39
Trees/shrubs and herbs	Mallow family	Malvaceae	3.26	3.26
Trees/shrubs and herbs		Saxifragales	0.19	0.19
Trees/shrubs and herbs		Asparagales	0.08	0.08
Trees/shrubs and herbs		Liliales	0.03	0.03
Grasses and relatives	Grass order	Poales	4.68	4.68
Grasses and relatives	Sedges	Cyperaceae	1.54	1.54
Grasses and relatives	True sedges	Carex	2.28	2.28
Grasses and relatives		Poeae	0.18	0.18
Grasses and relatives		**Panicoideae**	**1.62**	**1.62**
Ferns	Ferns	Polypodiopsida	0.07	0.07
Ferns	Horsetails	Equisetum	0.05	0.05
Ferns	Leptosporangiate ferns	Polypodiidae	0.41	0.41
Ferns	Polypod ferns	Polypodiales	1.37	1.37
Ferns	Spleenwort suborder	Aspleniineae	1.44	1.44
Ferns		Athyrium	0.14	0.14
Ferns		Phegopteridoideae	0.14	0.14
Ferns		Thelypteridoideae	0.08	0.08
Ferns		Polypodiineae	0.06	0.06
Ferns	Buckler and male ferns	Dryopteris	0.08	0.08
Ferns		Pteridaceae	0.11	0.11
Mixed		Tracheophyta	5.88	5.88
Mixed		Euphyllophyta	6.20	6.20
Mixed	Core angiosperms	Mesangiospermae	1168.16	1168.16
Mixed	Monocots	Liliopsida	508.64	508.64

Table 4.38: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF044A. Samples are in cm.

Ecological category	Group	Common name	Taxon	97
			Ingroup	6.73
Mixed aquatics	Fishes		Euacanthomorphacea	1.06
Mixed aquatics	Mammals	Toothed whales	Odontoceti	0.29
Mixed aquatics	Copepods		Podoplea	1.05
Mixed aquatics	Bivalves		Euheterodonta	0.58
Mixed aquatics	Bivalves	Mussels	Mytilidae	0.56
Terrestrial	Mammals	Pigs	Sus	0.32
Mixed	Lepidopterans		Ditrysia	1.56
Mixed	Coleopterans		Polyphaga	1.31

Embryophyta

As a particularly high-yield sample, it is unsurprising to see most biogenomic mass assigned to *Zostera* and its three most frequent parent taxa: Alismatales, Liliopsida, and Mesangiospermae. Of the low-frequency taxa, the reed group and *Typha* make a small appearance. Four halophytes and xerophytes that could be interpreted as such support the salt signal. Woody taxa are mostly represented by the willow group; grasses are limited. Overall, the plant profile is notable only for the sheer quantity of *Zostera*.

Metazoa

The single sample contains eight taxa, which is relatively diverse for Metazoa. Most are mixed aquatics: fishes, copepods, bivalves, and only the second appearance of Odontoceti (toothed whales; almost certainly marine). The only confirmed terrestrial taxon is *Sus* (pigs), which is most likely a false positive.

4.19.3 Pianka scores

Not applicable as ELF044A only has one sample.

4.19.4 Summary

ELF044A cannot reveal changes over time, but the single sample returned a large amount of data that clearly shows a predominantly saltwater environment. Like similar samples, the *Zostera* and higher taxa majority are accompanied by much smaller tree (mostly the willow group) and reed (with a small amount of other or ambiguous grasses) signals. However, the *Zostera* is particularly strong in ELF044A, as it was in ELF044. The two cores show a very similar, heavily marine environment.

4.20 ELF031

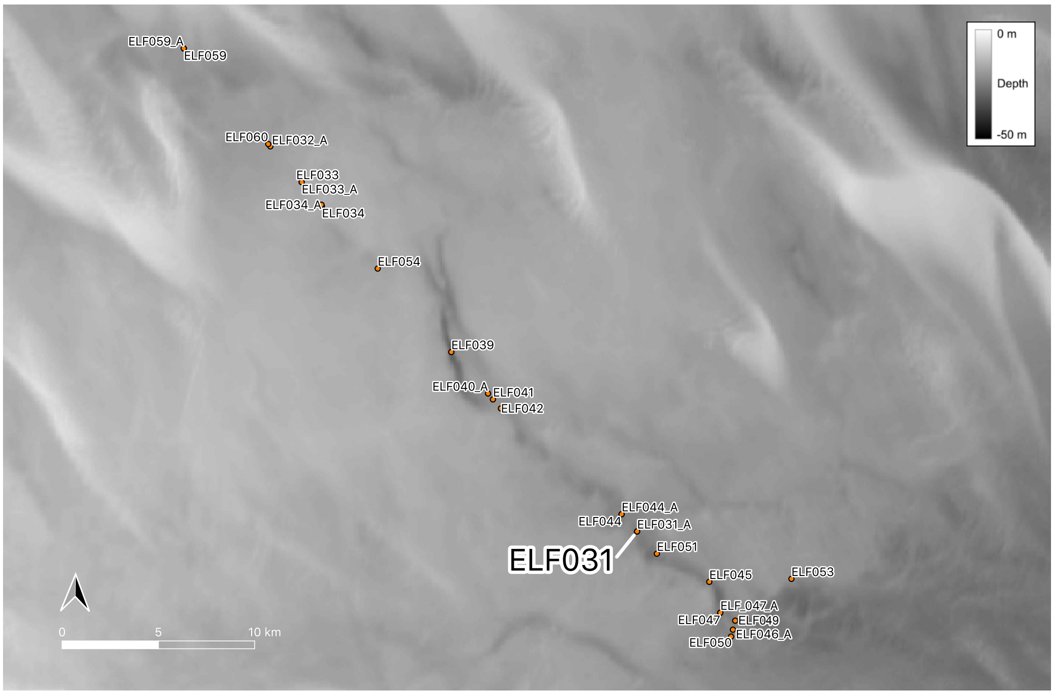


Figure 4.92: Bathymetry map of the palaeochannel transect highlighting ELF031. Darker grey indicates greater depth.

ELF031 is approximately 1 km southeast of ELF044 and ELF044A and 6 km from the mouth of the palaeochannel. It shares its location with ELF031A.

4.20.1 Read counts

There is very little data in this core. Sample 33 is empty, and the maximum number of Embryophyta reads is only 25.

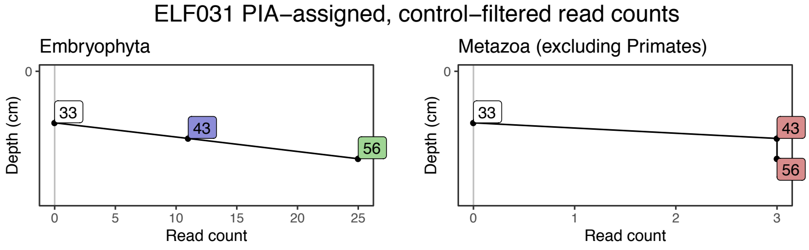
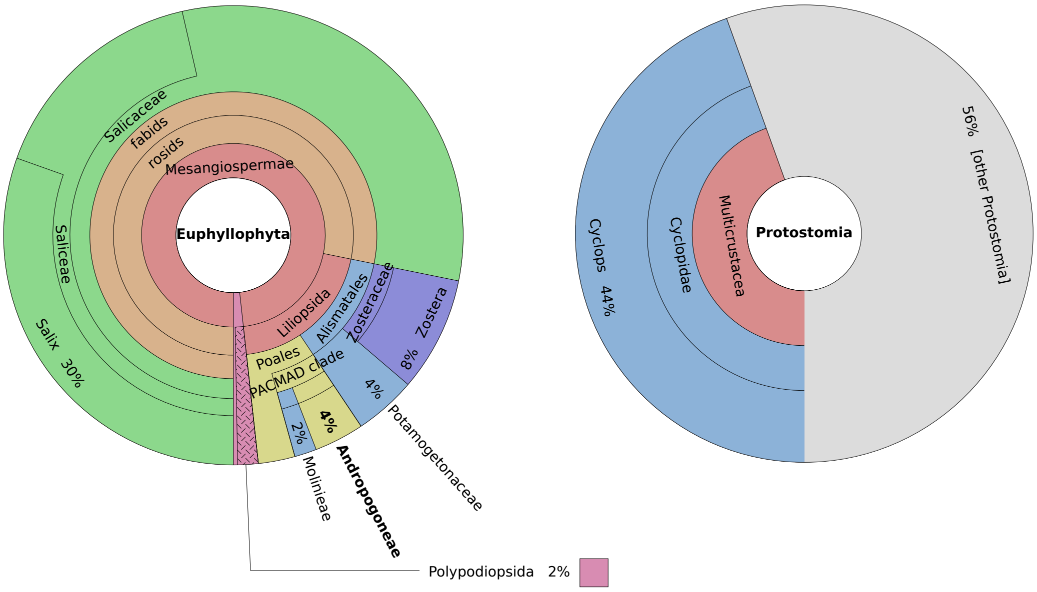


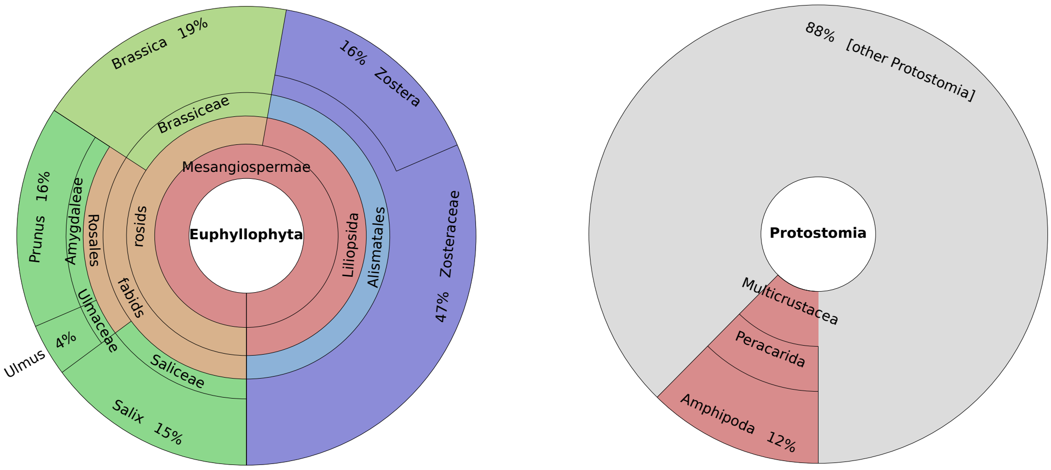
Figure 4.93: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF031. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.20.2 Taxonomic profiles

ELF031_056



ELF031_043



ELF031_033

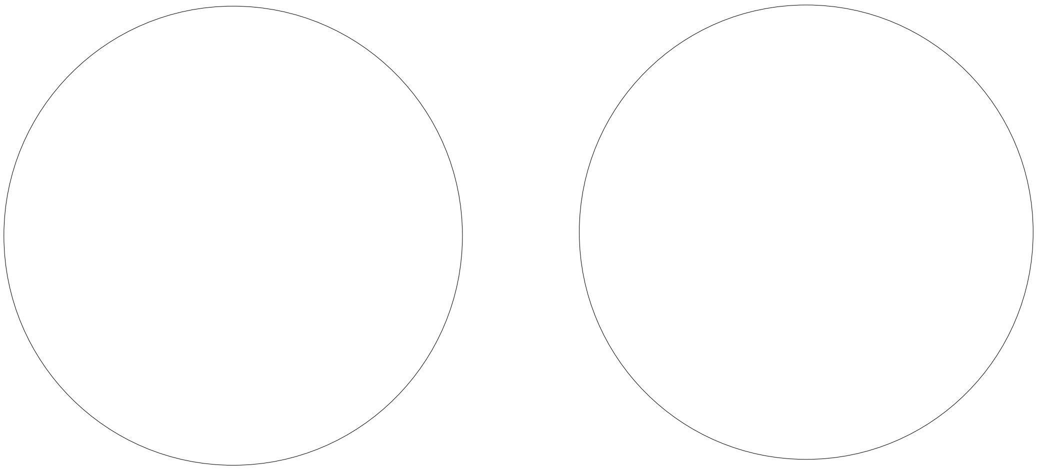


Figure 4.94: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF031. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

Table 4.39: Biogenomic masses of European taxa in Embryophyta from ELF031. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	56	43	33
		Ingroup	38.12	23.49	14.62	0.00
Salt/brackish aquatics	Eelgrass family	Zosteraceae	3.67	0.00	3.67	0.00
Salt/brackish aquatics	Eelgrasses	Zostera	3.67	1.83	1.83	0.00
Mixed aquatics	Pondweed family	Potamogetonaceae	0.99	0.99	0.00	0.00
Mixed aquatics	Common reed tribe	Molinieae	0.35	0.35	0.00	0.00
Trees/shrubs	Willow family	Salicaceae	7.22	7.22	0.00	0.00
Trees/shrubs	Willow tribe	Saliceae	3.63	3.63	0.00	0.00
Trees/shrubs	Willows	Salix	8.65	6.92	1.73	0.00
Trees/shrubs	Stone fruit trees	Prunus	1.83	0.00	1.83	0.00
Trees/shrubs	Elms	Ulmus	0.43	0.00	0.43	0.00
Herbs	Mustards	Brassica	2.17	0.00	2.17	0.00
Grasses and relatives	Grass order	Poales	0.59	0.59	0.00	0.00
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**0.80**	**0.80**	**0.00**	**0.00**
Ferns	Ferns	Polypodiopsida	0.07	0.07	0.00	0.00
Ferns	Polypod ferns	Polypodiales	0.16	0.16	0.00	0.00
Ferns	Spleenwort suborder	Aspleniineae	0.09	0.09	0.00	0.00
Ferns	Buckler and male ferns	Dryopteris	0.08	0.08	0.00	0.00
Mixed		Tracheophyta	0.17	0.17	0.00	0.00

Table 4.40: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF031.

Ecological category	Group	Common name	Taxon	Total	56	43	33
			Ingroup	3.06	1.87	1.19	0.00
Mixed aquatics	Copepods	Water fleas	Cyclops	0.83	0.83	0.00	0.00
Mixed	Invertebrates		Protostomia	2.08	1.04	1.04	0.00
Mixed	Crustaceans	Amphipods	Amphipoda	0.15	0.00	0.15	0.00

Embryophyta

The low read count in ELF031 is a pity because there are suggestions of environmental change.The most frequent ecological category in sample 56 is trees/shrubs (mostly the willow group), but in 43 it is salt/brackish aquatics (*Zostera*). There is significant taxonomic overlap between the samples, but quantities change. Nevertheless, the sampling of the environment is so limited, the difference may be insignificant.

Metazoa

Both samples only contain taxa from Protostomia, a higher invertebrate group. As the highest taxon in the Krona chart, it is coloured grey, but belongs to the mixed ecological category. The *Cyclops* (a water flea) in sample 56 is consistent with the minor aquatic signal from Embryophyta, but the reads in sample 43 are uninformative.

4.20.3 Pianka scores

Pianka scores can only be calculated between the deeper two samples, and the read counts are very low. However, the scores generally agree with the taxonomic profiles. For Embryophyta, taxa differ more by quantity than identity, and the Metazoa taxa all belong to an unusually low root taxon (although not very low), resulting in moderate overlap.

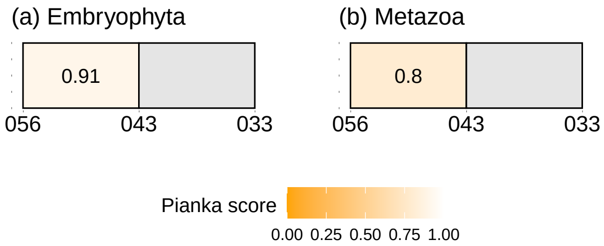


Figure 4.95: Pianka similarity scores between adjacent samples in ELF031 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores. Comparisons involving empty samples have no score.

4.20.4 Summary

The little data in ELF031 may suggest an increase in saltwater aquatics at the expense of woodland, consistent with inundation. However, the read counts too low to take this with confidence.

4.21 ELF031A

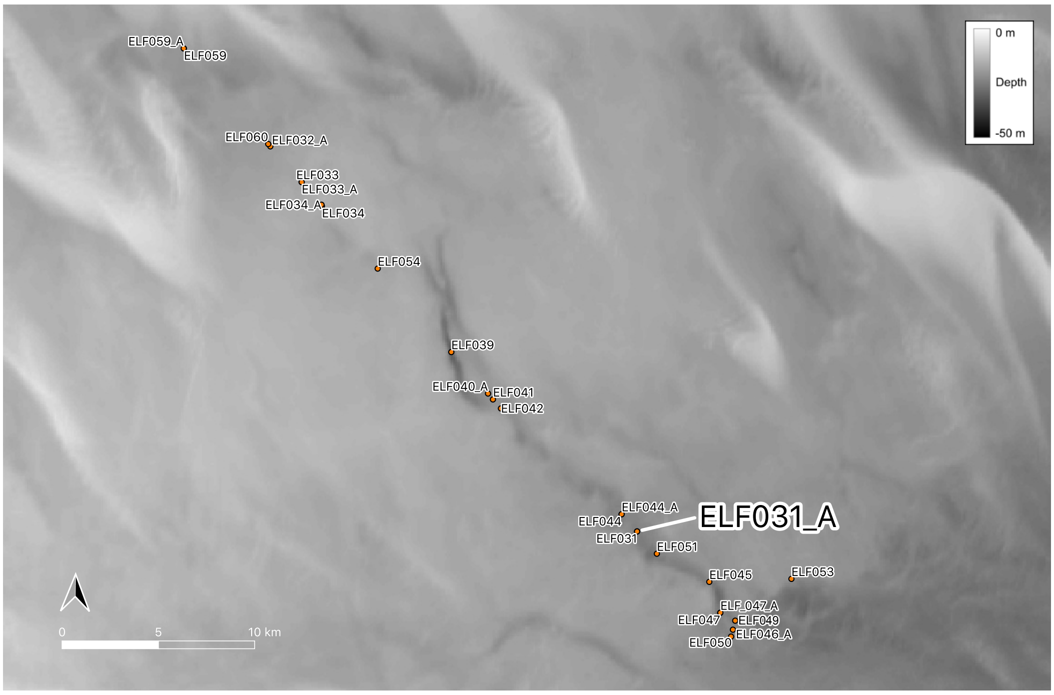


Figure 4.96: Bathymetry map of the palaeochannel transect highlighting ELF031A. Darker grey indicates greater depth.

ELF031A is in the same location as ELF031.

4.21.1 Read counts

ELF031A has eleven samples, the most along with ELF034A, although the deepest four have very little data. Samples 177 to 88 performed better, with several hundred Embryophyta reads, and samples 72 and 58 (*) returned several thousand. 58 was deep-sequenced alongside other samples potentially connected to the Storegga tsunami, yet its Embryophyta read count is actually smaller than 72. It did, however, return by far the largest Metazoa count at 167. If that were from Embryophyta, it would nearly be considered reasonable.

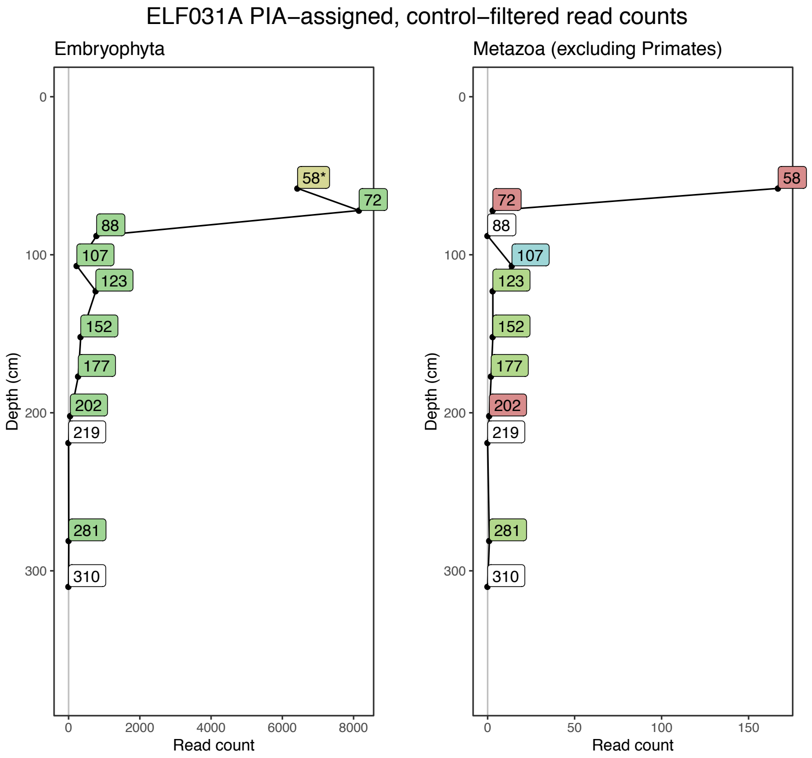
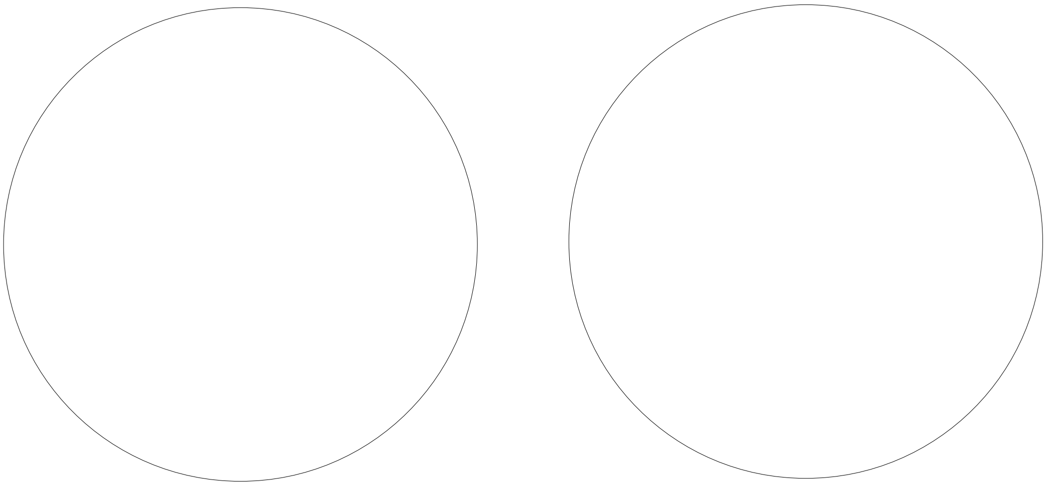


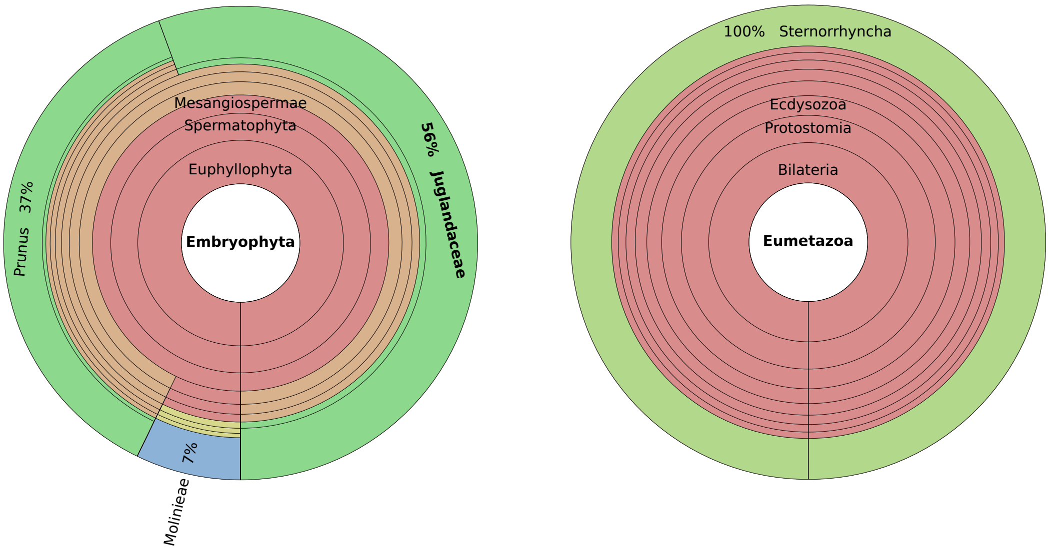
Figure 4.97: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF031A. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample 58 (*) was sequenced in greater depth. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.21.2 Taxonomic profiles

ELF031A_310



ELF031A_281



ELF031A_219

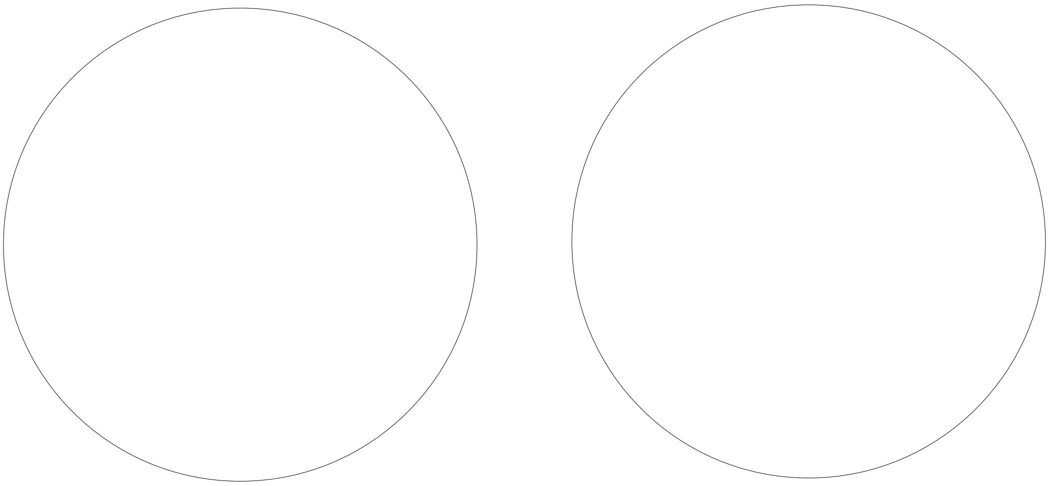
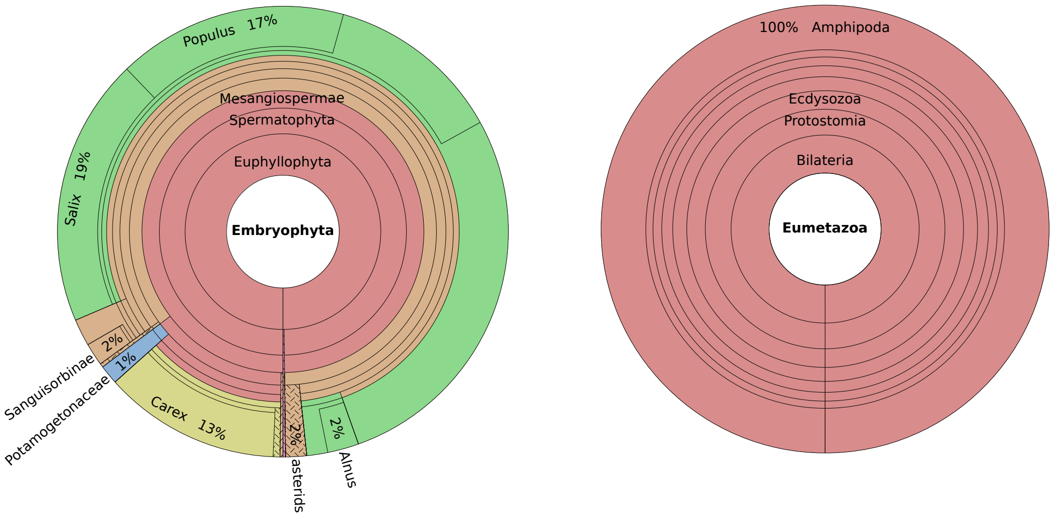
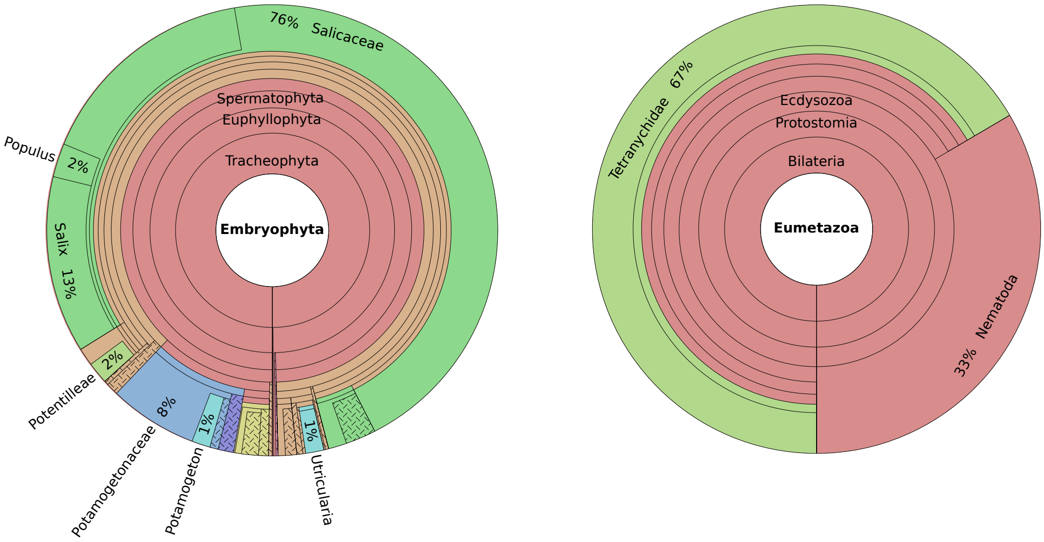


Figure 4.98: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF031A, samples 310, 281, and 219. Continued in figure 4.99. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

ELF031A_202



ELF031A_177



ELF031A_152

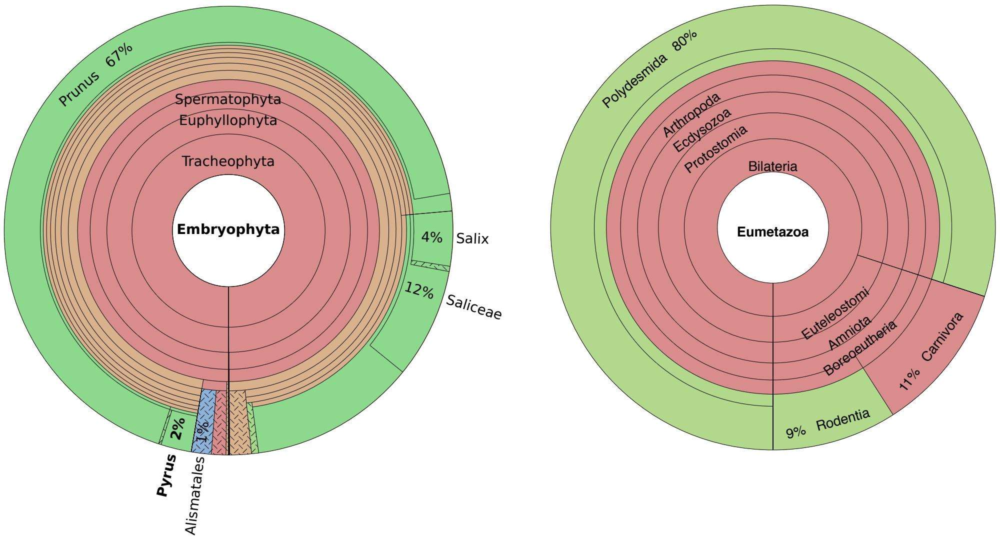
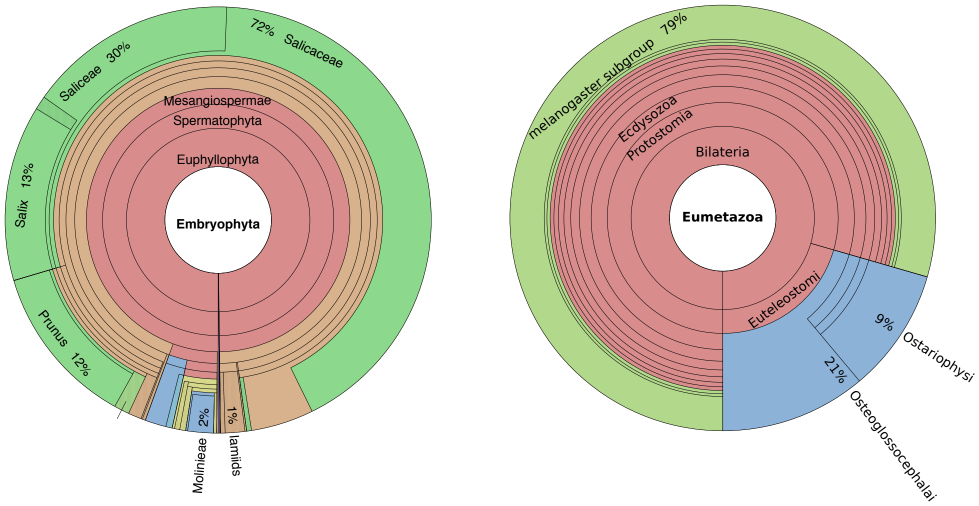
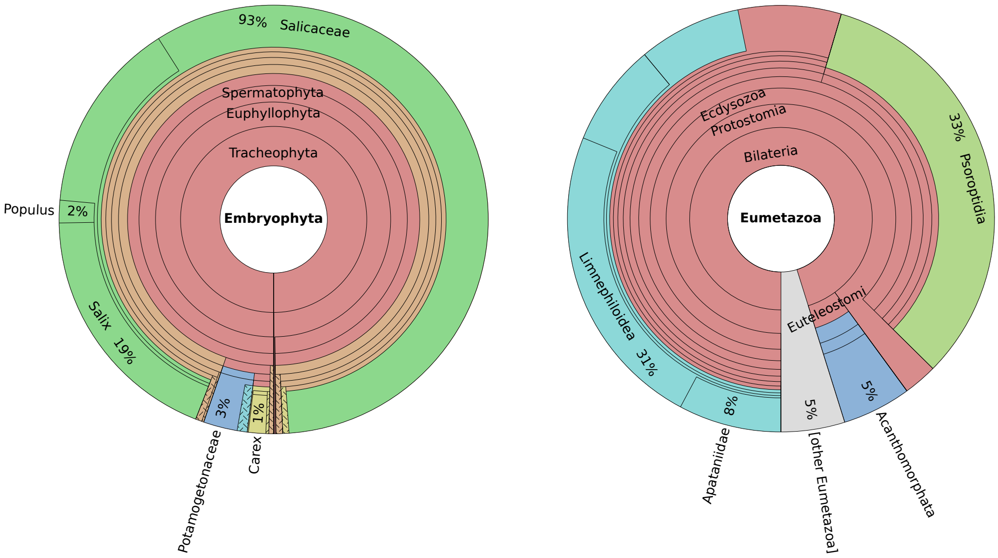


Figure 4.99: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF031A, samples 202, 177, and 152. Continued in figure 4.100. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

ELF031A_123



ELF031A_107



ELF031A_088

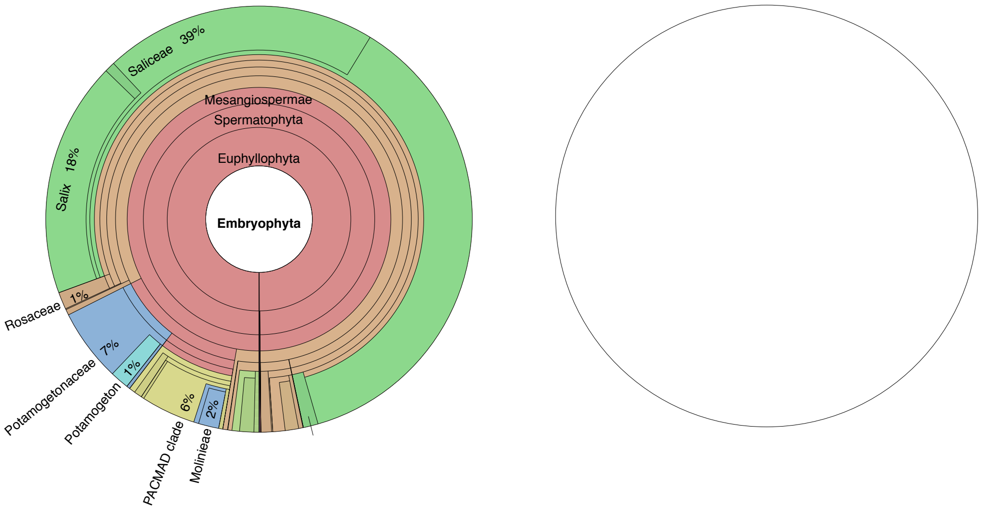
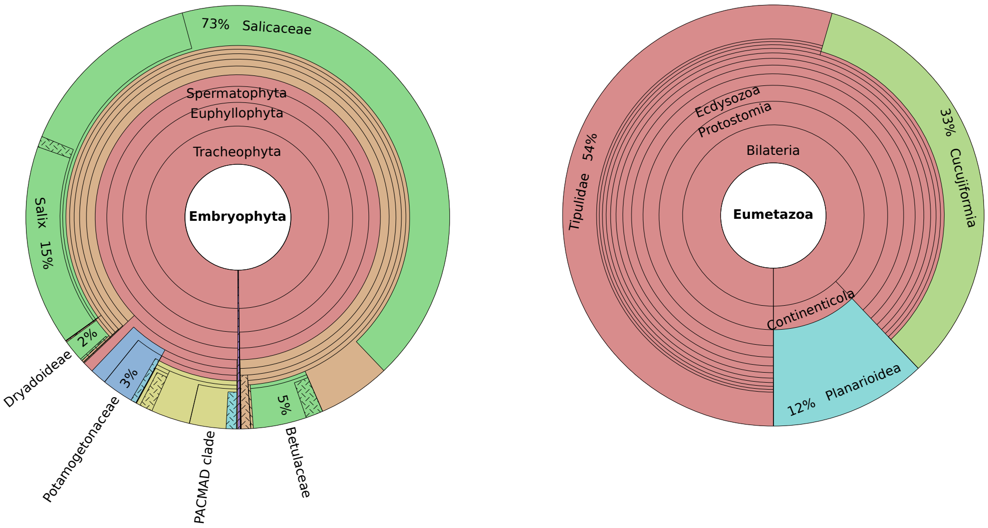


Figure 4.100: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF031A, samples 123, 107, and 88. Continued in figure 4.101. See figure 4.4 for colour key.

ELF031A_072



ELF031A_058*

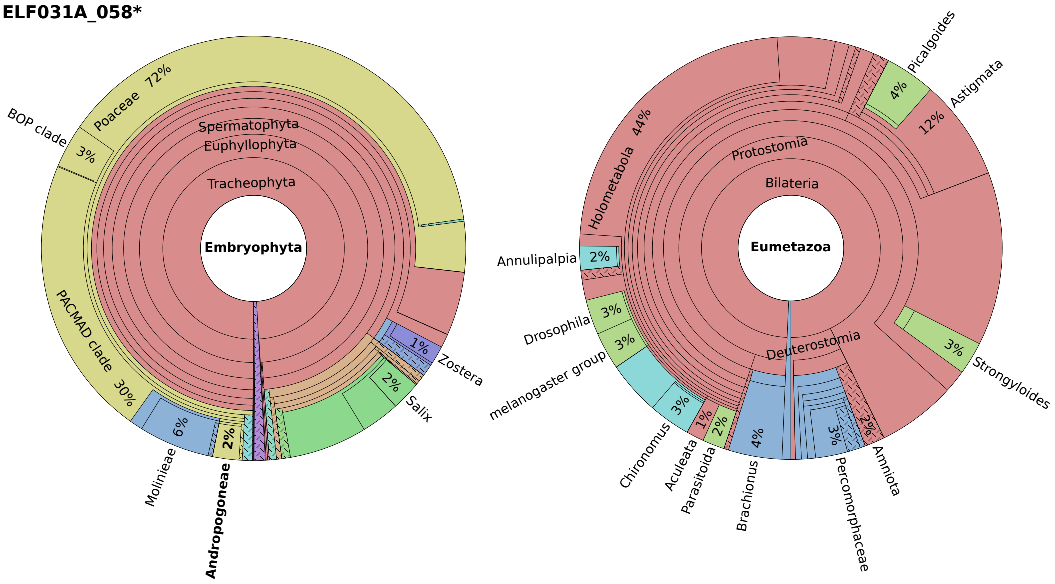


Figure 4.101: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF031A, samples 72 and 58. 58(*) was sequenced in greater depth. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

Table 4.41: Biogenomic masses of European taxa in Embryophyta from ELF031A. Samples are in cm. Sample 58 (*) was resequenced in greater depth. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	310	281	219	202	177	152	123	107	88	72	58*
		Ingroup	12847.55	0.00	5.61	0.00	71.91	328.80	534.26	977.29	350.01	764.60	7875.57	1939.50
Freshwater aquatics	Coontails	Ceratophyllum	10.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.14
Freshwater aquatics	Bladderworts	Utricularia	4.30	0.00	0.00	0.00	0.00	4.30	0.00	0.00	0.00	0.00	0.00	0.00
Freshwater aquatics	Water-milfoil family	Haloragaceae	13.85	0.00	0.00	0.00	0.00	2.31	0.00	0.00	0.00	2.31	2.31	6.92
Freshwater aquatics	Water-milfoils	Myriophyllum	4.62	0.00	0.00	0.00	0.00	0.00	0.00	2.31	0.00	0.00	0.00	2.31
Freshwater aquatics	Water-plantain family	Alismataceae	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07
Freshwater aquatics	Elodea family	Hydrocharitaceae	0.20	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00
Freshwater aquatics	Pondweeds	Potamogeton	39.89	0.00	0.00	0.00	0.00	4.31	1.08	2.16	2.16	10.78	18.87	0.54
Freshwater aquatics	Slenderleaf-pondweed	Stuckenia filiformis	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00
Freshwater aquatics	Rice tribe	Oryzeae	1.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.04
Freshwater aquatics	Bulrush family	Typhaceae	14.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	14.56
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	3.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.48
Salt/brackish aquatics	Tasselweeds	Ruppia	2.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.15
Salt/brackish aquatics	Eelgrass family	Zosteraceae	5.50	0.00	0.00	0.00	0.00	0.00	1.83	0.00	0.00	0.00	0.00	3.67
Salt/brackish aquatics	Eelgrasses	Zostera	29.36	0.00	0.00	0.00	0.00	3.67	0.00	0.00	0.00	0.00	0.00	25.69
Mixed aquatics	Alismatids	Alismatales	98.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	98.79	0.00
Mixed aquatics	Pondweed family	Potamogetonaceae	292.56	0.00	0.00	0.00	0.99	20.83	4.46	15.37	8.93	41.65	189.42	10.91
Mixed aquatics	Pondweeds	Stuckenia	20.33	0.00	0.00	0.00	0.00	1.98	0.00	2.48	0.50	1.98	13.39	0.00
Mixed aquatics	Reeds	Arundinoideae	35.37	0.00	0.00	0.00	0.00	0.35	0.00	1.41	0.00	2.83	12.02	18.74
Mixed aquatics	Common reed tribe	Molinieae	171.68	0.00	0.35	0.00	0.35	1.76	0.00	18.68	0.35	11.63	38.43	100.12
Mixed aquatics	Common reeds	Phragmites	7.56	0.00	0.00	0.00	0.00	0.00	0.42	0.00	0.42	0.00	0.42	6.30
Halophytes	Leadwort family	Plumbaginaceae	0.83	0.00	0.00	0.00	0.00	0.00	0.21	0.21	0.00	0.21	0.21	0.00
Halophytes	Sea-lavenders	Limonium	0.21	0.00	0.00	0.00	0.00	0.21	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs		Viburnum	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.53
Trees/shrubs		Cornales	0.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.61
Trees/shrubs	Heather family	Ericaceae	16.69	0.00	0.00	0.00	0.00	0.00	0.00	1.33	0.00	0.67	14.68	0.00
Trees/shrubs	Bearberry subfamily	Arbutoideae	4.80	0.00	0.00	0.00	0.00	0.00	0.00	0.80	0.00	0.00	4.00	0.00
Trees/shrubs	Crowberries	Empetrum	1.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.04	0.00
Trees/shrubs	Staff vine family	Celastraceae	1.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.20	0.00
Trees/shrubs	Beech order	Fagales	2.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.12
Trees/shrubs	Birch family	Betulaceae	339.08	0.00	0.00	0.00	1.08	4.32	1.08	0.00	0.00	7.56	318.56	6.48
Trees/shrubs	Alders	Alnus	66.13	0.00	0.00	0.00	1.61	3.23	0.00	1.61	0.00	0.00	58.06	1.61
Trees/shrubs	Birches	Betula	54.25	0.00	0.00	0.00	0.00	3.88	0.97	1.94	0.97	0.97	45.53	0.00
Trees/shrubs	Oak family	Fagaceae	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.98
Trees/shrubs	Oaks	Quercus	1.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.06
Trees/shrubs	**Walnut family**	**Juglandaceae**	**2.73**	**0.00**	**2.73**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Trees/shrubs	Willow family	Salicaceae	4530.69	0.00	0.00	0.00	19.86	148.01	63.18	404.33	202.17	277.98	3305.05	110.11
Trees/shrubs	Willow tribe	Saliceae	1659.30	0.00	0.00	0.00	9.08	52.65	41.75	152.50	50.83	156.13	1141.91	54.46
Trees/shrubs	Poplars	Populus	108.79	0.00	0.00	0.00	11.87	7.91	1.98	9.89	5.93	5.93	65.27	0.00
Trees/shrubs	Willows	Salix	1647.06	0.00	0.00	0.00	13.84	41.52	20.76	128.03	65.74	134.95	1200.69	41.52
Trees/shrubs	Buckthorn family	Rhamnaceae	1.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.92	0.00
Trees/shrubs		Amygdaleae	6.63	0.00	0.00	0.00	0.00	0.00	6.63	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Stone fruit trees	Prunus	482.12	0.00	1.83	0.00	0.00	0.00	354.28	118.70	0.00	0.00	7.30	0.00
Trees/shrubs	Apple tribe	Maleae	1.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.78
Trees/shrubs	Apples	Malus	2.97	0.00	0.00	0.00	0.00	0.00	0.99	1.98	0.00	0.00	0.00	0.00

Continued on next page

Table 4.41 continued

Ecological category	Common name	Taxon	Total	310	281	219	202	177	152	123	107	88	72	58*
		Ingroup	12847.55	0.00	5.61	0.00	71.91	328.80	534.26	977.29	350.01	764.60	7875.57	1939.50
Trees/shrubs	**Pears**	**Pyrus**	**13.22**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**11.57**	**1.65**	**0.00**	**0.00**	**0.00**	**0.00**
Trees/shrubs	Dryas subfamily	Dryadoideae	136.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.67	5.00	130.00	0.00
Trees/shrubs		Dryas	13.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.67	11.67	0.00
Trees/shrubs	Witch-hazel family	Hamamelidaceae	0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.77	0.00
Trees/shrubs	Barberry subfamily	Berberidoideae	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00
Herbs	Thistle tribe	Cardueae	0.48	0.00	0.00	0.00	0.00	0.48	0.00	0.00	0.00	0.00	0.00	0.00
Herbs	Chicory tribe	Cichorieae	0.56	0.00	0.00	0.00	0.00	0.00	0.28	0.00	0.00	0.00	0.00	0.28
Herbs	Cat's ear subtribe	Hypochaeridinae	0.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.54	0.00
Herbs	Bellflower family	Campanulaceae	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.47	0.00
Herbs	Wintergreen subfamily	Pyroloideae	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00
Herbs	Wintergreen tribe	Pyroleae	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.00	0.00	0.00	0.00
Herbs	Bedstraw family	Rubiaceae	2.02	0.00	0.00	0.00	0.00	0.00	1.01	0.00	0.00	0.00	0.00	1.01
Herbs		Rubioideae	0.72	0.00	0.00	0.00	0.00	0.00	0.00	0.72	0.00	0.00	0.00	0.00
Herbs	Madder tribe	Rubieae	0.71	0.00	0.00	0.00	0.00	0.71	0.00	0.00	0.00	0.00	0.00	0.00
Herbs	Bedstraw	Galium	0.73	0.00	0.00	0.00	0.00	0.73	0.00	0.00	0.00	0.00	0.00	0.00
Herbs	Dead-nettle subfamily	Lamioideae	0.63	0.00	0.00	0.00	0.00	0.63	0.00	0.00	0.00	0.00	0.00	0.00
Herbs	Catmint subfamily	Nepetoideae	4.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.39	1.13
Herbs	Mint tribe	Mentheae	18.41	0.00	0.00	0.00	0.00	0.00	0.00	1.23	0.00	4.91	12.28	0.00
Herbs	Oreganos	Origanum	1.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.34	0.00	0.00
Herbs	Broomrape family	Orobanchaceae	1.29	0.00	0.00	0.00	0.00	0.00	0.00	0.32	0.00	0.32	0.64	0.00
Herbs	Louseworts	Pedicularis	0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.00	0.00
Herbs	Speedwell tribe	Veroniceae	1.13	0.00	0.00	0.00	0.00	0.00	0.00	1.13	0.00	0.00	0.00	0.00
Herbs	Dodders	Cuscuta	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14
Herbs	Campions and catchflies	Silene	0.45	0.00	0.00	0.00	0.00	0.00	0.00	0.45	0.00	0.00	0.00	0.00
Herbs	Knotweed subfamily	Polygonoideae	2.23	0.00	0.00	0.00	0.00	0.00	1.12	0.00	0.00	0.00	1.12	0.00
Herbs		Hologalegina	1.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.67	0.00
Herbs		Galegeae	11.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00	10.77	0.00
Herbs	Violets	Viola	0.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.73	0.00
Herbs	**Bridal-wreath tribe**	**Spiraeeae**	**9.34**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**7.47**	**0.00**	**1.87**	**0.00**	**0.00**
Herbs	Cinquefoil tribe	Potentilleae	11.44	0.00	0.00	0.00	0.00	5.09	0.00	1.27	0.00	1.27	3.81	0.00
Herbs	Strawberry subtribe	Fragariinae	1.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.47	0.00
Herbs	Cinquefoils	Potentilla	4.25	0.00	0.00	0.00	0.00	0.00	0.00	4.25	0.00	0.00	0.00	0.00
Herbs	Meadowsweets	Filipendula	2.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.50	0.00
Herbs	Whitlow-grasses	Draba	2.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.41	0.00
Herbs		Camelineae	3.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.14
Herbs		Saxifragaceae	0.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.69	0.00
Herbs	Saxifrages	Saxifraga	0.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.69	0.00
Herbs	Arum subfamily	Aroideae	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00
Herbs	Bluebell family	Hyacinthaceae	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16
Herbs	Irises	Iris	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00
Herbs	Orchids	Orchidaceae	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00
Herbs	Buttercup subfamily	Ranunculoideae	6.47	0.00	0.00	0.00	0.00	0.32	0.00	0.11	0.22	4.31	1.40	0.11
Herbs	Buttercup tribe	Ranunculeae	2.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.73	0.16	0.31
Herbs	Buttercups	Ranunculus	9.98	0.00	0.00	0.00	0.15	0.30	0.00	0.00	0.15	6.40	2.53	0.45
Herbs		Thalictroideae	4.82	0.00	0.00	0.00	0.00	0.00	0.96	0.96	0.00	2.89	0.00	0.00
Trees/shrubs and herbs	Gymnosperms	Acrogymnospermae	0.11	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Gunneridae	1.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.88	0.00
Trees/shrubs and herbs		campanulids	1.54	0.00	0.00	0.00	0.00	0.00	0.00	0.62	0.00	0.00	0.00	0.92
Trees/shrubs and herbs	Ivy family	Araliaceae	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50

Continued on next page

Table 4.41 continued

Ecological category	Common name	Taxon	Total	310	281	219	202	177	152	123	107	88	72	58*
		Ingroup	12847.55	0.00	5.61	0.00	71.91	328.80	534.26	977.29	350.01	764.60	7875.57	1939.50
Trees/shrubs and herbs		Asterales	4.56	0.00	0.00	0.00	0.00	1.71	0.00	0.29	0.00	0.86	1.71	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	5.70	0.00	0.00	0.00	0.54	0.81	0.00	0.00	0.27	0.81	3.26	0.00
Trees/shrubs and herbs		Asteroideae	2.51	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.25	0.75	1.25	0.00
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	5.30	0.00	0.00	0.00	0.18	1.10	0.18	0.00	0.00	1.46	2.38	0.00
Trees/shrubs and herbs		Artemisiinae	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00	0.00
Trees/shrubs and herbs		Gnaphalieae	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.97	0.00	0.00
Trees/shrubs and herbs		Dipsacales	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.44
Trees/shrubs and herbs		Ericales	1.42	0.00	0.00	0.00	0.36	0.00	0.00	0.00	0.00	0.36	0.71	0.00
Trees/shrubs and herbs		lamiids	2.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.68	0.68	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	9.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.02	0.88	1.76
Trees/shrubs and herbs	Dead-nettle family	Lamiaceae	2.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.78	1.57	0.00
Trees/shrubs and herbs		Solanales	2.63	0.00	0.00	0.00	0.00	0.00	0.75	1.13	0.00	0.38	0.38	0.00
Trees/shrubs and herbs	Bindweed family	Convolvulaceae	19.19	0.00	0.00	0.00	0.00	0.35	4.19	9.77	0.00	2.09	2.79	0.00
Trees/shrubs and herbs	Knotweed family	Polygonaceae	4.97	0.00	0.00	0.00	0.00	0.55	0.55	0.55	0.55	0.00	2.76	0.00
Trees/shrubs and herbs	Milkvetches	Astragalus	2.42	0.00	0.00	0.00	0.00	0.00	0.61	0.00	0.00	0.00	1.82	0.00
Trees/shrubs and herbs		Malpighiales	481.50	0.00	0.00	0.00	0.00	0.00	0.00	45.73	0.00	0.00	435.77	0.00
Trees/shrubs and herbs	Rose subfamily	Rosoideae	13.91	0.00	0.00	0.00	1.39	4.17	0.00	4.17	0.00	0.00	4.17	0.00
Trees/shrubs and herbs	Burnet subtribe	Sanguisorbinae	1.17	0.00	0.00	0.00	1.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mustard order	Brassicales	1.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.30
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	2.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.56	0.56	0.56
Trees/shrubs and herbs		Myrtales	1.98	0.00	0.00	0.00	0.00	0.99	0.00	0.99	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Saxifragales	2.89	0.00	0.00	0.00	0.19	0.58	0.00	0.00	0.00	0.96	0.58	0.58
Trees/shrubs and herbs		Asparagales	0.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.24
Trees/shrubs and herbs	**Palms**	**Arecaceae**	**0.29**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.29**
Trees/shrubs and herbs		Liliales	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00
Trees/shrubs and herbs	Buttercup order	Ranunculales	5.29	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.78	2.64	1.55	0.16
Trees/shrubs and herbs	Buttercup family	Ranunculaceae	4.25	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.11	2.30	1.49	0.23
Grasses and relatives	Grass order	Poales	78.17	0.00	0.00	0.00	0.00	0.29	0.29	1.76	0.29	3.51	0.88	71.14
Grasses and relatives	Sedges	Cyperaceae	9.25	0.00	0.00	0.00	0.00	1.54	1.54	1.54	0.00	0.00	3.08	1.54
Grasses and relatives		Cyperoideae	27.00	0.00	0.00	0.00	0.00	1.69	0.00	0.00	0.00	0.00	23.62	1.69
Grasses and relatives	True sedges	Carex	102.50	0.00	0.00	0.00	9.11	2.28	2.28	2.28	4.56	4.56	77.45	0.00
Grasses and relatives	Grass family	Poaceae	938.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	232.60	706.37
Grasses and relatives		BOP clade	61.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	61.67
Grasses and relatives		Poeae Chloroplast Group 1 (Aveneae type)	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00
Grasses and relatives	Fescues	Festuca	0.36	0.00	0.00	0.00	0.00	0.18	0.00	0.00	0.00	0.00	0.18	0.00
Grasses and relatives	Ryegrasses	Lolium	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31
Grasses and relatives		Triticodae	3.37	0.00	0.00	0.00	0.00	0.00	0.45	1.80	0.00	1.12	0.00	0.00
Grasses and relatives	Barley subtribe	Hordeinae	1.59	0.00	0.00	0.00	0.00	0.00	0.12	0.85	0.00	0.12	0.49	0.00
Grasses and relatives	Barlies	Hordeum	0.26	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.00
Grasses and relatives	**Ryes**	**Secale**	**1.26**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.13**	**0.75**	**0.00**	**0.13**	**0.25**	**0.00**
Grasses and relatives	**Goat grasses**	**Aegilops**	**0.19**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.10**	**0.10**	**0.00**	**0.00**	**0.00**	**0.00**
Grasses and relatives	Wheat grasses	Thinopyrum	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00
Grasses and relatives	**Einkorn wheat**	**Triticum monococcum**	**1.07**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.15**	**0.61**	**0.00**	**0.00**	**0.31**	**0.00**
Grasses and relatives		PACMAD clade	646.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	31.25	217.78	397.37
Grasses and relatives		Chloridoideae	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.48	0.00	0.00
Grasses and relatives		**Panicoideae**	**1.08**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**1.08**	**0.00**	**0.00**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**50.31**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**1.60**	**0.00**	**0.80**	**10.78**	**37.13**
Grasses and relatives		**Paniceae**	**5.36**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.67**	**0.00**	**0.00**	**0.67**	**4.02**
Grasses and relatives		**Cenchrinae**	**0.58**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.58**

Continued on next page

Table 4.41 continued

Ecological category	Common name	Taxon	Total	310	281	219	202	177	152	123	107	88	72	58*
		Ingroup	12847.55	0.00	5.61	0.00	71.91	328.80	534.26	977.29	350.01	764.60	7875.57	1939.50
Ferns	Ferns	Polypodiopsida	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.27	0.00
Ferns	Horsetails	Equisetum	0.97	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.92	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	0.16	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.00	0.00	0.00	0.00
Ferns	Polypod ferns	Polypodiales	1.61	0.00	0.00	0.00	0.00	0.24	0.08	0.24	0.00	0.00	0.00	1.05
Ferns	Spleenwort suborder	Aspleniineae	1.26	0.00	0.00	0.00	0.09	0.09	0.00	0.09	0.09	0.00	0.00	0.90
Ferns		Thelypteridaceae	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10
Ferns		Phegopteridoideae	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14
Ferns		Thelypteridoideae	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17
Bryophytes	Mosses	Bryophyta	1.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.94	0.00
Bryophytes		Bryophytina	3.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.95	1.95
Bryophytes		Bryopsida	3.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.97	0.00
Bryophytes		Bryanae	1.29	0.00	0.00	0.00	0.00	0.00	0.00	1.29	0.00	0.00	0.00	0.00
Bryophytes		Bryaceae	1.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.15	0.00
Bryophytes		Pottiaceae	5.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.56	0.00
Bryophytes	Liverworts	Marchantiophyta	1.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.53	0.00	1.06
Bryophytes		Jungermanniopsida	1.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.08
Bryophytes		Jungermanniidae	0.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.94
Bryophytes		Jungermanniales	3.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.35
Bryophytes		Cephaloziineae	1.62	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.62
Bryophytes		Porellales	1.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.59
Bryophytes		Lejeuneaceae	4.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.76
Mixed		Tracheophyta	3.63	0.00	0.00	0.00	0.00	0.35	0.17	0.00	0.17	0.00	1.73	1.21
Mixed		Euphyllophyta	4.13	0.00	0.00	0.00	0.00	0.52	0.00	0.00	0.34	0.52	0.69	2.07
Mixed	Monocots	Liliopsida	65.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	65.17	0.00
Mixed		Petrosaviidae	19.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	19.10
Mixed		commelinids	89.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	89.53

Table 4.42: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF031A. Samples are in cm. Sample 58 (*) was resequenced in greater depth. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	310	281	219	202	177	152	123	107	88	72	58*
			Ingroup	241.70	0.00	1.93	0.00	0.15	17.68	3.05	7.44	20.30	0.00	4.22	186.94
Freshwater aquatics	Flatworms		Planarioidea	0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.51	0.00
Freshwater aquatics	Crustaceans		Niphargus	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18
Freshwater aquatics	Trichopterans	Caddisflies	Trichoptera	1.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.58	0.00	0.00	0.00
Freshwater aquatics	Trichopterans	Fixed-retreat makers	Annulipalpia	3.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.16
Freshwater aquatics	Trichopterans		Plenitentoria	1.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.58	0.00	0.00	0.00
Freshwater aquatics	Trichopterans		Limnephiloidea	4.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.75	0.00	0.00	0.00
Freshwater aquatics	Trichopterans		Apataniidae	1.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.58	0.00	0.00	0.00
Freshwater aquatics	Dipterans	Non-biting midges	Chironomidae	7.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.11
Freshwater aquatics	Dipterans		Chironomus	5.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.53
Salt/brackish aquatics	Bivalves		Mytilus	0.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.61
Mixed aquatics	Fishes		Actinopteri	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.80
Mixed aquatics	Fishes		Neopterygii	0.81	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.81
Mixed aquatics	Fishes		Osteoglossocephalai	0.83	0.00	0.00	0.00	0.00	0.00	0.00	0.83	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Acanthomorphata	2.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.06	0.00	0.00	1.06
Mixed aquatics	Fishes		Percomorphaceae	4.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.22
Mixed aquatics	Fishes		Ovalentaria	0.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.93
Mixed aquatics	**Fishes**	**Toothcarps**	**Cyprinodontiformes**	**0.78**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.78**
Mixed aquatics	Fishes	Salmon family	Salmonidae	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35
Mixed aquatics	Fishes	Salmon subfamily	Salmoninae	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35
Mixed aquatics	Fishes	Carp subcohort	Ostariophysi	0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.00	0.00	0.00	0.00
Mixed aquatics	Crabs		Eubrachyura	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33
Mixed aquatics	Rotifers		Brachionus	7.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.10
Mixed aquatics	Cnidarians	Cnidarians	Cnidaria	1.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.17
Terrestrial	Mammals	Rodents	Rodentia	0.28	0.00	0.00	0.00	0.00	0.00	0.28	0.00	0.00	0.00	0.00	0.00
Terrestrial	Mammals	Cattle subfamily	Bovinae	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.27
Terrestrial	Nematodes		Strongyloides	4.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.35
Terrestrial	Mites		Psoroptidia	6.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.67	0.00	0.00	0.00
Terrestrial	Mites	Group of feather mites	Picalgoides	6.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.67
Terrestrial	Mites	Spider mites	Tetranychidae	11.76	0.00	0.00	0.00	0.00	11.76	0.00	0.00	0.00	0.00	0.00	0.00
Terrestrial	Spiders	Thin-legged wolf spiders	Pardosa	0.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.49
Terrestrial	Arthropods	Flat-backed millipedes	Polydesmida	2.44	0.00	0.00	0.00	0.00	0.00	2.44	0.00	0.00	0.00	0.00	0.00
Terrestrial	Coleopterans		Cucujiformia	1.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.41	0.00
Terrestrial	Dipterans		Drosophila	4.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.63
Terrestrial	Dipterans		melanogaster group	4.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.95
Terrestrial	Dipterans		melanogaster subgroup	5.91	0.00	0.00	0.00	0.00	0.00	0.00	5.91	0.00	0.00	0.00	0.00
Terrestrial	Hymenopterans		Parasitoida	2.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.92
Terrestrial	Hemipterans	Aphids and relatives	Sternorrhyncha	1.93	0.00	1.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Animals	Animals	Metazoa	1.74	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.74
Mixed	Animals		Eumetazoa	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.98	0.00	0.00	0.00
Mixed	Animals	Deuterostomes	Deuterostomia	0.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.58
Mixed	Mammals	Mammals	Mammalia	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31
Mixed	Mammals	Carnivorans	Carnivora	0.33	0.00	0.00	0.00	0.00	0.00	0.33	0.00	0.00	0.00	0.00	0.00
Mixed	Vertebrates	Diapsids	Sauria	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.60
Mixed	Vertebrates		Archelosauria	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.68

Continued on next page

Table 4.42 continued

Ecological category	Group	Common name	Taxon	Total	310	281	219	202	177	152	123	107	88	72	58*
			Ingroup	241.70	0.00	1.93	0.00	0.15	17.68	3.05	7.44	20.30	0.00	4.22	186.94
Mixed	Testudines		Durocryptodira	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35
Mixed	Squamates	Snakes and related lizards	Toxicofera	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.47
Mixed	Invertebrates		Protostomia	10.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.52	0.00	0.00	10.39
Mixed	Invertebrates		Ecdysozoa	3.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.00
Mixed	Nematodes	Nematode worms	Nematoda	5.92	0.00	0.00	0.00	0.00	5.92	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Arthropods	Arthropods	Arthropoda	22.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	22.92
Mixed	Mites		Astigmata	13.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	13.33
Mixed	Spiders	Spiders	Araneae	0.42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.42
Mixed	Spiders	Cross-fanged spiders	Araneomorphae	0.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.83
Mixed	Spiders		Dionycha	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35
Mixed	Arthropods		Pancrustacea	1.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.86
Mixed	Crustaceans		Pleocyemata	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15
Mixed	Crustaceans	Amphipods	Amphipoda	0.15	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mixed	Arthropods	Insects and relatives	Hexapoda	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91
Mixed	Insects	Winged insects	Pterygota	1.84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.84
Mixed	Insects		Neoptera	7.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.66
Mixed	Insects	Metamorphosing insects	Holometabola	39.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	39.54
Mixed	Insects		Amphiesmenoptera	3.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.58	0.00	0.00	1.58
Mixed	Coleopterans		Polyphaga	1.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.31
Mixed	Dipterans	True flies	Diptera	2.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.67
Mixed	Dipterans	Crane flies and relatives	Tipulidae	2.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.30	0.00
Mixed	Hymenopterans	Stinging wasps, ants and bees	Aculeata	2.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.33

Embryophyta

Aquatics

Aquatics are typically discussed first because they are more frequent than terrestrial taxa in the majority of samples, but this core mostly returned trees. Nevertheless, aquatics can be found in most samples with data. They include *Zostera*, although only in three samples and low quantity. The reed group and a small range of freshwater taxa are present in all samples but 202. Aquatics are not lacking, despite having proportionally low biogenomic masses.

Halophytes

Several higher samples contain a halophyte read, which suggests some coastal influence to support the occasional saltwater aquatic.

Trees/shrubs

The dominant ecological category for all but sample 58 is trees/shrubs. The willow group are most frequent, with a biogenomic mass of over 5,700 in sample 72, but *Prunus* is also significant. A third notable tree is Juglandaceae (the walnut family), although it is limited to two reads (of six) in sample 281. Juglandaceae also occurred in core ELF059A and was interpreted as a biogeographic oddity. *J. regia* (the only likely species) had colonised much of Europe in previous interglacial periods, but in the Holocene was native to warmer regions until spread by cultivation (de Rigo *et al.* 2016).

A final woody taxon is Dryadoideae. Besides one read in ELF049_295, it is limited to samples 107-72 in this core, and is proportionally frequent enough to appear on the Krona chart for 72 (figure 4.101). The only European species is the alpine or tundra shrub *Dryas octopetala*, known for is prevalence during the eponymous Younger Dryas cool period, an interruption in the general warming trend between the Pleistocene and Holocene. Today, *D. octopetala* in Great Britain is mostly found in northern uplands, reaching down to coastal sands in Scotland (Walker 2015). As these cores are expected to be of Holocene age and the samples containing *D. octopetala* are towards the top, so likely to be the youngest, this is unlikely to represent the Younger Dryas itself. However, the reads may be evidence of a relict *D. octopetala* in what was rapidly becoming warmer lowlands; individual plants are long-lived and can persist despite a warming climate (Walker 2015, de Witte and Stöcklin 2011).

Grasses and relatives

Apart from reeds, grasses are absent from the data-poor deeper samples and relatively few but consistent from sample 177 onward, before becoming dominant in sample 58. There is an unusually large range of taxa. The Panicoideae group are present again, but there is a more surprising set of crop-associated grasses in ELF031A: Triticodae and six child taxa. *Hordeum* (barleys), Hordeinae, and *Thinopyrum* (wheat grasses) contain British natives from a variety of habitats. However, *Secale* (ryes), *Aegilops* (goat grasses), and *Triticum monococcum* (einkorn wheat) are from southern Europe.

These taxa are difficult to take at face value. They are clearly vulnerable to over-assignment and related taxa are among the most frequent in negative controls. The taxa are very rare across the cores. ELF031A certainly has the highest concentration, but it has only 65 reads, which may fall into the small percentage expected to be mis-assigned by PIA. Age-authentication of cereals, separated into close wheat relatives and not, is attempted in Chapter 6 (Mesophilic taxa and human disturbance indicators) with mixed results, although a somewhat more positive signal for non-wheat taxa. Further investigation is required before we can confirm domesticated grasses in ELF031A.

Metazoa

While the additional sequence data for sample 58 failed to match the number of Embryophyta reads in sample 72, it did result in the greatest Metazoa read count across all cores. The majority of taxa in ELF031A occur only in sample 58; other read counts conform to typically low expectations.

Despite this, sample 107 contains an interesting signal: the order of freshwater insects Trichoptera (caddisflies) and several of its child taxa. There are small Trichoptera signals in sample 58 and two other cores (ELF032A and ELF051), but most reads are in 107. Trichoptera are very rare across all cores, yet concentrated in this one sample. However, Trichoptera occupy a variety of freshwater habitats, so this does not suggest a particular environment in 108.

Otherwise, samples other than 58 show a mixture of mixed aquatic and terrestrial taxa, supporting the small aquatic signal in Embryophyta. Most biogenomic mass in 58 is assigned to the two higher taxa Holometabola (metamorphosing insects) and its parent Arthropoda, resembling the tendency of high-yield Embryophyta samples to return higher taxa.

Another taxon of note in sample 58 is Durocryptodira, a large group of turtles and tortoises. The two most likely taxa are *Dermochelys* (the leatherback sea turtle) and *Emys orbicularis* (the European pond terrapin), which was discussed with core ELF032A (also containing Trichoptera). Considering the terrestrial nature of Embryophyta in ELF031A, *E. orbicularis* may be the most likely interpretation.

4.21.3 Pianka scores

The Embryophyta scores clearly show overlap between the deeper samples, all tree-dominated although by different trees and with low read counts inflating some discrepancies. The contrast between those samples and 58, with its suddenly substantial grass signal, is also apparent. The Metazoa scores are less helpful due to data scarcity.

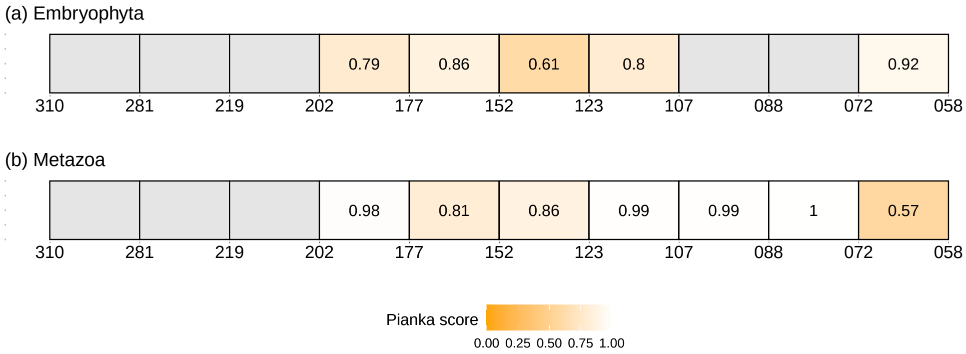


Figure 4.102: Pianka similarity scores between adjacent samples in ELF031A for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

4.21.4 Summary

ELF031A is an unusual core in many ways. It has an odd pattern of data yield which does not seem to correlate with taxonomic profile, unlike in previous cores where particularly well-performing samples could be explained by a saltwater environment. Indeed, there is hardly any saltwater signal in ELF031A, with almost all samples comprising mostly willow group trees or *Prunus*, followed by a very strong grass signal in the top sample, 58. The reed group is present but not particularly frequent in 58, so it may be better interpreted as a terrestrial grassland than the reed beds implied in previous cores.

There are a number of unusual plant taxa of varying plausibility. 108 has two reads from Juglandaceae but very little data otherwise. The upper samples besides 58 contain Dryadoideae, which could indicate cold conditions or be from a relict population. Finally, those upper samples also contain grasses in Triticodae associated with crops, although this signal is as yet unconvincing for such over-represented and archaeologically-loaded taxa.

The Metazoa data, as is often the case, was much smaller in quantity and generally less useful. However, sample 108 has an unexplained surfeit of Trichoptera, and sample 58 (the Metazoa data of which benefited greatly from the additional sequencing) contains a higher testudine taxon that may be best attributed to *Emys orbicularis*, a warm-climate terrapin.

4.22 ELF051

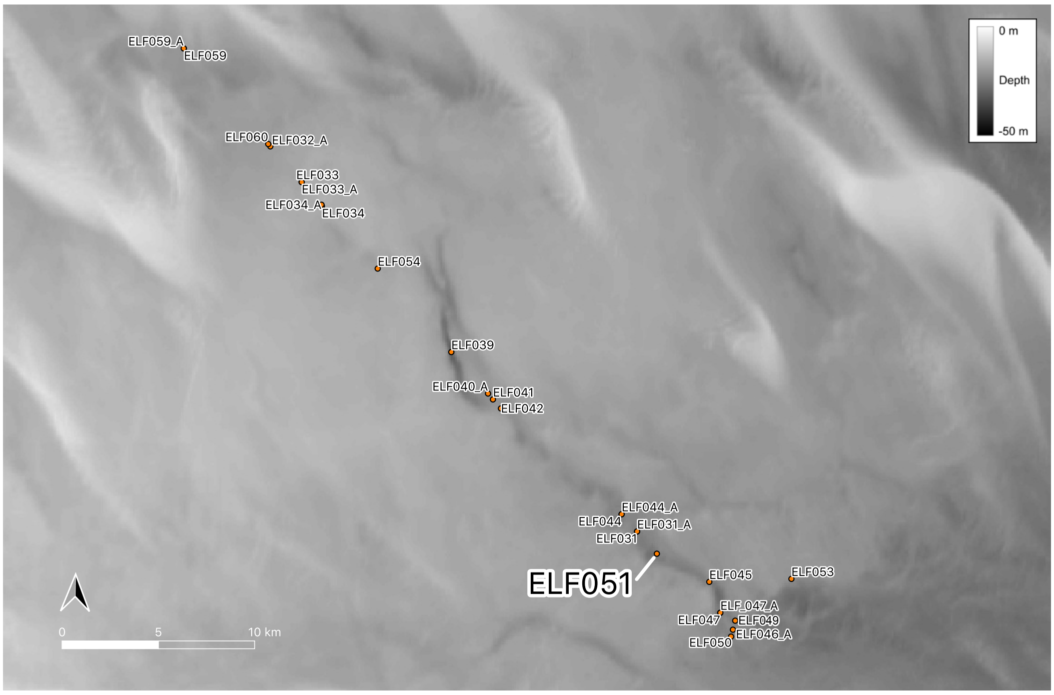


Figure 4.103: Bathymetry map of the palaeochannel transect highlighting ELF051. Darker grey indicates greater depth.

ELF051 is approximately 5 km from the mouth of the visible palaeochannel. There are no other cores nearby.

4.22.1 Read counts

Unusually, read counts in ELF051 broadly increase with depth. The top three samples have very little data. The largest Embryophyta count is for the predominantly saltwater sample, as expected, but this sample is also the deepest. We would expect inundation to result in saltwater samples at the tops of cores. This suggests that the core may be inverted, or may have captured a temporary reversal of inundation.

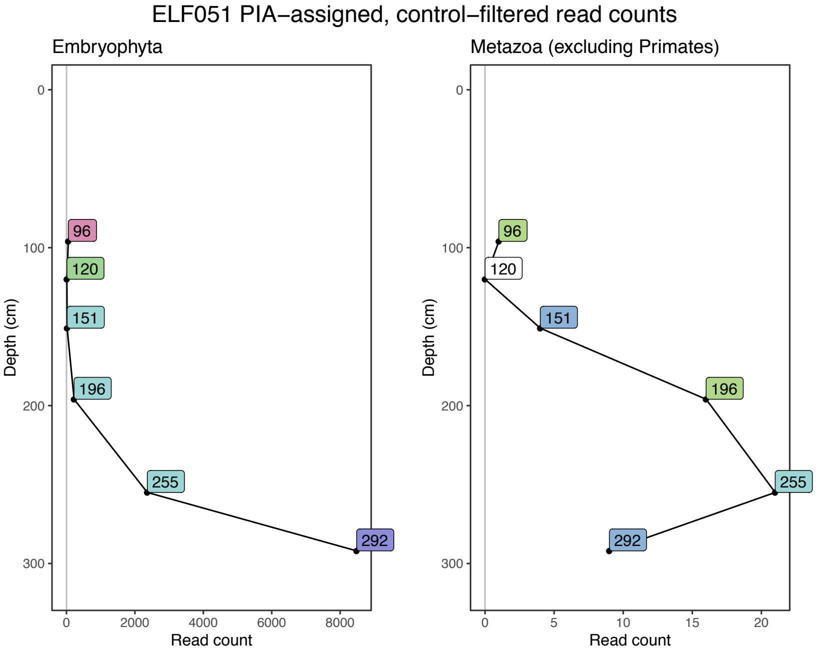


Figure 4.104: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF033. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.22.2 Taxonomic profiles

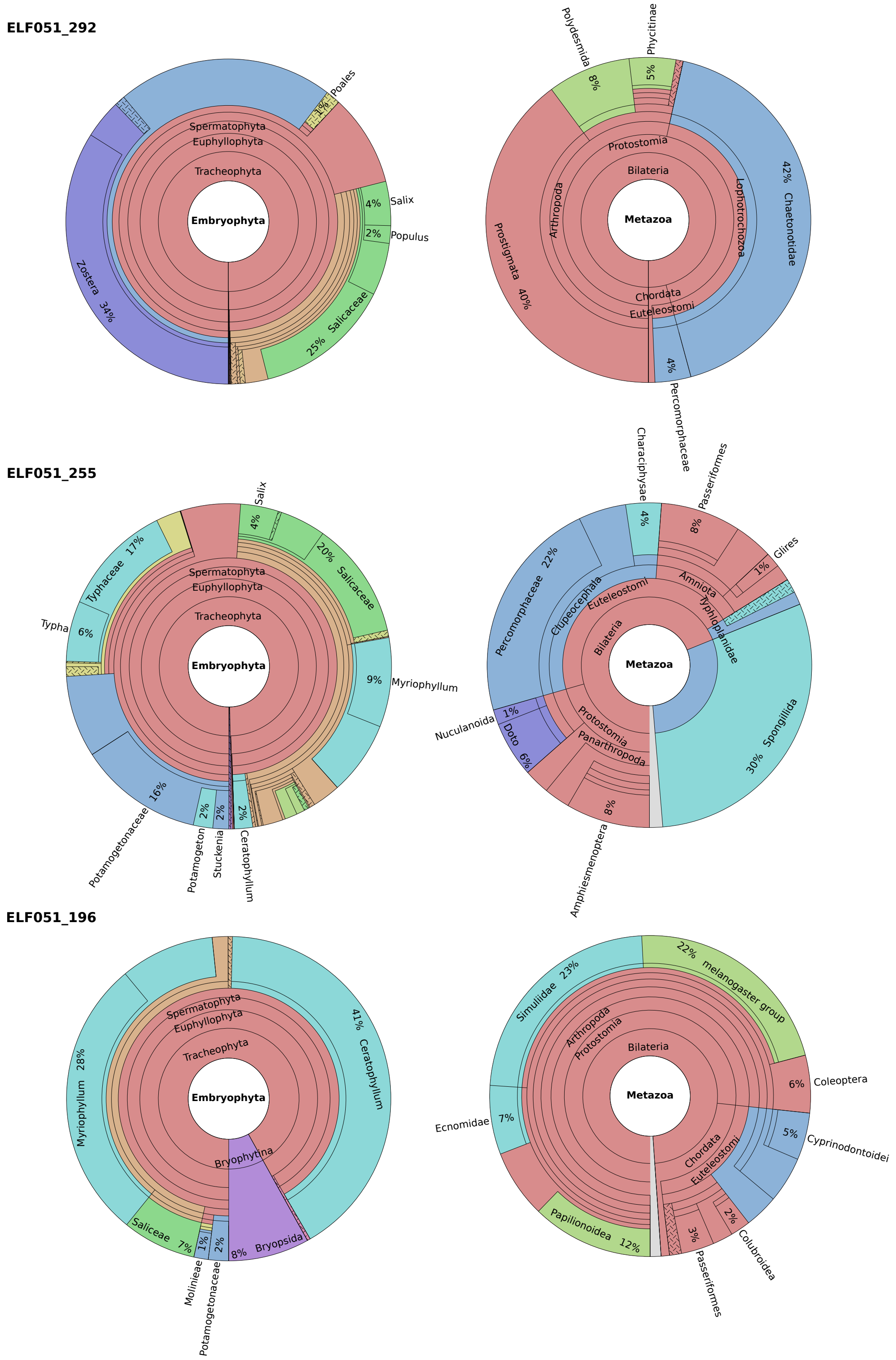
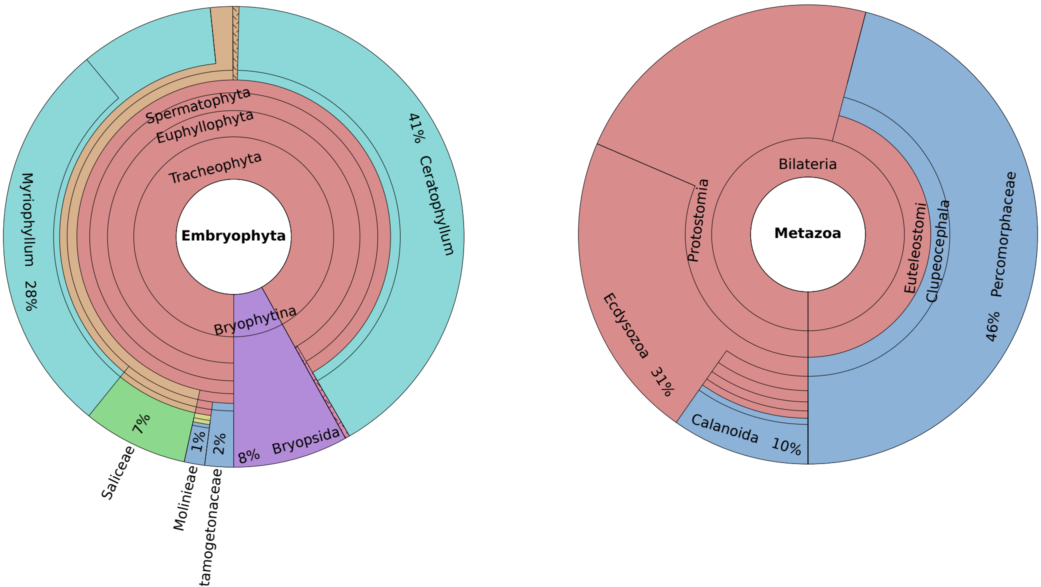
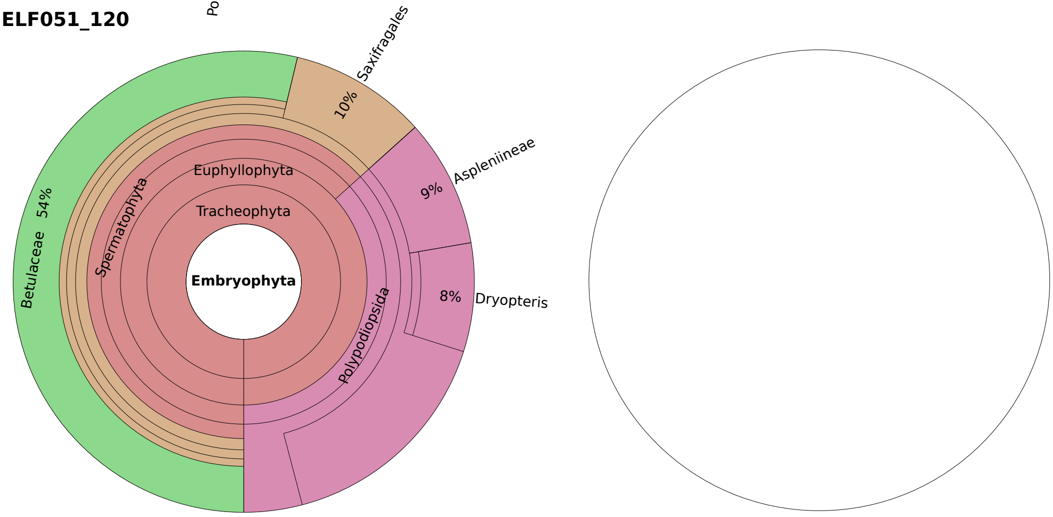


Figure 4.105: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF051, samples 292, 255, and 196. Continued in figure 4.106. See figure 4.4 for colour key.

ELF051_151



ELF051_120



ELF051_096

Figure 4.106: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF051, samples 151, 120, and 96. See figure 4.4 for colour key.

Table 4.43: Biogenomic masses of European taxa in Embryophyta from ELF051. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	292	255	196	151	120	96
		Ingroup	4638.74	3274.75	1200.76	131.64	24.59	2.01	4.99
Freshwater aquatics	Coontails	Ceratophyllum	86.96	0.00	21.74	55.07	10.14	0.00	0.00
Freshwater aquatics	Gipsyworts	Lycopus	1.60	0.00	1.60	0.00	0.00	0.00	0.00
Freshwater aquatics	Water-milfoil family	Haloragaceae	101.54	6.92	87.69	4.62	2.31	0.00	0.00
Freshwater aquatics	Water-milfoils	Myriophyllum	122.31	2.31	106.15	6.92	6.92	0.00	0.00
Freshwater aquatics	Water-plantain family	Alismataceae	0.14	0.00	0.14	0.00	0.00	0.00	0.00
Freshwater aquatics	Elodea family	Hydrocharitaceae	0.20	0.20	0.00	0.00	0.00	0.00	0.00
Freshwater aquatics	Pondweeds	Potamogeton	27.49	1.62	23.18	2.70	0.00	0.00	0.00
Freshwater aquatics	Bulrush family	Typhaceae	133.98	0.00	133.98	0.00	0.00	0.00	0.00
Freshwater aquatics	Bulrushes	Typha	75.47	0.00	71.70	3.77	0.00	0.00	0.00
Freshwater aquatics	Quillworts	Isoetes	0.15	0.00	0.00	0.15	0.00	0.00	0.00
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	2.09	2.09	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Tasselweeds	Ruppia	3.43	3.43	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	**Neptune-grasses**	**Posidonia**	**0.32**	**0.32**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Salt/brackish aquatics	Eelgrass family	Zosteraceae	126.61	126.61	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Eelgrasses	Zostera	1110.09	1110.09	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Alismatids	Alismatales	812.24	704.79	97.10	10.35	0.00	0.00	0.00
Mixed aquatics	Pondweed family	Potamogetonaceae	198.84	24.30	147.77	26.28	0.50	0.00	0.00
Mixed aquatics	Pondweeds	Stuckenia	23.31	2.48	18.84	1.98	0.00	0.00	0.00
Mixed aquatics	Reeds	Arundinoideae	1.77	1.77	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Common reed tribe	Molinieae	5.64	3.88	0.71	0.71	0.35	0.00	0.00
Halophytes	Sea-lavenders	Limonium	0.21	0.21	0.00	0.00	0.00	0.00	0.00
Halophytes	Scurvy-grasses	Cochlearia	1.42	1.42	0.00	0.00	0.00	0.00	0.00
Xerophytes	**Joint-pine**	**Ephedra**	**0.07**	**0.00**	**0.07**	**0.00**	**0.00**	**0.00**	**0.00**
Xerophytes	Goosefoot family	Chenopodiaceae	0.80	0.80	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Conifers	Pinidae	0.06	0.00	0.00	0.06	0.00	0.00	0.00
Trees/shrubs		Viburnum	0.26	0.26	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Birch family	Betulaceae	5.40	0.00	3.24	1.08	0.00	1.08	0.00
Trees/shrubs	Alders	Alnus	1.61	1.61	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Willow family	Salicaceae	592.06	447.65	138.99	5.42	0.00	0.00	0.00
Trees/shrubs	Willow tribe	Saliceae	226.93	170.65	54.46	0.00	1.82	0.00	0.00
Trees/shrubs	Poplars	Populus	61.32	57.36	3.96	0.00	0.00	0.00	0.00
Trees/shrubs	Willows	Salix	188.58	143.60	44.98	0.00	0.00	0.00	0.00
Trees/shrubs	Stone fruit trees	Prunus	3.65	3.65	0.00	0.00	0.00	0.00	0.00
Trees/shrubs	Elm family	Ulmaceae	0.50	0.50	0.00	0.00	0.00	0.00	0.00
Herbs	Water-parsnip tribe	Oenantheae	0.79	0.00	0.79	0.00	0.00	0.00	0.00
Herbs	Bellflower family	Campanulaceae	0.47	0.47	0.00	0.00	0.00	0.00	0.00
Herbs	Madder tribe	Rubieae	0.71	0.00	0.71	0.00	0.00	0.00	0.00
Herbs	Catmint subfamily	Nepetoideae	14.68	0.00	14.68	0.00	0.00	0.00	0.00
Herbs	Mint tribe	Mentheae	9.82	0.00	9.82	0.00	0.00	0.00	0.00
Herbs	Mints	Mentha	3.05	0.00	3.05	0.00	0.00	0.00	0.00
Herbs	Sages	Salvia	1.31	0.00	1.31	0.00	0.00	0.00	0.00
Herbs	Nightshades	Solanum	1.43	0.00	1.43	0.00	0.00	0.00	0.00
Herbs	Pink family	Caryophyllaceae	0.68	0.68	0.00	0.00	0.00	0.00	0.00
Herbs	Knotweed subfamily	Polygonoideae	1.12	1.12	0.00	0.00	0.00	0.00	0.00

Continued on next page

Table 4.43 continued

Ecological category	Common name	Taxon	Total	292	255	196	151	120	96
		Ingroup	4638.74	3274.75	1200.76	131.64	24.59	2.01	4.99
Herbs	Rhubarb family	Rumiceae	0.83	0.41	0.41	0.00	0.00	0.00	0.00
Herbs		Hologalegina	0.56	0.00	0.56	0.00	0.00	0.00	0.00
Herbs		Fabeae	0.74	0.00	0.74	0.00	0.00	0.00	0.00
Herbs	Vetchlings	Lathyrus	0.13	0.00	0.13	0.00	0.00	0.00	0.00
Herbs		Galegeae	0.67	0.67	0.00	0.00	0.00	0.00	0.00
Herbs	Cinquefoils	Potentilla	2.83	1.42	1.42	0.00	0.00	0.00	0.00
Herbs	Geranium family	Geraniaceae	0.81	0.00	0.00	0.81	0.00	0.00	0.00
Herbs	Crane's-bills	Geranium	0.57	0.00	0.00	0.57	0.00	0.00	0.00
Herbs	Saxifrages	Saxifraga	0.69	0.69	0.00	0.00	0.00	0.00	0.00
Herbs	Arum family	Araceae	0.55	0.37	0.00	0.18	0.00	0.00	0.00
Herbs	Bluebell family	Hyacinthaceae	0.08	0.08	0.00	0.00	0.00	0.00	0.00
Herbs	Orchids	Orchidaceae	0.17	0.00	0.00	0.17	0.00	0.00	0.00
Herbs		Epidendroideae	0.32	0.00	0.32	0.00	0.00	0.00	0.00
Herbs	Buttercup subfamily	Ranunculoideae	0.54	0.11	0.11	0.22	0.11	0.00	0.00
Herbs	Buttercup tribe	Ranunculeae	0.63	0.00	0.16	0.47	0.00	0.00	0.00
Herbs	Buttercups	Ranunculus	0.30	0.00	0.15	0.15	0.00	0.00	0.00
Herbs		Thalictroideae	4.82	1.93	2.89	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Gymnosperms	Acrogymnospermae	0.05	0.00	0.00	0.00	0.00	0.00	0.05
Trees/shrubs and herbs		Gunneridae	3.30	0.47	2.35	0.47	0.00	0.00	0.00
Trees/shrubs and herbs		campanulids	1.23	0.31	0.92	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Asterales	2.57	1.71	0.86	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	1.63	0.54	1.09	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Asteroideae	1.50	0.00	1.50	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	0.18	0.18	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Aster tribe	Astereae	0.52	0.52	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Ragwort tribe	Senecioneae	1.15	0.00	1.15	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		lamiids	5.48	1.37	4.11	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Gentianales	0.90	0.00	0.90	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	23.71	0.00	23.71	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Dead-nettle family	Lamiaceae	3.14	0.00	3.14	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Knotweed family	Polygonaceae	0.55	0.00	0.55	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Malpighiales	73.83	73.83	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Rose subfamily	Rosoideae	9.74	8.35	1.39	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	8.48	7.18	1.31	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	1.11	0.56	0.00	0.56	0.00	0.00	0.00
Trees/shrubs and herbs		Myrtales	0.99	0.00	0.00	0.99	0.00	0.00	0.00
Trees/shrubs and herbs		Saxifragales	39.10	2.12	33.32	3.08	0.39	0.19	0.00
Trees/shrubs and herbs		Asparagales	0.32	0.08	0.24	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	**Palms**	**Arecaceae**	**0.29**	**0.00**	**0.29**	**0.00**	**0.00**	**0.00**	**0.00**
Trees/shrubs and herbs		Liliales	0.03	0.03	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.16	0.16	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Buttercup family	Ranunculaceae	0.34	0.34	0.00	0.00	0.00	0.00	0.00
Grasses and relatives	Grass order	Poales	32.21	2.34	29.28	0.59	0.00	0.00	0.00
Grasses and relatives	Sedges	Cyperaceae	7.71	3.08	4.63	0.00	0.00	0.00	0.00
Grasses and relatives		Cyperoideae	10.12	10.12	0.00	0.00	0.00	0.00	0.00
Grasses and relatives	True sedges	Carex	34.17	22.78	11.39	0.00	0.00	0.00	0.00
Grasses and relatives	Rushes	Juncaceae	1.04	1.04	0.00	0.00	0.00	0.00	0.00
Grasses and relatives		Chloridoideae	0.48	0.48	0.00	0.00	0.00	0.00	0.00
Grasses and relatives		**Panicoideae**	**2.70**	**1.08**	**1.08**	**0.54**	**0.00**	**0.00**	**0.00**

Continued on next page

Table 4.43 continued

Ecological category	Common name	Taxon	Total	292	255	196	151	120	96
		Ingroup	4638.74	3274.75	1200.76	131.64	24.59	2.01	4.99
Grasses and relatives		**Paniceae**	**0.67**	**0.67**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Grasses and relatives	**Panicgrass**	**Panicum**	**0.64**	**0.64**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Ferns	Ferns	Polypodiopsida	0.07	0.00	0.00	0.00	0.00	0.00	0.07
Ferns	Horsetails	Equisetum	0.09	0.05	0.05	0.00	0.00	0.00	0.00
Ferns		Ophioglossidae	0.06	0.00	0.00	0.00	0.00	0.00	0.06
Ferns	Leptosporangiate ferns	Polypodiidae	0.49	0.00	0.00	0.00	0.00	0.08	0.41
Ferns	Polypod ferns	Polypodiales	2.41	0.64	0.00	0.00	0.00	0.32	1.45
Ferns	Spleenwort suborder	Aspleniineae	1.62	1.08	0.00	0.27	0.00	0.18	0.09
Ferns		Athyriaceae	0.07	0.07	0.00	0.00	0.00	0.00	0.00
Ferns		Thelypteridaceae	0.10	0.10	0.00	0.00	0.00	0.00	0.00
Ferns	Bracken family	Dennstaedtiaceae	0.10	0.10	0.00	0.00	0.00	0.00	0.00
Ferns		Dryopteridoideae	0.16	0.00	0.00	0.00	0.00	0.00	0.16
Ferns	Buckler and male ferns	Dryopteris	1.83	0.00	0.00	0.08	0.08	0.15	1.53
Bryophytes		Bryopsida	3.97	0.00	1.98	0.00	1.98	0.00	0.00
Bryophytes	Liverworts	Marchantiophyta	0.53	0.00	0.53	0.00	0.00	0.00	0.00
Bryophytes		Marchantiidae	1.50	0.00	1.50	0.00	0.00	0.00	0.00
Mixed		Tracheophyta	4.50	2.25	1.56	0.52	0.00	0.00	0.17
Mixed		Euphyllophyta	3.44	1.72	0.86	0.86	0.00	0.00	0.00
Mixed	Monocots	Liliopsida	369.70	298.65	71.04	0.00	0.00	0.00	0.00

Table 4.44: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF051. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	292	255	196	151	120	96
			Ingroup	76.19	31.26	18.90	23.42	2.30	0.00	0.32
Freshwater aquatics	**Fishes**		**Characiphysae**	**0.67**	**0.00**	**0.67**	**0.00**	**0.00**	**0.00**	**0.00**
Freshwater aquatics	Flatworms		Mesostominae	0.27	0.00	0.27	0.00	0.00	0.00	0.00
Freshwater aquatics	Trichopterans		Ecnomidae	1.58	0.00	0.00	1.58	0.00	0.00	0.00
Freshwater aquatics	Dipterans	Blackflies	Simuliidae	5.26	0.00	0.00	5.26	0.00	0.00	0.00
Freshwater aquatics	Poriferans	Freshwater sponges	Spongillida	5.63	0.00	5.63	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bivalves		Nuculanoida	0.28	0.00	0.28	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Gastropods		Doto	1.05	0.00	1.05	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes	Ray-finned fishes	Actinopterygii	0.78	0.00	0.00	0.78	0.00	0.00	0.00
Mixed aquatics	Fishes		Euteleosteomorpha	0.90	0.00	0.90	0.00	0.00	0.00	0.00
Mixed aquatics	Fishes		Acanthomorphata	1.06	0.00	0.00	1.06	0.00	0.00	0.00
Mixed aquatics	Fishes		Percomorphaceae	6.34	1.06	4.22	0.00	1.06	0.00	0.00
Mixed aquatics	**Fishes**		**Cyprinodontoidei**	**1.08**	**0.00**	**0.00**	**1.08**	**0.00**	**0.00**	**0.00**
Mixed aquatics	Flatworms		Typhloplanidae	0.25	0.00	0.25	0.00	0.00	0.00	0.00
Mixed aquatics	Copepods		Calanoida	0.22	0.00	0.00	0.00	0.22	0.00	0.00
Mixed aquatics	Crustaceans		Gammaridea	0.03	0.03	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Crustaceans		Gammaroidea	0.18	0.18	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Gastrotrichs		Chaetonotidae	12.50	12.50	0.00	0.00	0.00	0.00	0.00
Terrestrial	Mammals	Rodents	Rodentia	0.28	0.00	0.00	0.28	0.00	0.00	0.00
Terrestrial	Mammals	Pigs	Sus	0.32	0.00	0.00	0.00	0.00	0.00	0.32
Terrestrial	Arthropods	Flat-backed millipedes	Polydesmida	2.44	2.44	0.00	0.00	0.00	0.00	0.00
Terrestrial	Butterflies	Butterflies	Papilionoidea	2.79	0.00	0.00	2.79	0.00	0.00	0.00
Terrestrial	Moths		Phycitinae	1.37	1.37	0.00	0.00	0.00	0.00	0.00
Terrestrial	Dipterans		melanogaster group	4.95	0.00	0.00	4.95	0.00	0.00	0.00
Mixed	Animals	Animals	Metazoa	0.50	0.00	0.25	0.25	0.00	0.00	0.00
Mixed	Animals	Chordates	Chordata	0.38	0.19	0.00	0.19	0.00	0.00	0.00
Mixed	Mammals	Mammals	Mammalia	0.31	0.00	0.31	0.00	0.00	0.00	0.00
Mixed	Mammals	Rodents and lagomorphs	Glires	0.28	0.00	0.28	0.00	0.00	0.00	0.00
Mixed	Vertebrates	Archosaurs	Archosauria	0.73	0.00	0.73	0.00	0.00	0.00	0.00
Mixed	Birds	Perching birds	Passeriformes	2.27	0.00	1.51	0.76	0.00	0.00	0.00
Mixed	Squamates	Snakes and related lizards	Toxicofera	0.47	0.00	0.00	0.47	0.00	0.00	0.00
Mixed	Squamates	Colubrid snake superfamily	Colubroidea	0.46	0.00	0.00	0.46	0.00	0.00	0.00
Mixed	Invertebrates		Protostomia	0.52	0.00	0.00	0.00	0.52	0.00	0.00
Mixed	Invertebrates		Ecdysozoa	0.50	0.00	0.00	0.00	0.50	0.00	0.00
Mixed	Invertebrates		Panarthropoda	0.48	0.00	0.48	0.00	0.00	0.00	0.00
Mixed	Arthropods	Arthropods	Arthropoda	0.49	0.00	0.49	0.00	0.00	0.00	0.00
Mixed	Mites	Sucking mites	Prostigmata	11.76	11.76	0.00	0.00	0.00	0.00	0.00
Mixed	Insects		Amphiesmenoptera	1.58	0.00	1.58	0.00	0.00	0.00	0.00
Mixed	Lepidopterans		Obtectomera	1.56	0.00	0.00	1.56	0.00	0.00	0.00
Mixed	Coleopterans	Beetles	Coleoptera	1.32	0.00	0.00	1.32	0.00	0.00	0.00

Embryophyta

Aquatics

Aquatics of some kind are the most frequent ecological category in the four lower samples but absent from the top two. The deepest sample, 292, has a typical high-yield, *Zostera*-dominated pattern with most reads to *Zostera*, Alismatales, and Liliopsida. However, saltwater aquatics are not found elsewhere. Samples 255-151 instead mostly contain freshwater taxa, which is highly unusual across these cores. *Myriophyllum* (water-milfoils) and its family Haloragaceae are most consistent, but *Typha* and Typhaceae are most frequent in 255 and *Ceratophyllum*(coontails) in 196 and 151.

Mixed aquatics besides Alismatales are present at lower frequency in the deeper samples; their absence from 151 may be an artefact of its low read count (only 15). This includes a very small signal from the reed group.

Terrestrial

Halophytes are only found in sample 292. The two xerophyte taxa, Chenopodiaceae and the gymnosperm shrub *Ephedra* (joint-pine), can also be associated with coastal habitats (Huang *et al.* 2005, Schneider-Binder and Kuhlke 2015), which in a lowland temperate region are more likely than other water-stressed environments. These are split between 292 and the next deepest sample, 255. This supports the restriction of coastal habitats to the bottom of the core.

All but the top sample have a clear, if proportionally variable, signal from woody taxa. Most biogenomic mass is from the willow group. Sample 196 also has a single read from the rarely-seen Pinidae (pine family; despite its common name, *Ephedra* is in a different family). The other usually frequent group, grasses, are present in the lower three samples but with low biogenomic masses. They may be related to the small reed signal.

Finally, ferns often go unmentioned because they are rarely ecologically informative and usually occur at very low frequency, but ferns occupy a substantial proportion of the biogenomic mass in samples 120 and 96. A dominant fern signal has only been seen in one other sample: ELF040A_192. It is unclear what environment this could represent.

Metazoa

Metazoa results for the four deeper samples are consistent with the mostly aquatic Embryophyta profiles, and samples 255 and 196 share an unusually strong freshwater signal. However, the upper three samples contain very little data, and the one read in 96 (*Sus*) is likely to have been mis-assigned. They do not add any further information to the odd fern situation.

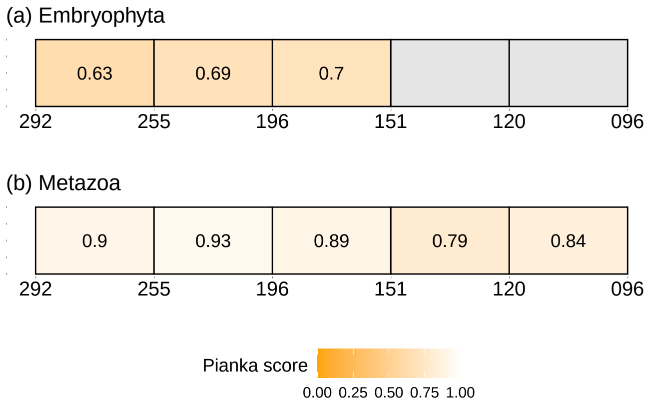


Figure 4.107: Pianka similarity scores between adjacent samples in ELF051 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

4.22.3 Pianka scores

The Embryophyta Pianka scores are surprisingly high at first glance, but considering the inflation that typically occurs in samples with more data, a score of 0.81 between the *Zostera*-dominated and first primarily freshwater sample does convey some of their dissimilarity. The three freshwater samples are found to be more similar, before a transition to the two samples unusually rich in ferns. These two samples have relatively low Pianka scores, but this can be attributed to their very low read counts.

Similarly, the four samples where Metazoa comparisons were possible appear very dissimilar, but this could reflect a lack of taxon sampling as much as ecological differences.

4.22.4 Summary

ELF051 appears inverted. The deepest sample is clearly saltwater, probably an estuarine environment rather than saltmarsh due to the lack of reeds or grasses, but subsequent samples appear freshwater followed by terrestrial. The freshwater signal is particularly unusual, and it is derived more from aquatic herbs than reeds or rushes that could indicate marsh succession. The sedimentary sequence may have been inverted. Alternatively, perhaps the location started some distance away from the mouth of the channel, far enough for the freshwater signal to be swamped by estuarine, but the channel then widened or wandered over it. Or, there may have been a temporary reversal of inundation, possibly due to changes in local topography. It is certainly an unusual core, and it is unfortunate that the upper samples are lacking in data.

4.23 ELF045

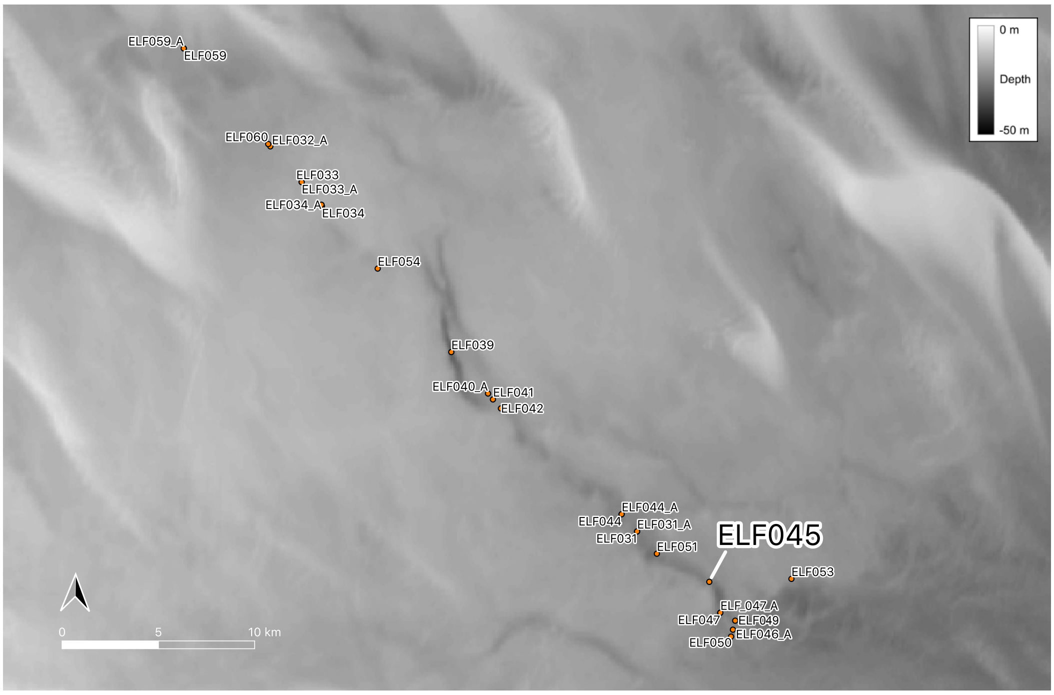


Figure 4.108: Bathymetry map of the palaeochannel transect highlighting ELF045. Darker grey indicates greater depth.

ELF045 is the final core in the transect before the mouth of the palaeochannel, which begins approximately 2 km southeast.

4.23.1 Read counts

The data yield is generally good across the core, with the lowest Embryophyta read count at a respectable 512. Sample 145 performed particularly well, returning nearly 30,000 Embryophyta reads.

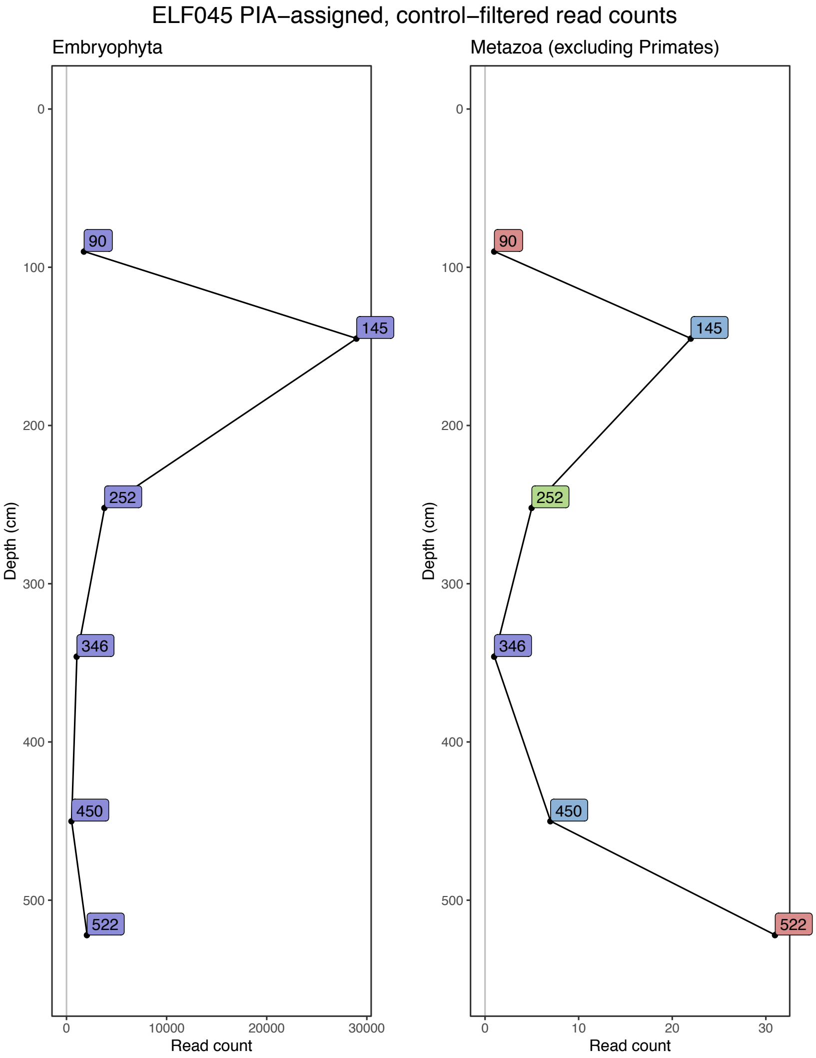
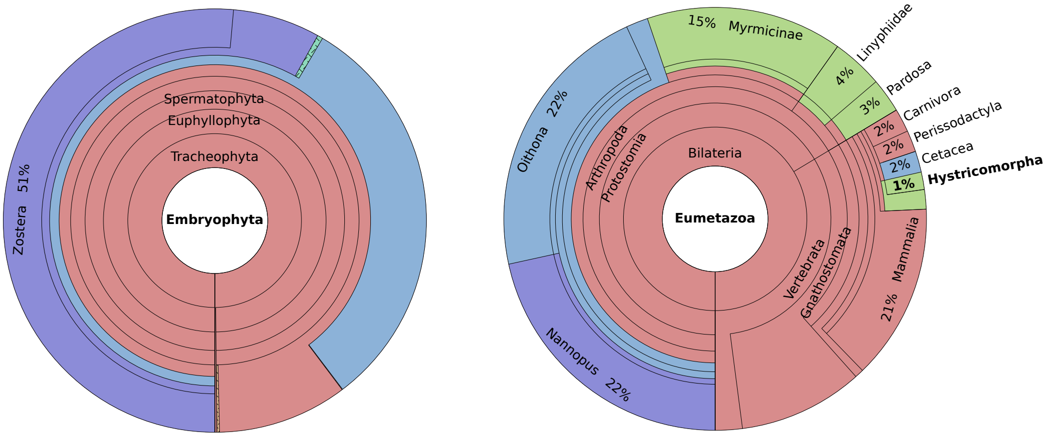


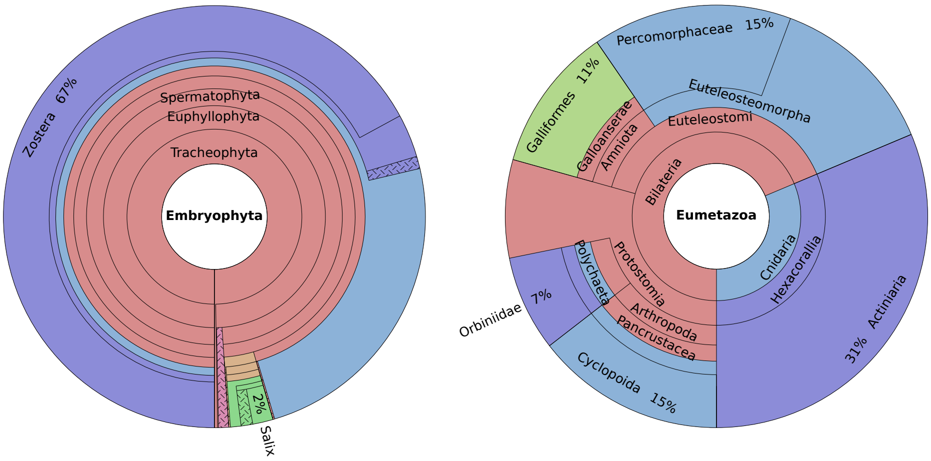
Figure 4.109: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF045. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.23.2 Taxonomic profiles

ELF045_522



ELF045_450



ELF045_346

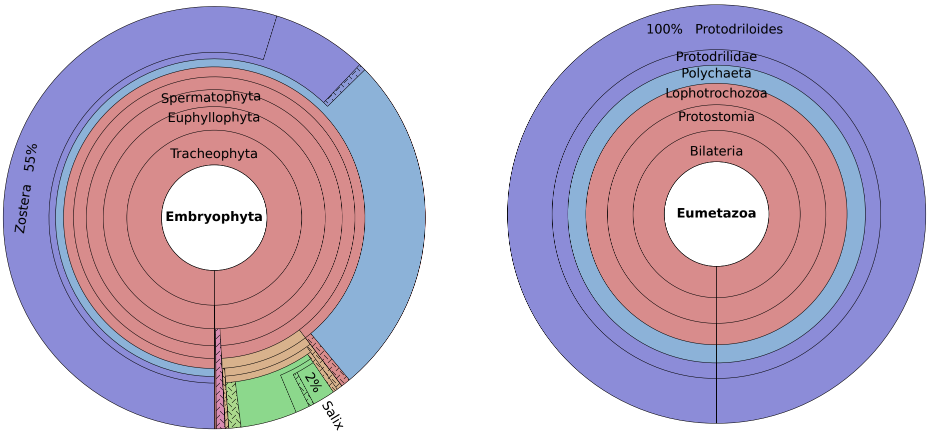
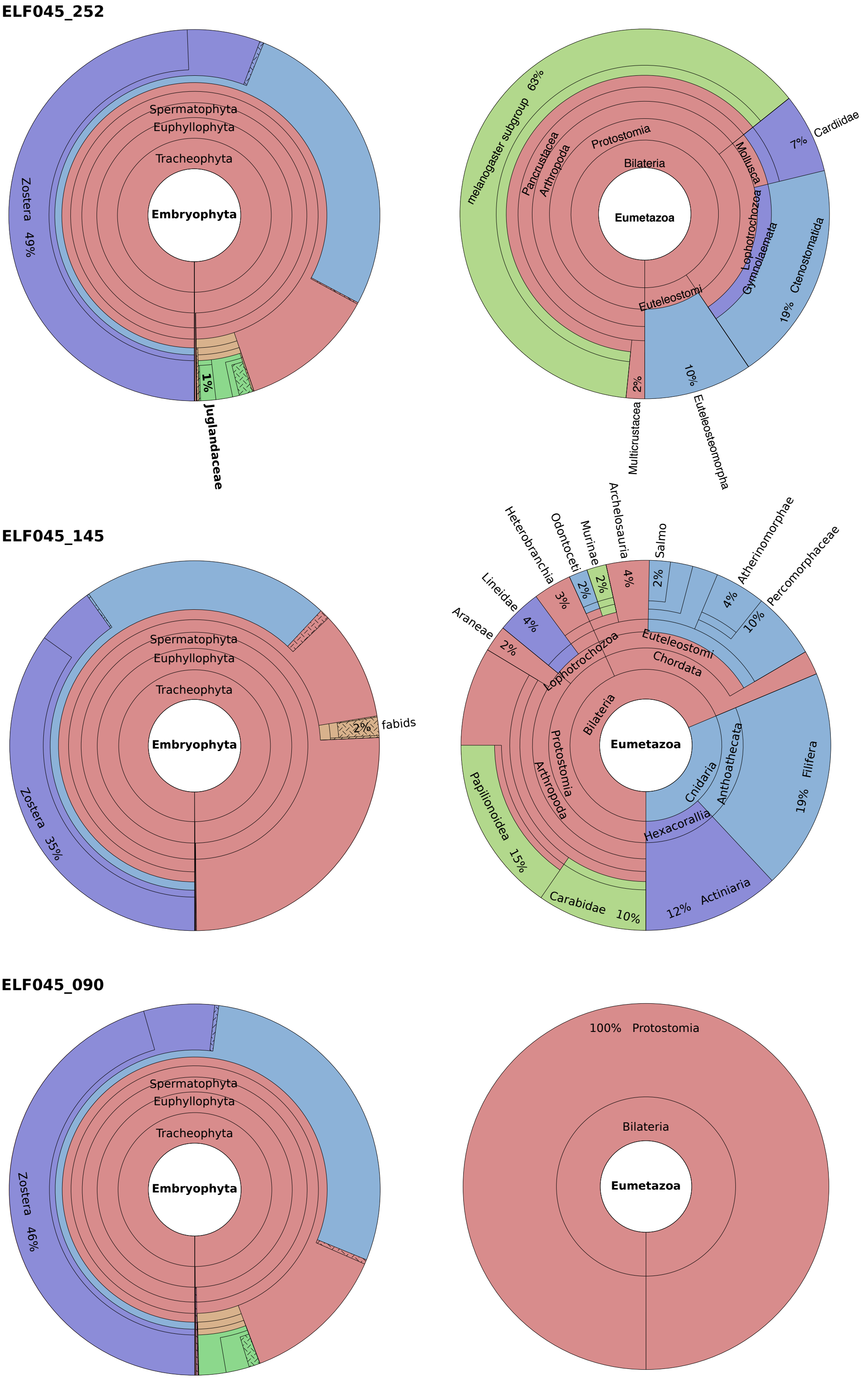


Figure 4.110: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF033, samples 522, 450, and 346. Continued in figure 4.111. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.



ELF045_252

ELF045_145

ELF045_090

Figure 4.111: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF045, samples 252, 145, and 90. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

Table 4.45: Biogenomic masses of European taxa in Embryophyta from ELF045. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	522	450	346	252	145	90
		Ingroup	11001.54	660.25	248.58	506.25	1243.70	7804.75	538.01
Freshwater aquatics	Bladderworts	Utricularia	4.30	0.00	0.00	0.00	0.00	4.30	0.00
Freshwater aquatics	Water-plantain family	Alismataceae	0.07	0.00	0.00	0.00	0.00	0.07	0.00
Freshwater aquatics	Duckweed family	Lemnoideae	2.44	1.22	0.00	0.00	0.00	1.22	0.00
Freshwater aquatics	Elodea family	Hydrocharitaceae	0.40	0.00	0.00	0.00	0.00	0.40	0.00
Freshwater aquatics	Pondweeds	Potamogeton	1.62	0.54	0.00	0.54	0.00	0.54	0.00
Freshwater aquatics	Spike sedges	Eleocharis	0.33	0.00	0.00	0.00	0.00	0.33	0.00
Freshwater aquatics	Bulrush family	Typhaceae	2.91	0.00	0.00	0.00	0.00	2.91	0.00
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	9.76	0.00	2.09	0.00	2.09	4.18	1.39
Salt/brackish aquatics	Tasselweeds	Ruppia	5.58	0.00	0.00	1.72	0.43	3.43	0.00
Salt/brackish aquatics	**Neptune-grasses**	**Posidonia**	**0.32**	**0.00**	**0.00**	**0.00**	**0.00**	**0.32**	**0.00**
Salt/brackish aquatics	Eelgrass family	Zosteraceae	605.50	44.04	7.34	36.70	78.90	405.50	33.03
Salt/brackish aquatics	Eelgrasses	Zostera	4343.12	339.45	148.62	275.23	612.84	2722.94	244.04
Mixed aquatics	Alismatids	Alismatales	2568.45	205.90	53.30	132.06	331.35	1690.01	155.83
Mixed aquatics	Pondweed family	Potamogetonaceae	8.43	0.99	0.00	0.50	1.49	4.96	0.50
Mixed aquatics	Pondweeds	Stuckenia	0.50	0.00	0.00	0.00	0.50	0.00	0.00
Mixed aquatics	Reeds	Arundinoideae	1.06	0.00	0.00	0.00	0.00	0.71	0.35
Mixed aquatics	Common reed tribe	Molinieae	9.17	0.00	0.00	1.06	0.00	6.35	1.76
Halophytes	Sea-blite subfamily	Suaedoideae	0.94	0.00	0.00	0.00	0.00	0.94	0.00
Halophytes	Leadwort family	Plumbaginaceae	0.21	0.00	0.00	0.00	0.00	0.21	0.00
Halophytes	Sea-lavenders	Limonium	0.21	0.00	0.00	0.00	0.00	0.00	0.21
Xerophytes	Goosefoot family	Chenopodiaceae	2.39	0.00	0.00	0.00	0.80	1.59	0.00
Trees/shrubs	**Spruces**	**Picea**	**0.05**	**0.00**	**0.00**	**0.00**	**0.05**	**0.00**	**0.00**
Trees/shrubs		Viburnum	0.26	0.00	0.00	0.00	0.00	0.26	0.00
Trees/shrubs		Cornales	0.61	0.00	0.00	0.00	0.00	0.61	0.00
Trees/shrubs	Birch family	Betulaceae	7.56	0.00	0.00	0.00	0.00	7.56	0.00
Trees/shrubs	Alders	Alnus	3.23	0.00	0.00	0.00	1.61	1.61	0.00
Trees/shrubs	Hazels	Corylus	9.09	0.00	0.00	2.27	0.00	6.82	0.00
Trees/shrubs	Oaks	Quercus	2.12	0.00	0.00	1.06	0.00	1.06	0.00
Trees/shrubs	**Walnut family**	**Juglandaceae**	**17.75**	**0.00**	**0.00**	**1.37**	**16.38**	**0.00**	**0.00**
Trees/shrubs	Willow family	Salicaceae	113.72	0.00	1.81	21.66	18.05	59.57	12.64
Trees/shrubs	Willow tribe	Saliceae	45.39	0.00	0.00	5.45	7.26	21.79	10.89
Trees/shrubs	Poplars	Populus	11.87	0.00	1.98	1.98	5.93	1.98	0.00
Trees/shrubs	Willows	Salix	34.60	0.00	3.46	8.65	8.65	8.65	5.19
Trees/shrubs	Stone fruit trees	Prunus	10.96	0.00	0.00	1.83	0.00	9.13	0.00
Trees/shrubs	Apple tribe	Maleae	0.89	0.00	0.00	0.00	0.00	0.89	0.00
Trees/shrubs	Barberry subfamily	Berberidoideae	0.25	0.00	0.00	0.00	0.00	0.00	0.25
Herbs	Bedstraw family	Rubiaceae	1.01	0.00	0.00	0.00	0.00	1.01	0.00
Herbs	Catmint subfamily	Nepetoideae	2.26	0.00	0.00	0.00	0.00	2.26	0.00
Herbs	Nightshades	Solanum	1.43	0.00	0.00	0.00	0.71	0.71	0.00
Herbs	Nettle family	Urticaceae	1.68	0.00	0.00	0.00	1.68	0.00	0.00
Herbs		Camelineae	3.14	0.00	0.00	0.00	0.00	3.14	0.00
Herbs	Pepperworts	Lepidium	1.49	0.00	0.00	1.49	0.00	0.00	0.00
Herbs	Arum family	Araceae	0.37	0.00	0.00	0.00	0.00	0.37	0.00
Herbs	Irises	Iris	0.11	0.00	0.00	0.00	0.11	0.00	0.00

Continued on next page

Table 4.45 continued

Ecological category	Common name	Taxon	Total	522	450	346	252	145	90
		Ingroup	11001.54	660.25	248.58	506.25	1243.70	7804.75	538.01
Herbs		Dioscoreales	2.80	0.00	0.00	0.00	0.93	1.87	0.00
Herbs		Dioscorea	0.90	0.00	0.00	0.00	0.00	0.90	0.00
Trees/shrubs and herbs	Gymnosperms	Acrogymnospermae	0.05	0.00	0.00	0.05	0.00	0.00	0.00
Trees/shrubs and herbs		campanulids	1.23	0.00	0.00	0.62	0.31	0.31	0.00
Trees/shrubs and herbs		Asterales	0.29	0.00	0.00	0.00	0.29	0.00	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	0.54	0.00	0.27	0.00	0.27	0.00	0.00
Trees/shrubs and herbs		Artemisiinae	0.19	0.00	0.00	0.00	0.00	0.19	0.00
Trees/shrubs and herbs	Aster tribe	Astereae	1.04	0.00	0.00	0.00	0.00	1.04	0.00
Trees/shrubs and herbs		Stachydeae	0.72	0.00	0.00	0.72	0.00	0.00	0.00
Trees/shrubs and herbs	Bindweed family	Convolvulaceae	0.35	0.00	0.00	0.00	0.00	0.35	0.00
Trees/shrubs and herbs	Nightshade subfamily	Solanoideae	0.47	0.00	0.00	0.00	0.00	0.47	0.00
Trees/shrubs and herbs	Rose subfamily	Rosoideae	4.17	0.00	0.00	0.00	0.00	4.17	0.00
Trees/shrubs and herbs	Agrimony subtribe	Agrimoniinae	0.46	0.00	0.00	0.00	0.00	0.46	0.00
Trees/shrubs and herbs	Mustard order	Brassicales	1.30	1.30	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Mallow family	Malvaceae	5.22	0.00	0.00	0.65	0.00	4.57	0.00
Trees/shrubs and herbs		Asparagales	0.08	0.00	0.00	0.00	0.00	0.08	0.00
Trees/shrubs and herbs	Buttercup order	Ranunculales	0.31	0.00	0.00	0.00	0.00	0.31	0.00
Grasses and relatives	Grass order	Poales	5.56	0.29	0.29	0.00	0.59	4.39	0.00
Grasses and relatives	Sedges	Cyperaceae	3.08	0.00	0.00	0.00	0.00	3.08	0.00
Grasses and relatives	True sedges	Carex	9.11	0.00	0.00	2.28	0.00	6.83	0.00
Grasses and relatives	Ryegrasses	Lolium	0.31	0.00	0.00	0.31	0.00	0.00	0.00
Grasses and relatives	Barley subtribe	Hordeinae	0.12	0.00	0.00	0.00	0.00	0.12	0.00
Grasses and relatives		PACMAD clade	38.69	0.00	0.00	0.00	0.00	38.69	0.00
Grasses and relatives		**Panicoideae**	**2.70**	**0.00**	**0.00**	**0.00**	**0.00**	**2.70**	**0.00**
Grasses and relatives	**Sorghum tribe**	**Andropogoneae**	**0.40**	**0.00**	**0.00**	**0.00**	**0.00**	**0.40**	**0.00**
Ferns	Ferns	Polypodiopsida	0.20	0.00	0.00	0.07	0.00	0.14	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	0.57	0.00	0.25	0.16	0.08	0.08	0.00
Ferns	Polypod ferns	Polypodiales	2.33	0.00	0.48	1.13	0.08	0.40	0.24
Ferns	Spleenwort suborder	Aspleniineae	3.51	0.00	0.00	1.44	0.18	1.53	0.36
Ferns		Athyriaceae	0.22	0.00	0.00	0.15	0.00	0.00	0.07
Ferns		Thelypteridaceae	0.49	0.00	0.00	0.10	0.10	0.29	0.00
Ferns		Phegopteridoideae	0.14	0.00	0.00	0.00	0.14	0.00	0.00
Ferns		Thelypteridoideae	0.25	0.00	0.00	0.00	0.00	0.17	0.08
Ferns	Bracken family	Dennstaedtiaceae	0.31	0.00	0.00	0.21	0.00	0.10	0.00
Ferns		Polypodiineae	0.12	0.00	0.06	0.00	0.00	0.06	0.00
Ferns	Buckler and male ferns	Dryopteris	0.08	0.00	0.08	0.00	0.00	0.00	0.00
Ferns		Polypodiaceae	0.85	0.00	0.85	0.00	0.00	0.00	0.00
Bryophytes		Bryopsida	1.98	0.00	0.00	0.00	0.00	1.98	0.00
Bryophytes	Liverworts	Marchantiophyta	0.53	0.00	0.00	0.00	0.00	0.53	0.00
Mixed		Tracheophyta	6.05	1.04	0.52	0.35	0.52	3.11	0.52
Mixed		Euphyllophyta	7.06	0.00	0.17	0.52	1.55	4.65	0.17
Mixed	Core angiosperms	Mesangiospermae	1991.64	0.00	0.00	0.00	0.00	1991.64	0.00
Mixed	Monocots	Liliopsida	1025.35	65.48	0.00	0.00	147.12	744.64	68.11

Table 4.46: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF045. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	522	450	346	252	145	90
			Ingroup	61.43	18.55	6.92	4.17	9.44	21.84	0.52
Salt/brackish aquatics	Copepods		Nannopus	4.00	4.00	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Annelids		Protodriloides	4.17	0.00	0.00	4.17	0.00	0.00	0.00
Salt/brackish aquatics	Annelids		Orbiniidae	0.51	0.00	0.51	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Bivalves	Cockles	Cardiidae	0.66	0.00	0.00	0.00	0.66	0.00	0.00
Salt/brackish aquatics	Nemerteans		Lineidae	0.71	0.00	0.00	0.00	0.00	0.71	0.00
Salt/brackish aquatics	Sea anemones	Sea anemones	Actiniaria	4.33	0.00	2.17	0.00	0.00	2.17	0.00
Mixed aquatics	Fishes		Euteleosteomorpha	1.79	0.00	0.90	0.00	0.90	0.00	0.00
Mixed aquatics	Fishes		Percomorphaceae	2.11	0.00	1.06	0.00	0.00	1.06	0.00
Mixed aquatics	Fishes		Atherinomorphae	0.80	0.00	0.00	0.00	0.00	0.80	0.00
Mixed aquatics	Fishes		Protacanthopterygii	0.40	0.00	0.00	0.00	0.00	0.40	0.00
Mixed aquatics	Fishes	Salmon subfamily	Salmoninae	0.35	0.00	0.00	0.00	0.00	0.35	0.00
Mixed aquatics	Fishes	European salmon and trout	Salmo	0.33	0.00	0.00	0.00	0.00	0.33	0.00
Mixed aquatics	Mammals	Whales	Cetacea	0.30	0.30	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Mammals	Toothed whales	Odontoceti	0.29	0.00	0.00	0.00	0.00	0.29	0.00
Mixed aquatics	Copepods		Neocopepoda	0.31	0.31	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Copepods		Cyclopoida	1.00	0.00	1.00	0.00	0.00	0.00	0.00
Mixed aquatics	Copepods		Oithona	4.00	4.00	0.00	0.00	0.00	0.00	0.00
Mixed aquatics	Bryozoans		Ctenostomatida	1.82	0.00	0.00	0.00	1.82	0.00	0.00
Mixed aquatics	Hydrozoans		Filifera	3.51	0.00	0.00	0.00	0.00	3.51	0.00
Terrestrial	Mammals	Rodents	Rodentia	0.28	0.28	0.00	0.00	0.00	0.00	0.00
Terrestrial	**Mammals**	**Guinea pig suborder**	**Hystricomorpha**	**0.25**	**0.25**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Terrestrial	Mammals	Old World rats and mice	Murinae	0.31	0.00	0.00	0.00	0.00	0.31	0.00
Terrestrial	Birds	Landfowl	Galliformes	0.77	0.00	0.77	0.00	0.00	0.00	0.00
Terrestrial	Spiders	Money spiders	Linyphiidae	0.73	0.73	0.00	0.00	0.00	0.00	0.00
Terrestrial	Spiders	Thin-legged wolf spiders	Pardosa	0.49	0.49	0.00	0.00	0.00	0.00	0.00
Terrestrial	Butterflies	Butterflies	Papilionoidea	2.79	0.00	0.00	0.00	0.00	2.79	0.00
Terrestrial	Coleopterans	Ground beetles	Carabidae	1.74	0.00	0.00	0.00	0.00	1.74	0.00
Terrestrial	Dipterans		melanogaster subgroup	5.91	0.00	0.00	0.00	5.91	0.00	0.00
Terrestrial	Hymenopterans		Myrmicinae	2.79	2.79	0.00	0.00	0.00	0.00	0.00
Mixed	Animals	Chordates	Chordata	0.38	0.00	0.00	0.00	0.00	0.38	0.00
Mixed	Vertebrates	Vertebrates	Vertebrata	0.38	0.38	0.00	0.00	0.00	0.00	0.00
Mixed	Vertebrates	Jawed vertebrates	Gnathostomata	1.77	1.77	0.00	0.00	0.00	0.00	0.00
Mixed	Vertebrates	Tetropods and lungfish	Dipnotetrapodomorpha	0.13	0.13	0.00	0.00	0.00	0.00	0.00
Mixed	Mammals	Mammals	Mammalia	2.50	2.50	0.00	0.00	0.00	0.00	0.00
Mixed	Mammals	Carnivorans	Carnivora	0.33	0.33	0.00	0.00	0.00	0.00	0.00
Mixed	Mammals	Odd-toed ungulates	Perissodactyla	0.30	0.30	0.00	0.00	0.00	0.00	0.00
Mixed	Vertebrates		Archelosauria	0.68	0.00	0.00	0.00	0.00	0.68	0.00
Mixed	Invertebrates		Protostomia	1.04	0.00	0.52	0.00	0.00	0.00	0.52
Mixed	Spiders	Spiders	Araneae	0.42	0.00	0.00	0.00	0.00	0.42	0.00
Mixed	Crustaceans		Multicrustacea	0.15	0.00	0.00	0.00	0.15	0.00	0.00
Mixed	Lepidopterans		Ditrysia	1.56	0.00	0.00	0.00	0.00	1.56	0.00
Mixed	Gastropods		Heterobranchia	0.61	0.00	0.00	0.00	0.00	0.61	0.00

Embryophyta

Aquatics

At first glance, the six samples in ELF045 appear homogeneous for Embryophyta. All show a classic *Zostera*-dominated pattern, with high-yield samples also assigning increasing proportions to Alismatales, Liliopsida, and Mesangiospermae. Aquatic taxa besides *Zostera* include *Ruppia*, Typhaceae, and the reed group, often associated with *Zostera* in what was presumably an estuary with reed beds or saltmarsh. However, other aquatics seem to have low biogenomic masses compared to similar cores, despite the generally high read counts. *Zostera* is particularly dominant.

Terrestrial

Woody taxa are present in all samples but the deepest, with most biogenomic mass from the willow group. Note also *Picea* (spruces) in sample 252: no *Picea* are native to Great Britain, but *P. abies* (Norway spruce) is a key component of Boreal and mountainous forest across Europe and was part of the British flora during the last complete interglacial (Packham *et al.* 1992). The read may have been mis-assigned (*P. abies* is one of the few gymnosperm genomes available; Caudullo *et al.* 2016), but it would not be implausible.

A second surprising if not inexplicable taxon is Juglandaceae, previously discussed in ELF059A and ELF031A. It is present across two samples: 346 and 252. Walnut is associated with warm climates, so its occurrence in the same sample as *Picea* casts some doubt on them both, although *Picea* pollen can travel long distances (Pan *et al.* 2013), so they may have originated from different regions.

Other terrestrial ecological categories are only a minor component of the taxonomic profiles. This includes grasses and relatives, which have a clear signal in most samples, yet make up only a tiny percentage of ELF045.

Metazoa

Aside from the single ambiguous taxon in 90, all samples contain saltwater taxa that support the particularly strong marine signal from Embryophyta. Samples 522 and 145 also contain Cetacea and Odontoceti respectively, which while technically mixed aquatics are almost exclusively marine. Most samples also contain a terrestrial signal, although the tetrapods are close to well-studied animals and may have been over-assigned. Note also the substantial Mammalia signal in 522; reads assigned to Primates were removed during analysis, but we expect read counts for other parent taxa of humans to be inflated by the inevitable human contamination. The negative controls may not always remove these taxa, which is possibly what happened for Mammalia here.

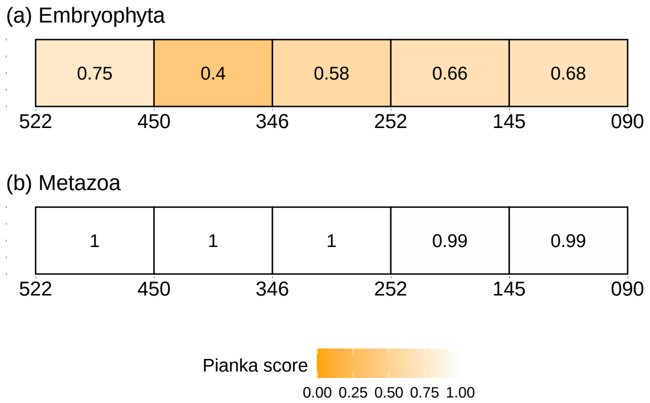


Figure 4.112: Pianka similarity scores between adjacent samples in ELF045 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

4.23.3 Pianka scores

The Embryophyta Pianka scores are very close to 1, reflecting the clear similarity between samples and probably also a product of the generally high read counts. The Metazoa scores are lower and appear strongly influenced by read count. In any case, there is no clear evidence of change across the core.

4.23.4 Summary

ELF045 is a strongly saltwater core. *Zostera* and Alismatales dominate, and the accompanying terrestrial signal is weaker than in other *Zostera*-type cores, although still includes two interesting and incongruous taxa: *Picea* and Juglandaceae. Considering its location near the mouth of the channel, ELF045 may have been less influenced by terrestrial or freshwater taxa than other aquatic cores because, even before the palaeochannel was inundated, ELF045 was in closer proximity to the sea and its currents.

4.24 ELF047

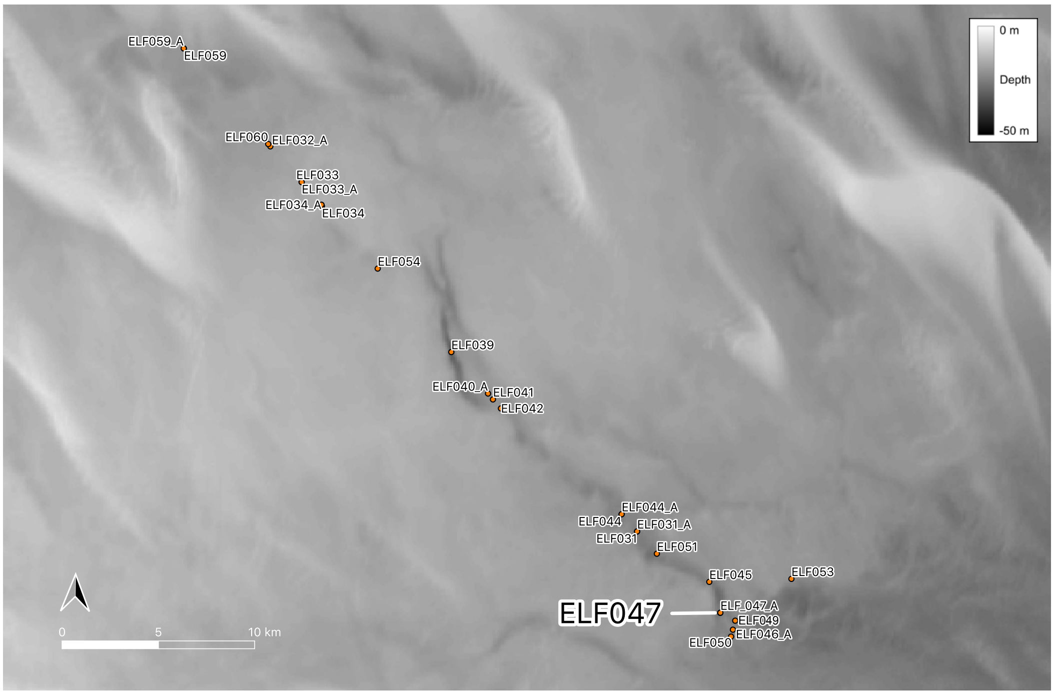


Figure 4.113: Bathymetry map of the palaeochannel transect highlighting ELF047. Darker grey indicates greater depth.

ELF047 is the first in a cluster of cores around the mouth of the palaeochannel. It shares the location with ELF047A.

4.24.1 Read counts

Read counts are variable with little pattern. Sample 70 (top) has only 63 Embryophyta reads, whereas the neighbouring 150 has over 5,500. The Metazoa data has similar peaks and troughs but with a maximum of only eight reads, and sample 241 is empty.

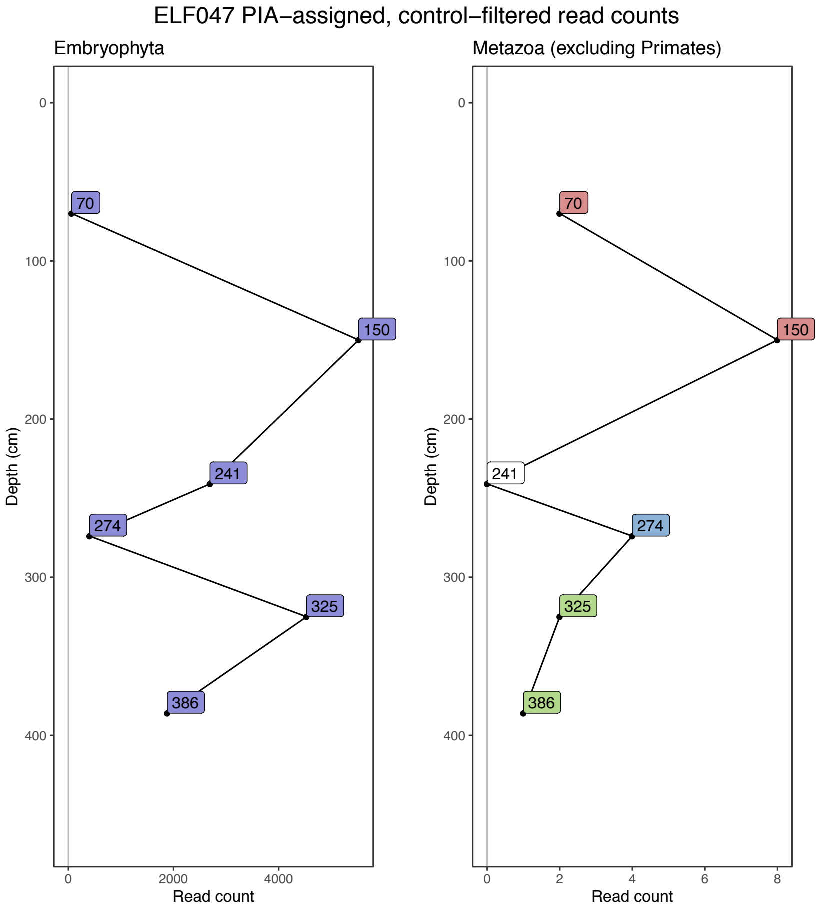
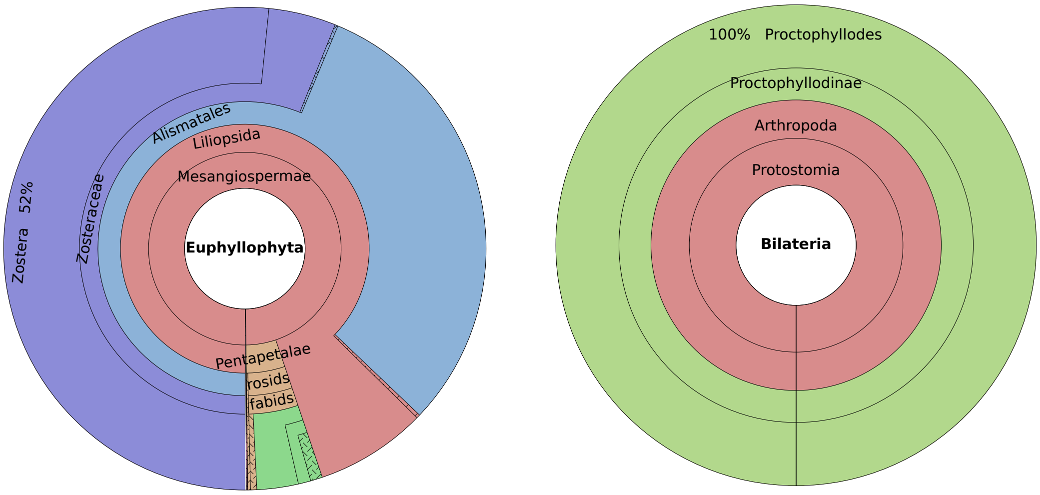


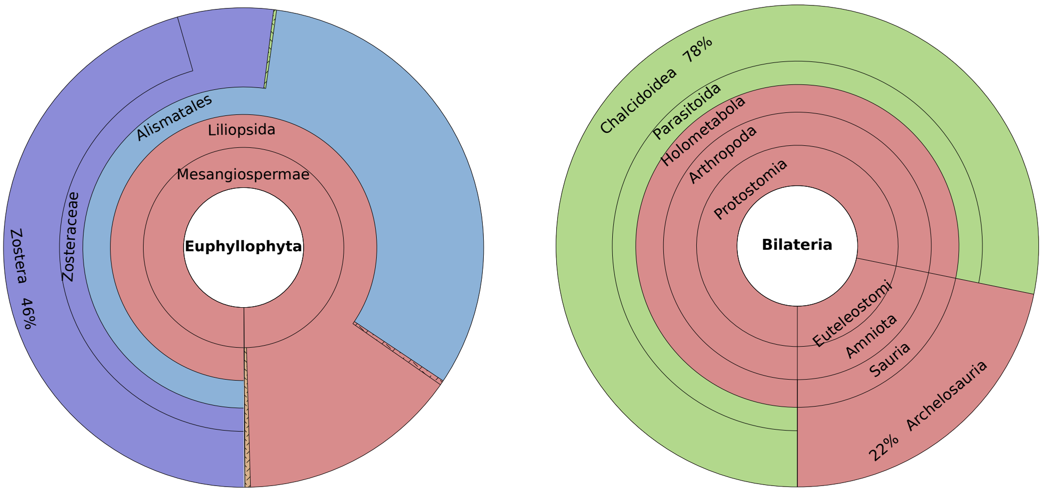
Figure 4.114: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF047. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.24.2 Taxonomic profiles

ELF047_386



ELF047_325



ELF047_274

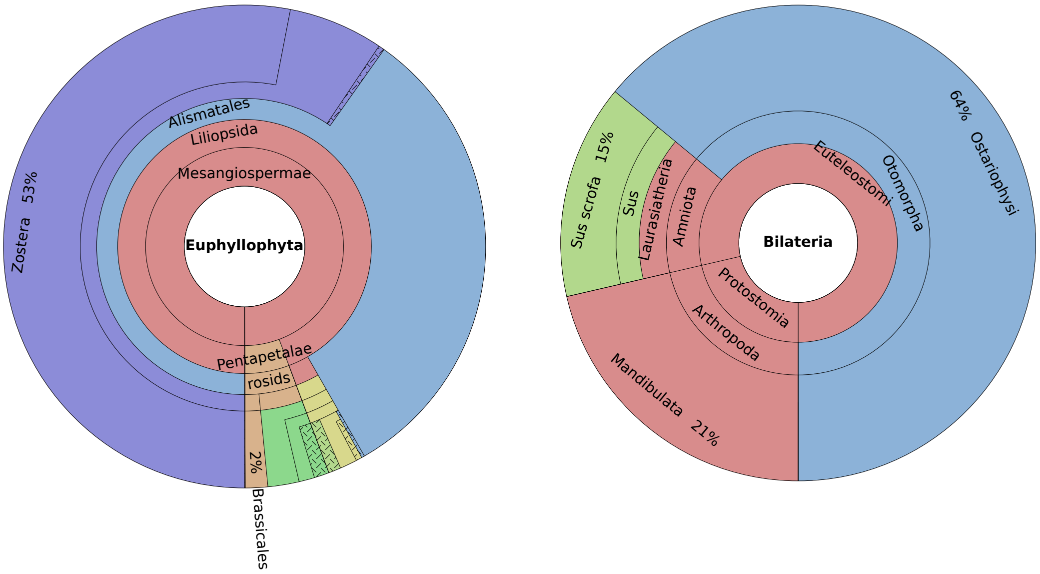
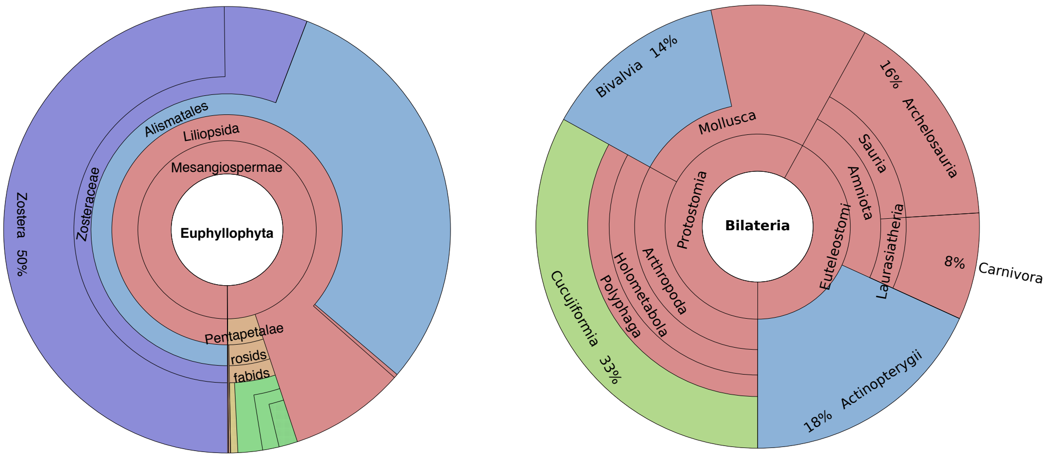
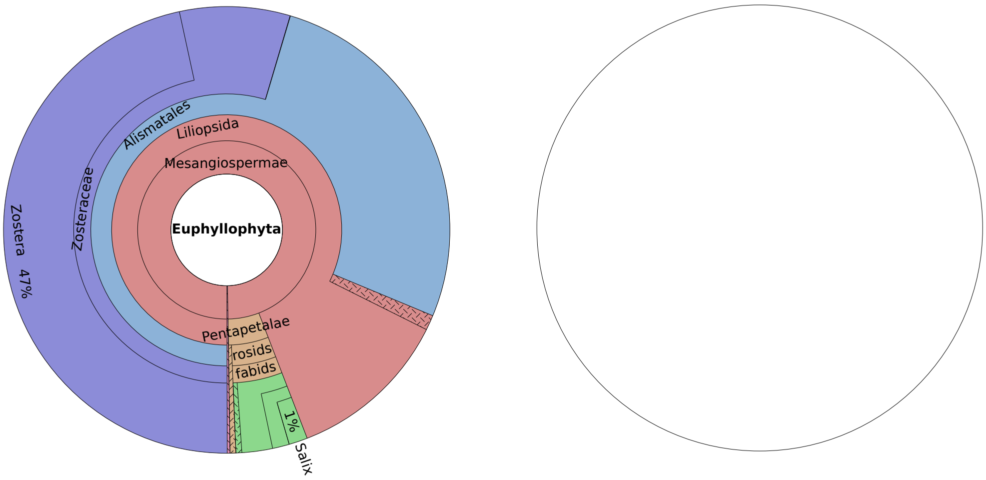


Figure 4.115: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF047, samples 386, 325, and 274. Continued in figure 4.116. See figure 4.4 for colour key.

ELF047_241



ELF047_150



ELF047_070

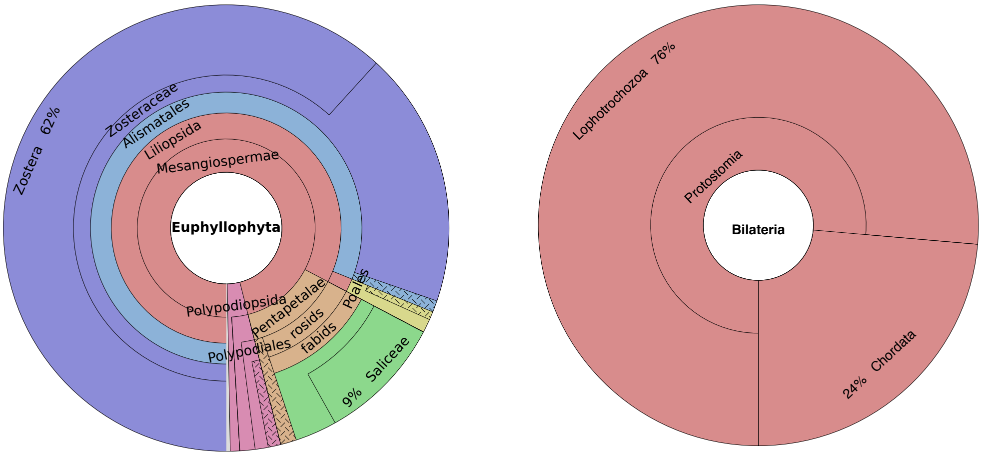


Figure 4.116: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF047, samples 241, 150, and 70. See figure 4.4 for colour key.

Table 4.47: Biogenomic masses of European taxa in Embryophyta from ELF047. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	386	325	274	241	150	70
		Ingroup	4856.26	650.82	1254.49	177.86	914.13	1799.54	59.42
Freshwater aquatics	Duckweed family	Lemnoideae	2.44	0.00	2.44	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Manatee-grass family	Cymodoceaceae	1.39	0.70	0.00	0.70	0.00	0.00	0.00
Salt/brackish aquatics	Tasselweeds	Ruppia	0.43	0.43	0.00	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Eelgrass family	Zosteraceae	330.28	29.36	80.73	11.01	55.05	143.12	11.01
Salt/brackish aquatics	Eelgrasses	Zostera	2324.77	335.78	570.64	91.74	455.05	834.86	36.70
Mixed aquatics	Alismatids	Alismatales	1414.34	200.47	403.15	55.17	277.88	477.67	0.00
Mixed aquatics	Pondweed family	Potamogetonaceae	0.99	0.00	0.00	0.00	0.00	0.50	0.50
Mixed aquatics	Reeds	Arundinoideae	1.77	0.00	0.00	0.00	0.35	1.06	0.35
Mixed aquatics	Common reed tribe	Molinieae	8.46	0.71	1.06	0.35	1.41	4.94	0.00
Mixed aquatics	Common reeds	Phragmites	0.42	0.00	0.00	0.00	0.00	0.42	0.00
Halophytes	Leadwort family	Plumbaginaceae	0.21	0.00	0.00	0.00	0.00	0.21	0.00
Trees/shrubs		Viburnum	0.26	0.00	0.00	0.00	0.26	0.00	0.00
Trees/shrubs	Birch family	Betulaceae	7.56	0.00	1.08	0.00	1.08	5.40	0.00
Trees/shrubs	Hazels	Corylus	2.27	0.00	0.00	0.00	0.00	2.27	0.00
Trees/shrubs	Willow family	Salicaceae	81.23	18.05	1.81	3.61	16.25	39.71	1.81
Trees/shrubs	Willow tribe	Saliceae	45.39	5.45	0.00	1.82	10.89	21.79	5.45
Trees/shrubs	Poplars	Populus	1.98	0.00	0.00	0.00	1.98	0.00	0.00
Trees/shrubs	Willows	Salix	43.25	5.19	1.73	1.73	10.38	24.22	0.00
Trees/shrubs	Stone fruit trees	Prunus	3.65	1.83	0.00	0.00	1.83	0.00	0.00
Trees/shrubs	Apple tribe	Maleae	0.89	0.89	0.00	0.00	0.00	0.00	0.00
Herbs	Mints	Mentha	3.05	0.00	0.00	0.00	0.00	3.05	0.00
Herbs	Rhubarb family	Rumiceae	0.41	0.00	0.00	0.00	0.41	0.00	0.00
Herbs	Arum family	Araceae	0.18	0.18	0.00	0.00	0.00	0.00	0.00
Herbs	Bluebell family	Hyacinthaceae	0.08	0.08	0.00	0.00	0.00	0.00	0.00
Herbs	Orchids	Orchidaceae	0.17	0.00	0.00	0.00	0.00	0.17	0.00
Herbs		Nartheciaceae	2.56	0.00	0.00	0.00	0.00	2.56	0.00
Herbs	Fritillaries	Fritillaria	0.04	0.00	0.00	0.00	0.00	0.04	0.00
Trees/shrubs and herbs		campanulids	0.62	0.00	0.00	0.00	0.00	0.62	0.00
Trees/shrubs and herbs		Apiineae	0.47	0.47	0.00	0.00	0.00	0.00	0.00
Trees/shrubs and herbs		Asterales	0.86	0.00	0.00	0.00	0.00	0.86	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	1.36	0.27	0.00	0.00	0.27	0.81	0.00
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	0.37	0.00	0.00	0.00	0.00	0.37	0.00
Trees/shrubs and herbs	Aster tribe	Astereae	1.04	0.52	0.00	0.00	0.00	0.52	0.00
Trees/shrubs and herbs		lamiids	0.68	0.00	0.00	0.00	0.00	0.00	0.68
Trees/shrubs and herbs	Dead-nettle order	Lamiales	0.88	0.00	0.00	0.00	0.00	0.88	0.00
Trees/shrubs and herbs	Bindweed family	Convolvulaceae	0.35	0.00	0.00	0.00	0.35	0.00	0.00
Trees/shrubs and herbs		Rosoideae incertae sedis	1.88	0.00	0.00	0.00	1.88	0.00	0.00
Trees/shrubs and herbs	Mustard order	Brassicales	2.60	0.00	0.00	2.60	0.00	0.00	0.00
Trees/shrubs and herbs		Saxifragales	0.19	0.00	0.00	0.00	0.00	0.19	0.00
Trees/shrubs and herbs		Asparagales	0.08	0.00	0.00	0.00	0.00	0.08	0.00
Grasses and relatives	Grass order	Poales	3.51	0.59	0.59	0.00	0.88	0.88	0.59
Grasses and relatives	Sedges	Cyperaceae	1.54	0.00	0.00	0.00	0.00	1.54	0.00
Grasses and relatives	True sedges	Carex	9.11	0.00	2.28	0.00	0.00	6.83	0.00

Continued on next page

Table 4.47 continued

Ecological category	Common name	Taxon	Total	386	325	274	241	150	70
		Ingroup	4856.26	650.82	1254.49	177.86	914.13	1799.54	59.42
Grasses and relatives		Poeae	0.18	0.18	0.00	0.00	0.00	0.00	0.00
Grasses and relatives		Chloridoideae	2.40	0.00	0.00	1.44	0.00	0.96	0.00
Grasses and relatives		**Paniceae**	**2.01**	**0.00**	**0.00**	**2.01**	**0.00**	**0.00**	**0.00**
Grasses and relatives	**Bristle grasses**	**Setaria**	**0.69**	**0.00**	**0.00**	**0.69**	**0.00**	**0.00**	**0.00**
Ferns	Ferns	Polypodiopsida	0.47	0.00	0.00	0.00	0.00	0.07	0.41
Ferns	Horsetails	Equisetum	0.05	0.00	0.00	0.00	0.05	0.00	0.00
Ferns	Leptosporangiate ferns	Polypodiidae	0.82	0.00	0.00	0.00	0.00	0.16	0.66
Ferns	Polypod ferns	Polypodiales	2.41	0.16	0.00	0.00	0.24	1.45	0.56
Ferns	Spleenwort suborder	Aspleniineae	1.89	0.27	0.00	0.00	0.36	1.26	0.00
Ferns		Athyriaceae	0.15	0.00	0.00	0.00	0.00	0.07	0.07
Ferns		Thelypteridaceae	0.20	0.00	0.00	0.00	0.00	0.20	0.00
Ferns		Phegopteridoideae	0.14	0.00	0.00	0.00	0.00	0.14	0.00
Ferns		Thelypteridoideae	0.17	0.08	0.00	0.00	0.00	0.00	0.08
Ferns	Bracken family	Dennstaedtiaceae	0.62	0.00	0.10	0.00	0.00	0.21	0.31
Ferns		Polypodiineae	0.06	0.00	0.00	0.00	0.00	0.06	0.00
Ferns		Dryopteridoideae	0.08	0.00	0.00	0.08	0.00	0.00	0.00
Ferns		Polypodiaceae	0.08	0.00	0.00	0.00	0.00	0.00	0.08
Mixed		Tracheophyta	4.32	0.00	1.90	0.17	0.35	1.90	0.00
Mixed		Euphyllophyta	2.93	0.69	1.03	0.00	0.34	0.69	0.17
Mixed	Monocots	Liliopsida	521.86	48.48	185.95	0.00	76.61	210.82	0.00

Table 4.48: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF047. Samples are in cm.

Ecological category	Group	Common name	Taxon	Total	386	325	274	241	150	70
			Ingroup	20.48	6.67	3.13	2.19	0.00	7.68	0.81
Mixed aquatics	Fishes	Ray-finned fishes	Actinopterygii	0.78	0.00	0.00	0.00	0.00	0.78	0.00
Mixed aquatics	Fishes	Carp subcohort	Ostariophysi	1.40	0.00	0.00	1.40	0.00	0.00	0.00
Mixed aquatics	Bivalves	Bivalves	Bivalvia	0.58	0.00	0.00	0.00	0.00	0.58	0.00
Terrestrial	Mammals	Wild boar and domestic pig	Sus scrofa	0.32	0.00	0.00	0.32	0.00	0.00	0.00
Terrestrial	Mites	Group of feather mites	Proctophyllodes	6.67	6.67	0.00	0.00	0.00	0.00	0.00
Terrestrial	Coleopterans		Cucujiformia	1.41	0.00	0.00	0.00	0.00	1.41	0.00
Terrestrial	Hymenopterans	Chalcid wasps	Chalcidoidea	2.45	0.00	2.45	0.00	0.00	0.00	0.00
Mixed	Animals	Chordates	Chordata	0.19	0.00	0.00	0.00	0.00	0.00	0.19
Mixed	Mammals	Carnivorans	Carnivora	0.33	0.00	0.00	0.00	0.00	0.33	0.00
Mixed	Vertebrates		Archelosauria	1.36	0.00	0.68	0.00	0.00	0.68	0.00
Mixed	Arthropods		Mandibulata	0.47	0.00	0.00	0.47	0.00	0.00	0.00
Mixed	Invertebrates	Lophotrochozoans	Lophotrochozoa	0.62	0.00	0.00	0.00	0.00	0.00	0.62
Mixed	Molluscs	Molluscs	Mollusca	0.49	0.00	0.00	0.00	0.00	0.49	0.00

Embryophyta

The taxonomic profiles through ELF047 are generally similar. All show a typical *Zostera*-dominated pattern, with increasing assignment to Alismatales and Liliopsida with greater data yields. The only freshwater aquatic is Lemnoideae (duckweed family) which has a small signal in sample 325. Mixed aquatics besides Alismatales are similarly rare. Most samples have a very weak signal from the reed group, but the other common companion to *Zostera*, *Ruppia*, is absent. Terrestrial taxa are a little more diverse. All samples have some tree signal, primarily from Betulaceae, the willow group, and *Prunus*, and a few reads from grasses and relatives. Other groups are more scattered.

Metazoa

The Metazoa results are limited and most taxa are ambiguous. However, the presence of an aquatic in sample 274 suggests that Embryophyta did not give a full picture, and that 274 may be more similar to the rest of the core than it appears.

4.24.3 Pianka scores

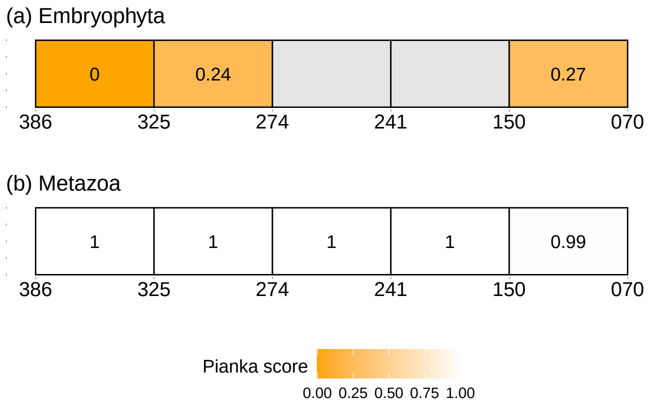


Figure 4.117: Pianka similarity scores between adjacent samples in ELF047 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

The Embryophyta Pianka scores are very high, reflecting the very similar taxonomic profile. In contrast, the Metazoa scores that could be calculated are among the lowest across all cores, although this is probably an artefact of the low read counts. There is no clear evidence of environmental change.

4.24.4 Summary

ELF047 is another largely saltwater aquatic core. Other ecological categories are limited. There is some evidence of floodplain woodland in the area, but little of reed beds or brackish influence, suggesting open water. This is consistent with the bathymetry, which shows that ELF047 was located in a river mouth or estuary at some point in time. Increased proximity to the sea may have diluted any signals from inland environments.

4.25 ELF047A

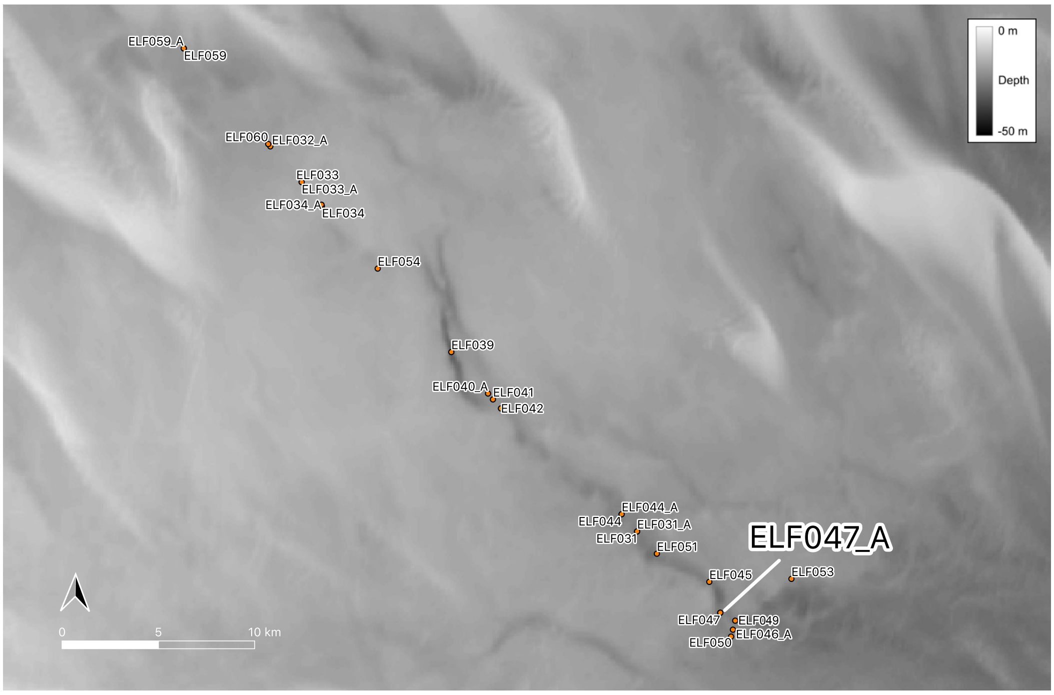


Figure 4.118: Bathymetry map of the palaeochannel transect highlighting ELF047A. Darker grey indicates greater depth.

ELF047A is in the same location as ELF047, inside the mouth of the palaeochannel.

4.25.1 Read counts

Data yield was low across most of this core, with the lower three samples returning no more than 20 Embryophyta reads each. The top sample returned nearly 2,000, although only ten for Metazoa. It will be difficult to compare such disparate samples.

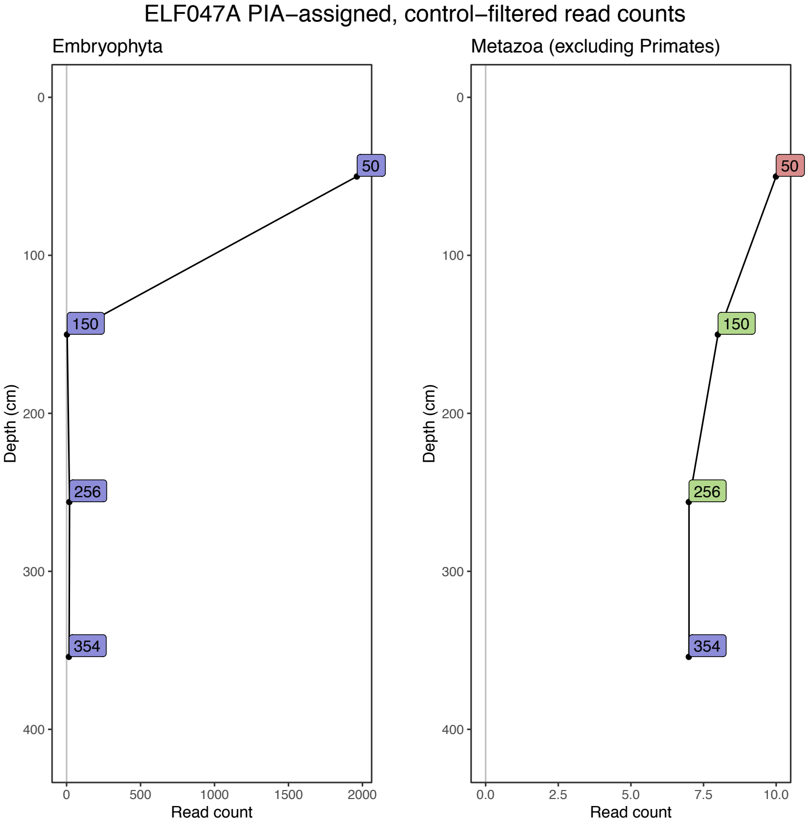


Figure 4.119: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF047A. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.25.2 Taxonomic profiles

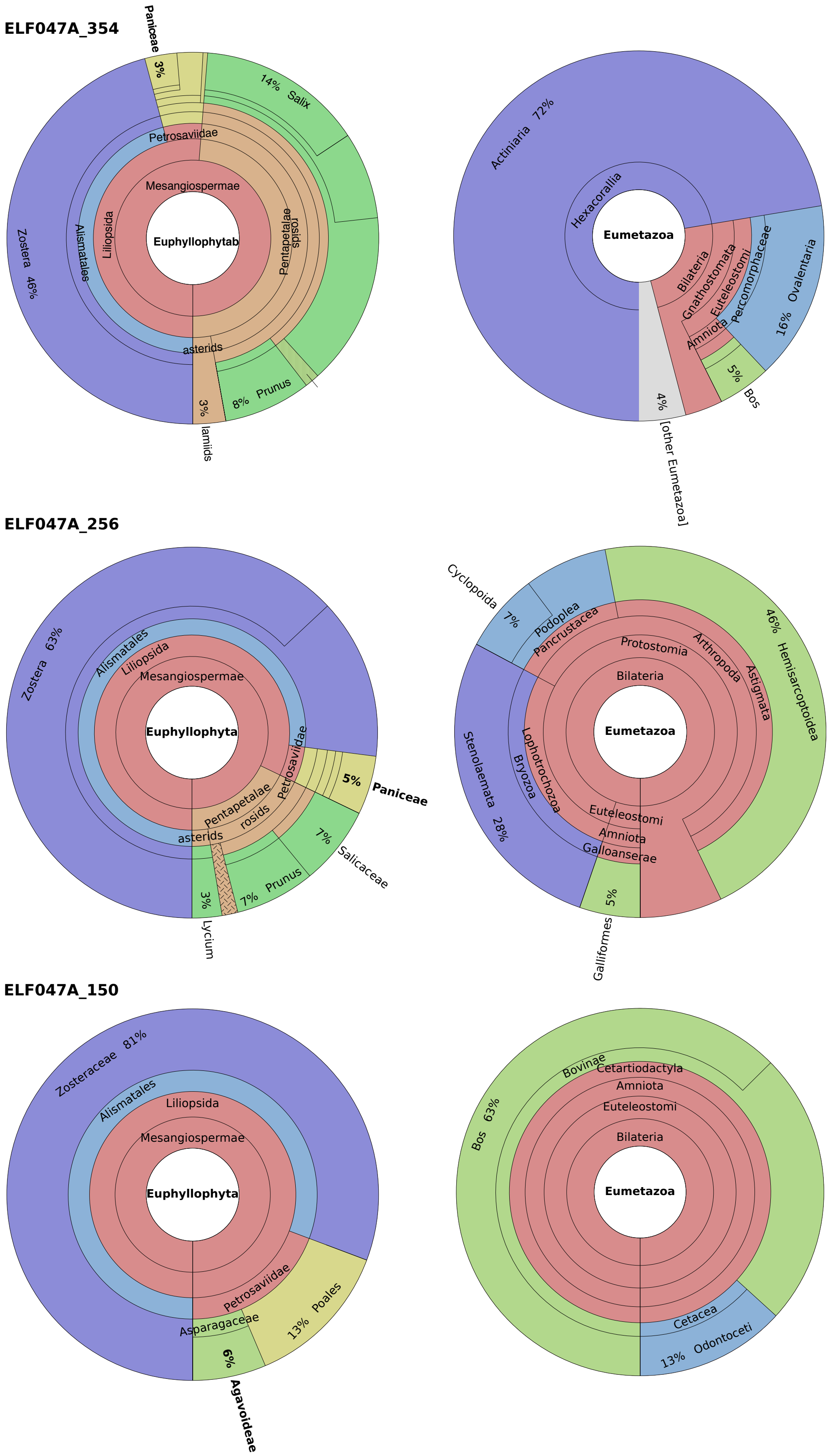


Figure 4.120: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF047A, samples 354, 256, and 150. Continued in figure 4.121. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

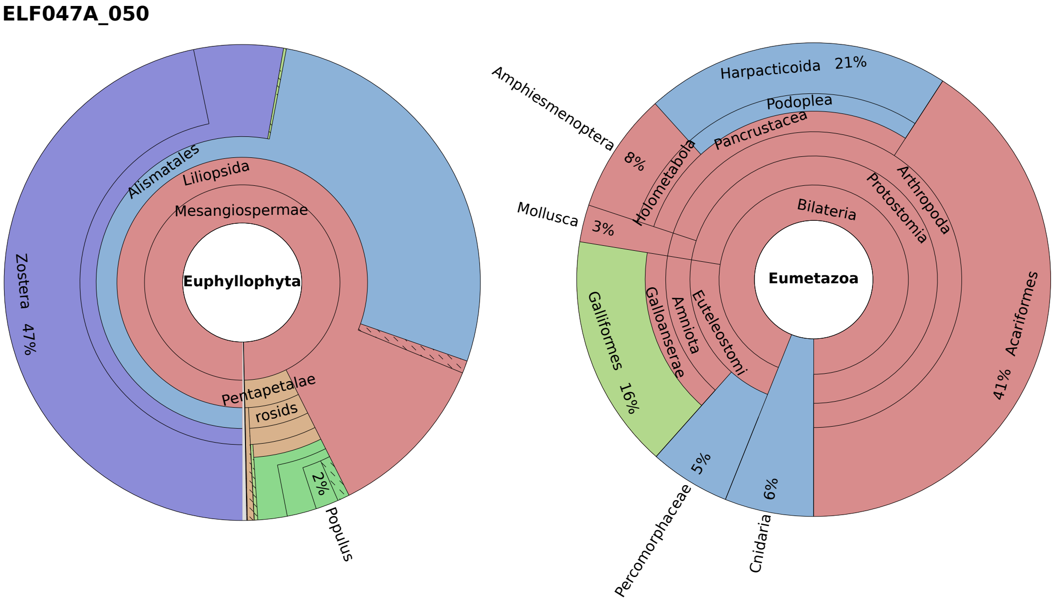


Figure 4.121: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF047A, sample 50. See figure 4.4 for colour key.

Table 4.49: Biogenomic masses of European taxa in Embryophyta from ELF047A. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	354	256	150	50
		Ingroup	693.07	24.02	27.65	2.98	638.42
Freshwater aquatics	Duckweed family	Lemnoideae	1.22	0.00	0.00	0.00	1.22
Salt/brackish aquatics	Eelgrass family	Zosteraceae	44.04	0.00	3.67	1.83	38.53
Salt/brackish aquatics	Eelgrasses	Zostera	324.77	11.01	16.51	0.00	297.25
Mixed aquatics	Alismatids	Alismatales	174.16	0.00	0.00	0.00	174.16
Mixed aquatics	Common reed tribe	Molinieae	1.76	0.00	0.00	0.00	1.76
Trees/shrubs	**Boxthorns**	**Lycium**	**0.68**	**0.00**	**0.68**	**0.00**	**0.00**
Trees/shrubs	Birch family	Betulaceae	1.08	0.00	0.00	0.00	1.08
Trees/shrubs	Willow family	Salicaceae	18.05	3.61	1.81	0.00	12.64
Trees/shrubs	Willow tribe	Saliceae	14.52	1.82	0.00	0.00	12.71
Trees/shrubs	Poplars	Populus	9.89	0.00	0.00	0.00	9.89
Trees/shrubs	Willows	Salix	8.65	3.46	0.00	0.00	5.19
Trees/shrubs	Stone fruit trees	Prunus	3.65	1.83	1.83	0.00	0.00
Herbs	Spurge family	Euphorbiaceae	0.31	0.31	0.00	0.00	0.00
Herbs	**Amyris subfamily**	**Amyridoideae**	**0.36**	**0.00**	**0.36**	**0.00**	**0.00**
Herbs		**Agavoideae**	**0.15**	**0.00**	**0.00**	**0.15**	**0.00**
Trees/shrubs and herbs		campanulids	0.31	0.00	0.00	0.00	0.31
Trees/shrubs and herbs	Ivy family	Araliaceae	0.50	0.00	0.00	0.00	0.50
Trees/shrubs and herbs		Asterales	0.29	0.00	0.00	0.00	0.29
Trees/shrubs and herbs	Daisy family	Asteraceae	0.27	0.00	0.00	0.00	0.27
Trees/shrubs and herbs		Asteroideae	0.25	0.00	0.00	0.00	0.25
Trees/shrubs and herbs	Honeysuckle family	Caprifoliaceae	0.57	0.00	0.00	0.00	0.57
Trees/shrubs and herbs		lamiids	0.68	0.68	0.00	0.00	0.00
Trees/shrubs and herbs	Knotweed tribe	Polygoneae	0.73	0.00	0.00	0.00	0.73
Trees/shrubs and herbs		Fabales	0.47	0.00	0.00	0.00	0.47
Grasses and relatives	Grass order	Poales	1.76	0.00	0.00	0.29	1.46
Grasses and relatives	Sedges	Cyperaceae	1.54	0.00	0.00	0.00	1.54
Grasses and relatives	**Goat grasses**	**Aegilops**	**0.10**	**0.10**	**0.00**	**0.00**	**0.00**
Grasses and relatives		**Panicoideae**	**0.54**	**0.54**	**0.00**	**0.00**	**0.00**
Grasses and relatives		**Paniceae**	**2.68**	**0.67**	**1.34**	**0.00**	**0.67**
Ferns	Horsetails	Equisetum	0.05	0.00	0.00	0.00	0.05
Ferns	Polypod ferns	Polypodiales	0.24	0.00	0.00	0.00	0.24
Mixed		Tracheophyta	0.35	0.00	0.00	0.00	0.35
Mixed		Euphyllophyta	1.89	0.00	0.00	0.00	1.89
Mixed	Monocots	Liliopsida	72.51	0.00	0.00	0.00	72.51

Table 4.50: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF047A. Samples are in cm.

Ecological category	Group	Common name	Taxon	Total	354	256	150	50
			Ingroup	42.89	6.94	14.53	2.22	19.21
Salt/brackish aquatics	Bryozoans		Stenolaemata	4.00	0.00	4.00	0.00	0.00
Salt/brackish aquatics	Sea anemones	Sea anemones	Actiniaria	4.33	4.33	0.00	0.00	0.00
Mixed aquatics	Fishes		Percomorphaceae	1.06	0.00	0.00	0.00	1.06
Mixed aquatics	Fishes		Ovalentaria	0.93	0.93	0.00	0.00	0.00
Mixed aquatics	Mammals	Toothed whales	Odontoceti	0.29	0.00	0.00	0.29	0.00
Mixed aquatics	Copepods		Podoplea	1.05	0.00	1.05	0.00	0.00
Mixed aquatics	Copepods		Cyclopoida	1.00	0.00	1.00	0.00	0.00
Mixed aquatics	Copepods		Harpacticoida	4.00	0.00	0.00	0.00	4.00
Mixed aquatics	Cnidarians	Cnidarians	Cnidaria	1.17	0.00	0.00	0.00	1.17
Terrestrial	Mammals	Cattle subfamily	Bovinae	0.54	0.00	0.00	0.54	0.00
Terrestrial	Mammals	Wild and domestic cattle	Bos	1.67	0.28	0.00	1.39	0.00
Terrestrial	Birds	Landfowl	Galliformes	3.84	0.00	0.77	0.00	3.07
Terrestrial	Mites		Hemisarcoptoidea	6.67	0.00	6.67	0.00	0.00
Mixed	Animals		Eumetazoa	0.25	0.25	0.00	0.00	0.00
Mixed	Vertebrates	Jawed vertebrates	Gnathostomata	0.20	0.20	0.00	0.00	0.00
Mixed	Invertebrates		Protostomia	1.04	0.00	1.04	0.00	0.00
Mixed	Mites		Acariformes	7.84	0.00	0.00	0.00	7.84
Mixed	Insects		Amphiesmenoptera	1.58	0.00	0.00	0.00	1.58
Mixed	Molluscs	Molluscs	Mollusca	0.49	0.00	0.00	0.00	0.49

Embryophyta

The samples appear similar at first glance because all contain mostly *Zostera*. Other aquatics, like most taxa, are found only in the top sample, as would be expected from its much higher data yield. Woody taxa are more consistent, with the willow group and *Prunus* found across multiple samples. Grasses are also a small component of every sample, although *Aegilops* and the Panicoideae group are questionable.

Metazoa

There is no clear pattern in ecological categories and little taxonomic overlap between samples. All contain a mixture of terrestrial and aquatics, although three of four terrestrial taxa contain domesticates so may be over-assigned.

4.25.3 Pianka scores

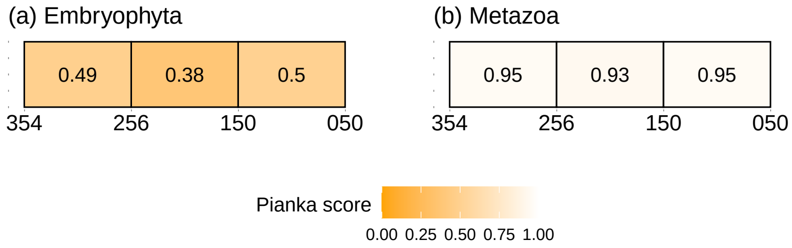


Figure 4.122: Pianka similarity scores between adjacent samples in ELF047A for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

The Embryophyta Pianka scores for samples 354-150 are surprisingly high considering the low read counts, strongly suggesting that the similarity across the core is genuine. The Metazoa scores are much lower, as expected; there is little sign of any pattern in the Metazoa data, whether change or consistency.

4.25.4 Summary

ELF047A, like its sister core, shows a predominantly saltwater environment with a small terrestrial signal, although the lack of data in the three deeper samples precludes much detail. The frequently observed floodplain woodland is present in all samples alongside very minor grasses and herbs, but there is no clear evidence of reed beds or marsh. Similar to ELF047, the most likely interpretation would be something like an estuary.

4.26 ELF049

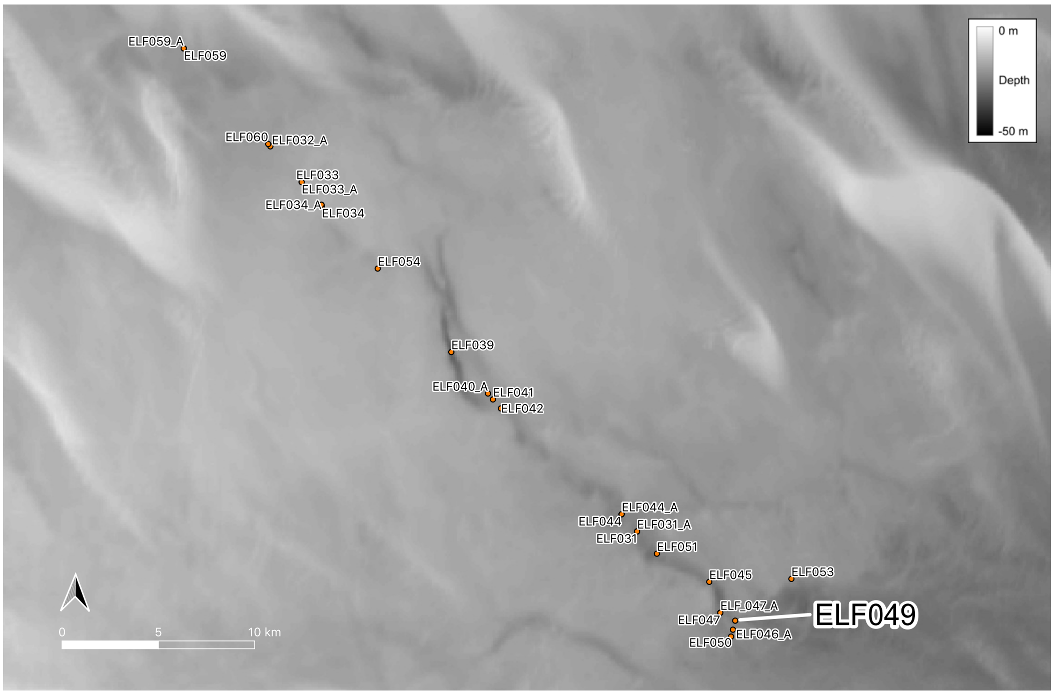


Figure 4.123: Bathymetry map of the palaeochannel transect highlighting ELF049. Darker grey indicates greater depth.

ELF049 is approximately 1 km southeast of ELF047 and ELF047A, further into the depression at the end of the palaeochannel.

4.26.1 Read counts

Only two samples were taken from this long core. Sample 361 returned only 136 Embryophyta reads, but 295 performed better with over 1,200. Data from Metazoa is limited to five reads each.

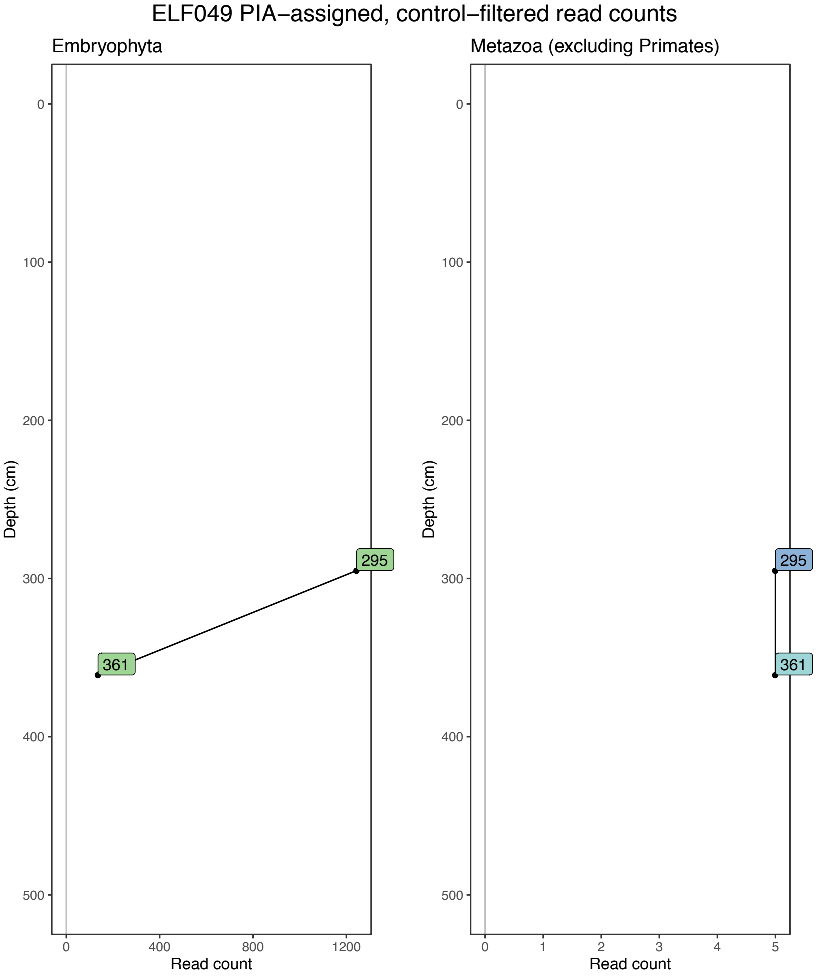
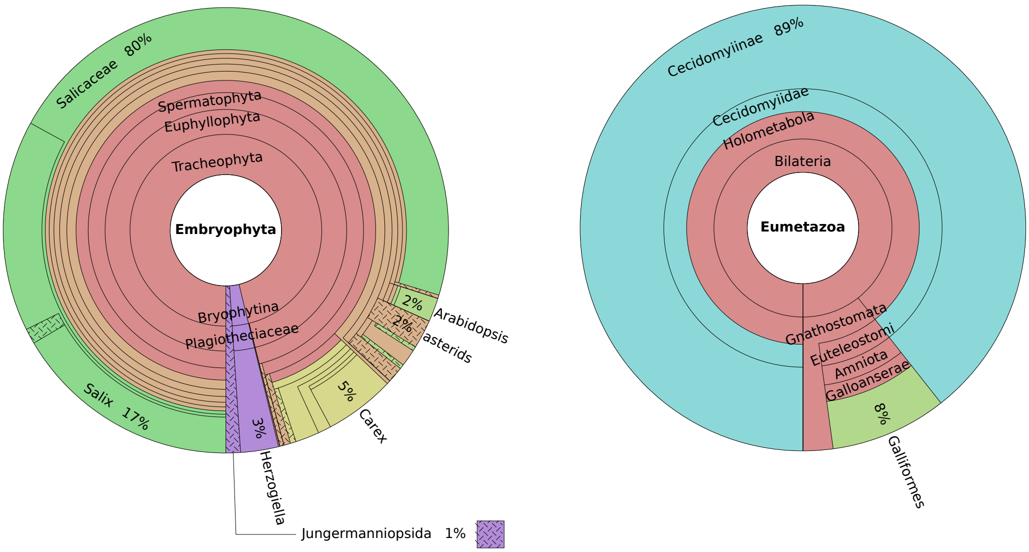


Figure 4.124: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF049. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.26.2 Taxonomic profiles

ELF049_361



ELF049_295

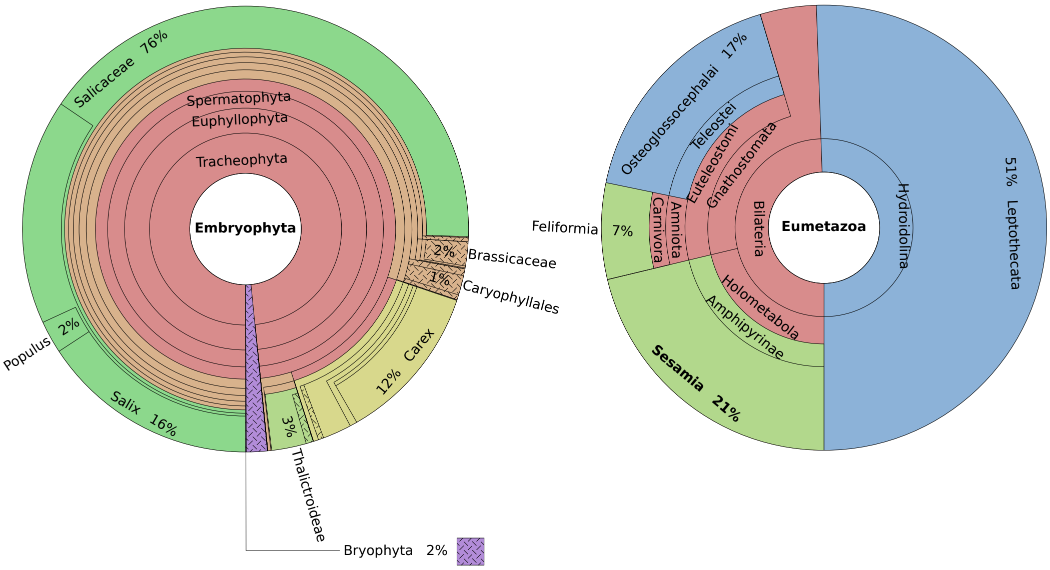


Figure 4.125: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF049. Taxa in bold are not native to Great Britain. See figure 4.4 for colour key.

Table 4.51: Biogenomic masses of European taxa in Embryophyta from ELF049. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	361	295
		Ingroup	2102.76	178.95	1923.80
Freshwater aquatics		Menyanthaceae	1.43	0.00	1.43
Mixed aquatics	Common reed tribe	Molinieae	0.35	0.00	0.35
Halophytes	Leadwort family	Plumbaginaceae	0.83	0.00	0.83
Halophytes	Sea-lavenders	Limonium	0.21	0.00	0.21
Halophytes	Scurvy-grass tribe	Cochlearieae	7.37	0.00	7.37
Halophytes	Scurvy-grasses	Cochlearia	1.42	0.00	1.42
Xerophytes	Goosefoot family	Chenopodiaceae	0.80	0.00	0.80
Trees/shrubs	Conifers	Pinidae	0.11	0.11	0.00
Trees/shrubs	**Spruces**	**Picea**	**0.21**	**0.21**	**0.00**
Trees/shrubs	Heather family	Ericaceae	0.33	0.33	0.00
Trees/shrubs	Willow family	Salicaceae	871.84	83.03	788.81
Trees/shrubs	Willow tribe	Saliceae	341.30	27.23	314.07
Trees/shrubs	Poplars	Populus	47.47	1.98	45.49
Trees/shrubs	Willows	Salix	332.18	29.41	302.77
Trees/shrubs	Dryas subfamily	Dryadoideae	1.67	0.00	1.67
Herbs	Thistle subfamily	Carduoideae	0.48	0.00	0.48
Herbs	Thistle tribe	Cardueae	0.95	0.00	0.95

Continued on next page

Table 4.51 continued

Ecological category	Common name	Taxon	Total	361	295
		Ingroup	2102.76	178.95	1923.80
Herbs	Bellflower family	Campanulaceae	0.94	0.00	0.94
Herbs	Bellflowers	Campanula	0.55	0.00	0.55
Herbs	Primrose family	Primulaceae	0.35	0.00	0.35
Herbs	**Rock jasmine**	**Androsace**	**0.35**	**0.00**	**0.35**
Herbs	Catmint subfamily	Nepetoideae	1.13	1.13	0.00
Herbs	Plantains	Plantago	0.46	0.00	0.46
Herbs	Pink family	Caryophyllaceae	3.42	0.00	3.42
Herbs		Alsineae	1.16	0.58	0.58
Herbs	Knotweed subfamily	Polygonoideae	2.79	0.56	2.23
Herbs	Knotweed tribe	Persicarieae	0.45	0.00	0.45
Herbs	Rhubarb family	Rumiceae	0.83	0.00	0.83
Herbs		Hologalegina	0.28	0.28	0.00
Herbs	Vetches	Vicia	0.18	0.18	0.00
Herbs		Galegeae	2.02	0.00	2.02
Herbs	Flaxes	Linum	0.65	0.00	0.65
Herbs	Cinquefoils	Potentilla	1.42	0.00	1.42
Herbs		Arabideae	17.51	0.00	17.51
Herbs	Whitlow-grasses	Draba	2.41	0.00	2.41
Herbs	Thale cresses	Arabidopsis	6.15	3.07	3.07
Herbs	Geranium family	Geraniaceae	0.81	0.00	0.81
Herbs		Saxifragaceae	3.43	1.37	2.06
Herbs	Saxifrages	Saxifraga	4.86	0.69	4.17
Herbs		Epidendroideae	0.32	0.00	0.32
Herbs	Poppy subfamily	Papaveroideae	0.34	0.34	0.00
Herbs	Buttercup subfamily	Ranunculoideae	0.11	0.11	0.00
Herbs	Buttercup tribe	Ranunculeae	0.16	0.00	0.16
Herbs		Thalictroideae	48.24	0.00	48.24
Herbs	Meadow-rues	Thalictrum	10.08	0.00	10.08
Herbs	Core clubmoss family	Lycopodiaceae	0.27	0.00	0.27
Trees/shrubs and herbs	Gymnosperms	Acrogymnospermae	0.22	0.22	0.00
Trees/shrubs and herbs		Gunneridae	1.41	0.47	0.94
Trees/shrubs and herbs		campanulids	0.31	0.00	0.31
Trees/shrubs and herbs		Asterales	0.57	0.00	0.57
Trees/shrubs and herbs	Daisy family	Asteraceae	2.71	1.09	1.63
Trees/shrubs and herbs		Asteroideae	0.50	0.25	0.25
Trees/shrubs and herbs	Chamomile tribe	Anthemideae	0.55	0.00	0.55
Trees/shrubs and herbs	Dead-nettle family	Lamiaceae	0.78	0.00	0.78
Trees/shrubs and herbs	Plantain family	Plantaginaceae	0.92	0.92	0.00
Trees/shrubs and herbs	Knotweed family	Polygonaceae	19.33	2.21	17.12
Trees/shrubs and herbs	Knotweed tribe	Polygoneae	0.73	0.00	0.73
Trees/shrubs and herbs	Milkvetches	Astragalus	0.61	0.00	0.61
Trees/shrubs and herbs	Flax family	Linaceae	0.65	0.00	0.65
Trees/shrubs and herbs	Mustard order	Brassicales	2.60	0.00	2.60
Trees/shrubs and herbs		**Alyssum**	**1.69**	**0.00**	**1.69**
Trees/shrubs and herbs	Mallow subfamily	Malvoideae	0.56	0.00	0.56
Trees/shrubs and herbs		Saxifragales	0.77	0.00	0.77
Trees/shrubs and herbs		Asparagales	0.08	0.00	0.08
Trees/shrubs and herbs	Buttercup order	Ranunculales	4.51	0.31	4.20
Trees/shrubs and herbs	Buttercup family	Ranunculaceae	1.26	0.11	1.15
Grasses and relatives	Grass order	Poales	7.03	0.00	7.03
Grasses and relatives	Sedges	Cyperaceae	43.18	3.08	40.10
Grasses and relatives		Cyperoideae	11.81	1.69	10.12
Grasses and relatives	True sedges	Carex	230.06	9.11	220.95
Grasses and relatives	Rushes	Juncaceae	1.04	0.00	1.04
Grasses and relatives		Poeae	1.11	0.00	1.11
Grasses and relatives		Aveninae	0.29	0.00	0.29
Grasses and relatives		Poeae Chloroplast Group 1 (Aveneae type)	0.17	0.17	0.00
Grasses and relatives		Agrostidinae	0.16	0.00	0.16
Grasses and relatives	Ryegrass subtribe	Loliinae	0.91	0.00	0.91
Grasses and relatives	Fescues	Festuca	1.97	0.18	1.79
Grasses and relatives	Meadow grass subtribe	Poinae	0.48	0.00	0.48
Grasses and relatives	Meadow grasses	Poa	0.82	0.00	0.82
Grasses and relatives	Barley subtribe	Hordeinae	0.12	0.12	0.00
Grasses and relatives	Barlies	Hordeum	0.26	0.26	0.00
Grasses and relatives		**Panicoideae**	**0.54**	**0.00**	**0.54**
Ferns	Horsetails	Equisetum	0.05	0.00	0.05
Bryophytes	Mosses	Bryophyta	1.94	0.00	1.94
Bryophytes		Bryophytina	7.81	0.00	7.81
Bryophytes		Bryopsida	17.86	0.00	17.86
Bryophytes		Herzogiella	4.88	4.88	0.00
Bryophytes	Hair moss family	Polytrichaceae	1.35	0.00	1.35
Bryophytes	Liverworts	Marchantiophyta	1.06	0.00	1.06
Bryophytes		Jungermanniidae	1.88	1.88	0.00
Mixed		Tracheophyta	0.52	0.17	0.35

Table 4.52: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF049. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Group	Common name	Taxon	Total	361	295
			Ingroup	15.1581111	10.3332314	4.82487968
Freshwater aquatics	Dipterans		Cecidomyiinae	8.10810811	8.10810811	0
Mixed aquatics	Fishes		Osteoglossocephalai	0.82647233	0	0.82647233
Mixed aquatics	Hydrozoans	Thecate hydroids	Leptothecata	2.43902439	0	2.43902439
Terrestrial	Mammals	Cat-like carnivorans	Feliformia	0.33723022	0	0.33723022
Terrestrial	Birds	Landfowl	Galliformes	0.76757753	0.76757753	0
Terrestrial	**Moths**		**Sesamia**	**1.02564103**	**0**	**1.02564103**
Mixed	Vertebrates	Jawed vertebrates	Gnathostomata	0.39302343	0.19651171	0.19651171

Embryophyta

The Embryophyta present a curious picture. Both samples contain mostly the willow group, which is found across most cores as a minor accompaniment to *Zostera*. But here, aquatics are nearly absent, the exception being two very small signals from the freshwater Menyanthaceae and mixed Molinieae (common reed tribe) in sample 295. Despite the absence of saltwater aquatics, 295 contains several halophytes and Chenopodiaceae, which can be interpreted as such.

All woody taxa besides the willow group are oddities. 361 has *Picea* (spruces; see ELF045), its parent family Pinidae, and the heather family Ericaceae. 295 contains Dryadoideae, of which the only European species is *Dryas octopetala* (see ELF031A). All are associated with cool upland habitats, as is the herb *Androsace* (rock jasmine), an alpine that is not currently native to Britain, but was during the last glacial (Bell 1969). The willow group has mostly been interpreted as floodplain woodland in these cores, but *Salix* and *Populus* can also occur in boreal or tundra environments (Pividori *et al.* 2016, Anderson *et al.* 1994).

Finally, there are two groups of taxa that are often present in cores at low frequency but are not usually ecologically informative: *Carex* (true sedges) and its parent family Cyperaceae, and the bryophytes (liverworts, mosses, and hornworts). Both are unusually frequent in this core, suggesting environments more suitable for sedges and bryophytes than we have previously seen. While occurring widely, sedges most often dominate montane and wet habitats, including arctic tundra (Ball and Reznicek 2008). Mosses are similarly cosmopolitan with higher prevalence in harsh environments where higher plants compete less effectively (Bliss 2000). They are consistent with a surprisingly cool climate.

With the exception of *Picea*, none of these taxa are particularly well represented in genomic databases, so they are not likely candidates for mis-assignment. Similarly, only Pinidae was also found in negative controls. There were too few reads to age-authenticate each taxon individually, but both samples in ELF049 passed (Chapter 5: Authentication results). This excludes significant modern contamination. The current evidence supports a genuine cold signal in this core.

Metazoa

Metazoa reads are limited and inconclusive, although freshwater and mixed aquatics have the most biogenomic mass (saltwater are again absent).

4.26.3 Pianka scores

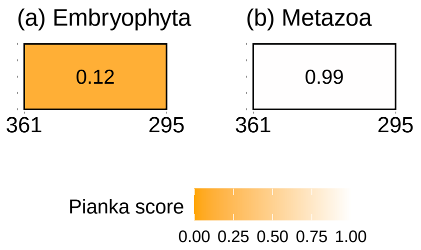


Figure 4.126: Pianka similarity scores between adjacent samples in ELF049 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores. Comparisons involving empty samples have no score.

The Embryophyta and Metazoa scores are nearly at opposite ends of the scale. The two samples appear highly similar in Embryophyta, despite the difference in read counts, whereas there is hardly any overlap in the few Metazoa taxa.

4.26.4 Summary

ELF049 contains mostly the willow group, which are associated with wet habitats and have in other cores been taken to indicate floodplain near the river that formed the palaeochannel. However, several other taxa present a surprisingly cool and upland profile, particularly as ELF049 was taken in the channel mouth. The read counts of most are very low, but the variety of taxa and their absence or reduction (sedges) across other cores is intriguing, and willows can be associated with these habitats as well as temperate floodplains (Boulanger-Lapointe *et al.* 2016). The samples in ELF049 may date from before most other cores, when the climate was cooler and the river mouth was still some distance away.

4.27 ELF046A

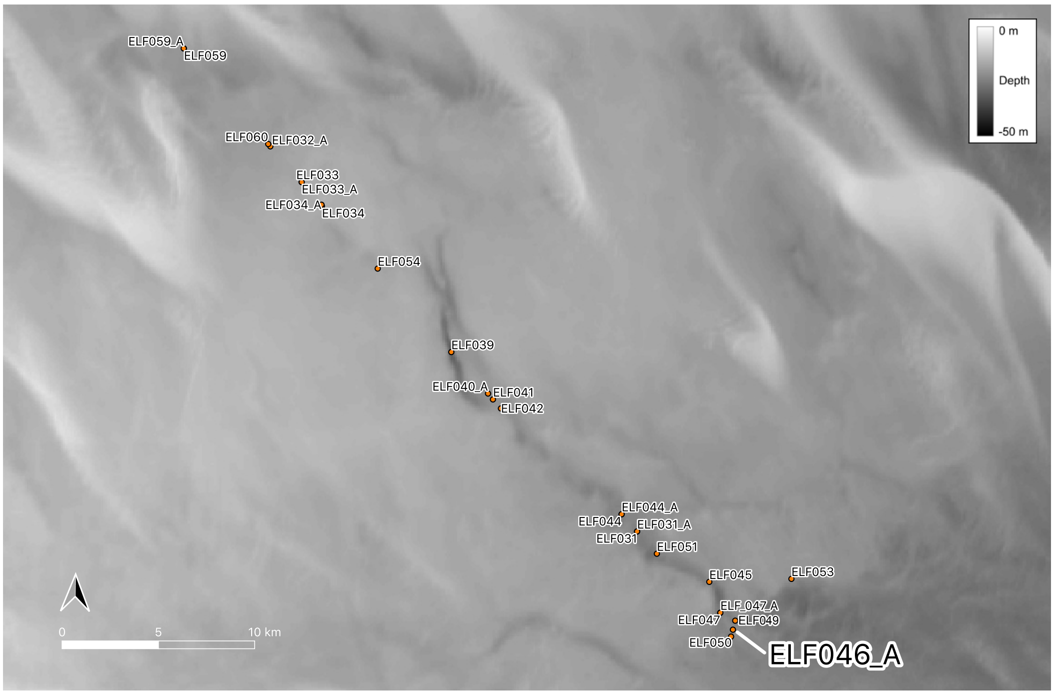


Figure 4.127: Bathymetry map of the palaeochannel transect highlighting ELF046A. Darker grey indicates greater depth.

ELF046A is within 1 km of ELF049 and ELF050. One sample was taken. Unfortunately, the initial sequencing run returned no data, and the replicate was associated with excessive contamination and discarded. Therefore, there is no usable sedaDNA data for ELF046A.

4.28 ELF050

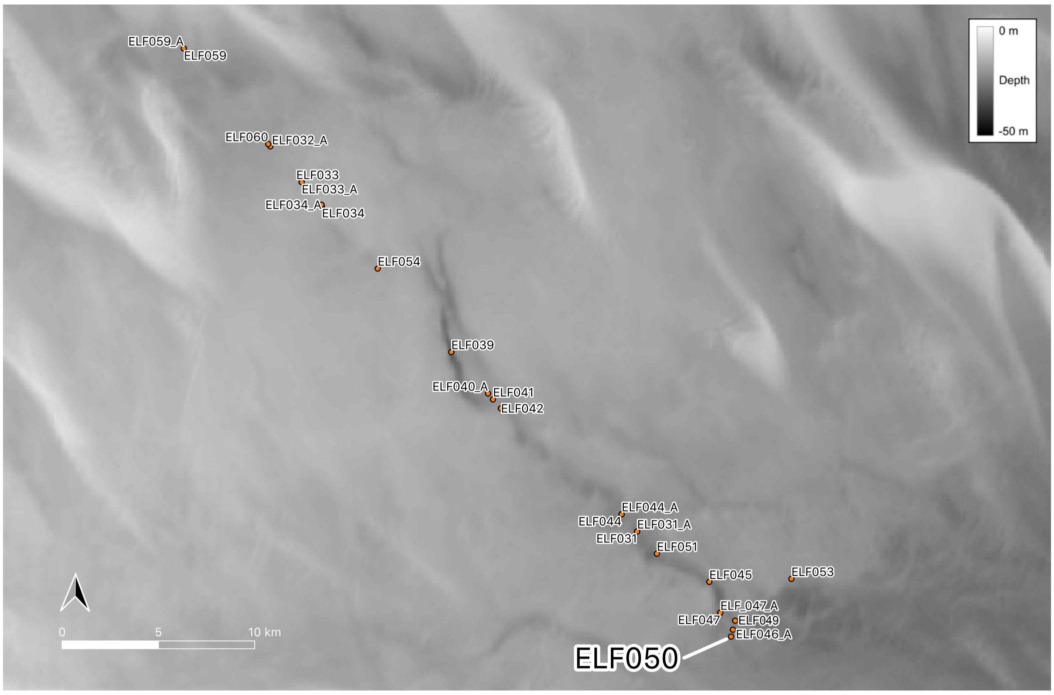


Figure 4.128: Bathymetry map of the palaeochannel transect highlighting ELF050. Darker grey indicates greater depth.

ELF050 is the furthest southeast of these cores, approximately level with the headlands at the mouth of the palaeochannel.

4.28.1 Read counts

The read counts in ELF050 are very low for both Embryophyta and Metazoa. Note that the single Embryophyta read in sample 250 was assigned to the non-European *Citrullus lanatus* (watermelon), so is not included in the taxonomic results, hence the uncoloured label.

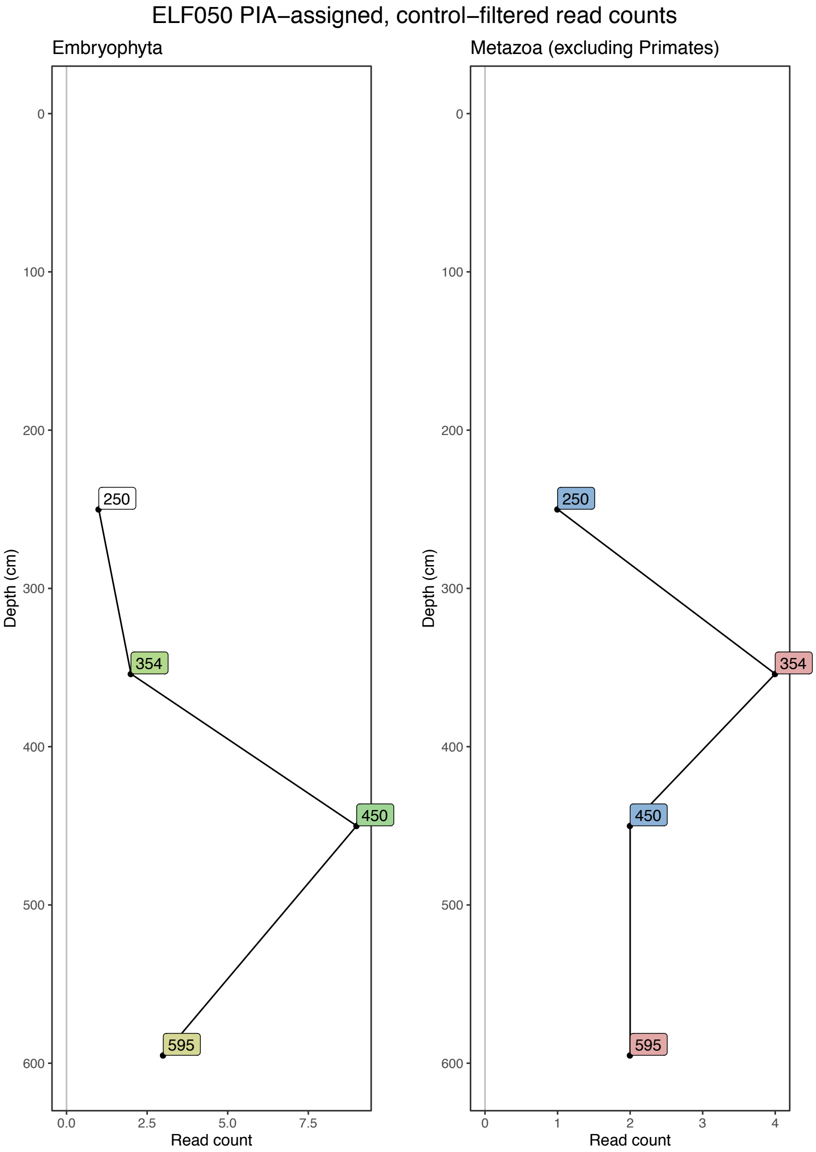
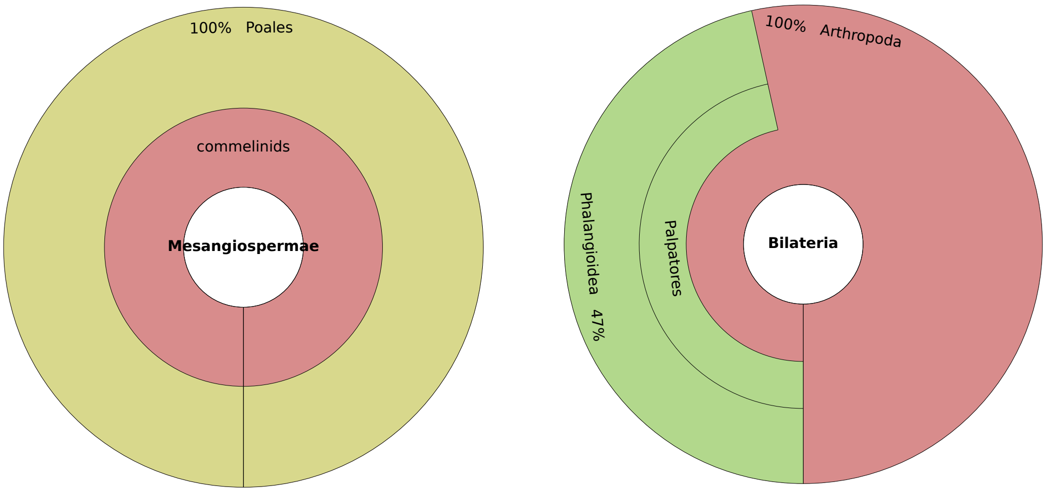


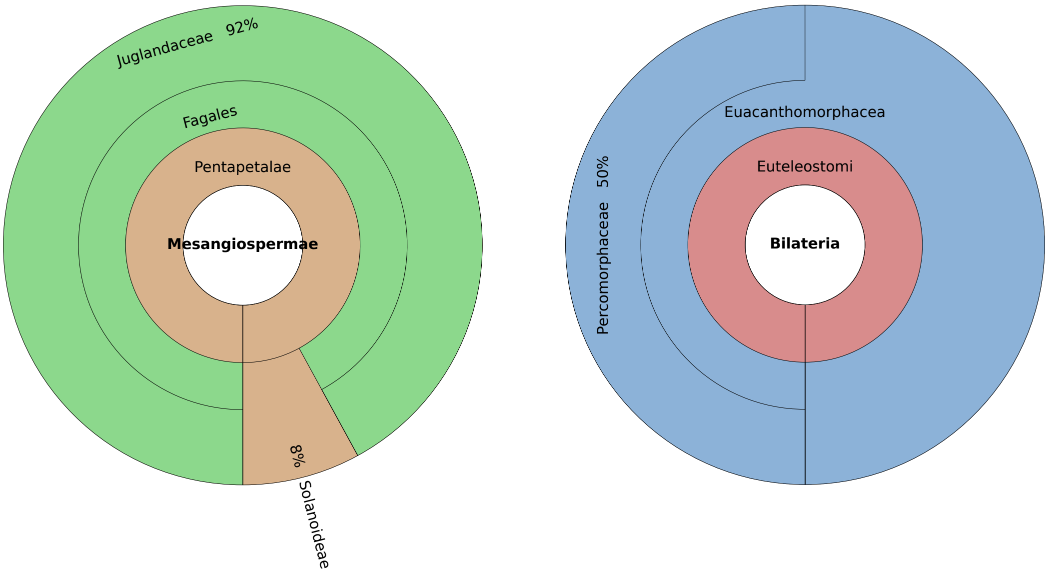
Figure 4.129: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF050. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.28.2 Taxonomic profiles

ELF050_595



ELF050_450



ELF050_354

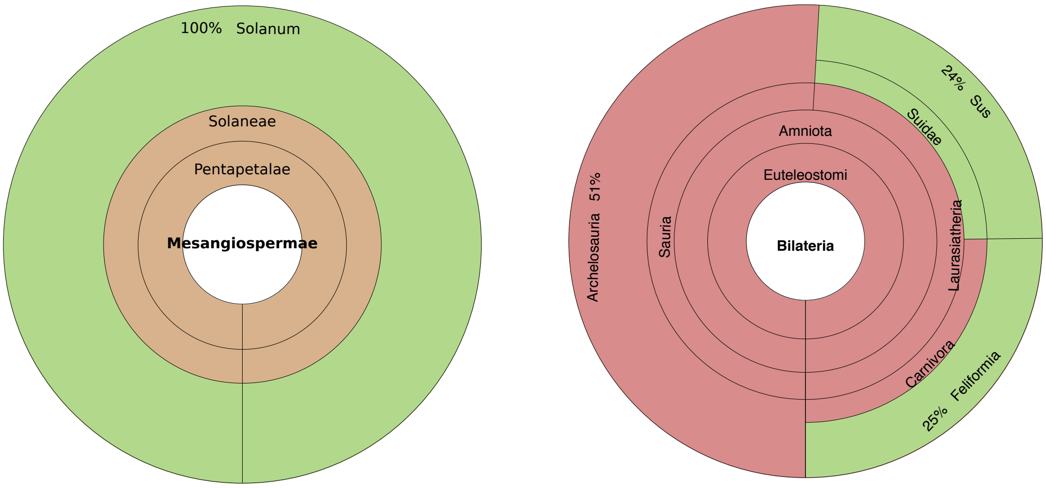


Figure 4.130: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF050, samples 595, 450, and 354. Continued in figure 4.131. See figure 4.4 for colour key.

ELF050_250

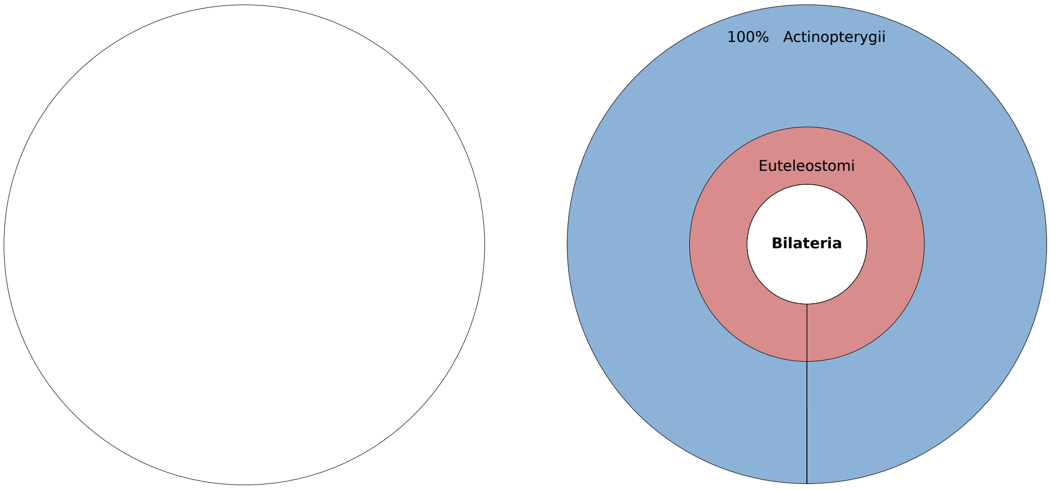


Figure 4.131: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF050, sample 250. See figure 4.4 for colour key.

Table 4.53: Biogenomic masses of European taxa in Embryophyta from ELF050. Samples are in cm. Taxa in bold are not native to Great Britain.

Ecological category	Common name	Taxon	Total	595	450	354	250
		Ingroup	19.19	1.87	14.06	1.26	2.00
Trees/shrubs	**Walnut family**	**Juglandaceae**	**5.46**	**0.00**	**5.46**	**0.00**	**0.00**
Herbs	Nightshades	Solanum	0.71	0.00	0.00	0.71	0.00
Trees/shrubs and herbs	Nightshade subfamily	Solanoideae	0.47	0.00	0.47	0.00	0.00
Grasses and relatives	Grass order	Poales	0.29	0.29	0.00	0.00	0.00

Table 4.54: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF050. Samples are in cm.

Ecological category	Group	Common name	Taxon	Total	595	450	354	250
			Ingroup	5.94	0.91	2.11	2.13	0.78
Mixed aquatics	Fishes	Ray-finned fishes	Actinopterygii	0.78	0.00	0.00	0.00	0.78
Mixed aquatics	Fishes		Euacanthomorphacea	1.06	0.00	1.06	0.00	0.00
Mixed aquatics	Fishes		Percomorphaceae	1.06	0.00	1.06	0.00	0.00
Terrestrial	Mammals	Cat-like carnivorans	Feliformia	0.34	0.00	0.00	0.34	0.00
Terrestrial	Mammals	Pigs	Sus	0.32	0.00	0.00	0.32	0.00
Terrestrial	Harvestmen		Phalangioidea	0.42	0.42	0.00	0.00	0.00
Mixed	Vertebrates		Archelosauria	0.68	0.00	0.00	0.68	0.00
Mixed	Arthropods	Arthropods	Arthropoda	0.49	0.49	0.00	0.00	0.00

Embryophyta

The four Embryophyta taxa are all terrestrial, but further interpretation may be difficult to justify. Juglandaceae (see ELF059A) would require careful interpretation in the best of situations. *Solanum* (nightshades) and its subfamily Solanoideae have not been discussed before, despite appearing with some regularity, because *Solanum lycopersicum* is the cultivated tomato, which is highly likely to be over-represented in reference databases. The most reliable Embryophyta taxon in this core is Poales in sample 595.

Metazoa

Unusually, ELF050 has more Metazoa taxa than Embryophyta, although read counts are very low and *Sus* in sample 354 should be treated with caution. Two samples contain mixed aquatic taxa, suggesting some aquatic influence not seen in Embryophyta. Beyond this, there is little to interpret.

4.28.3 Pianka scores

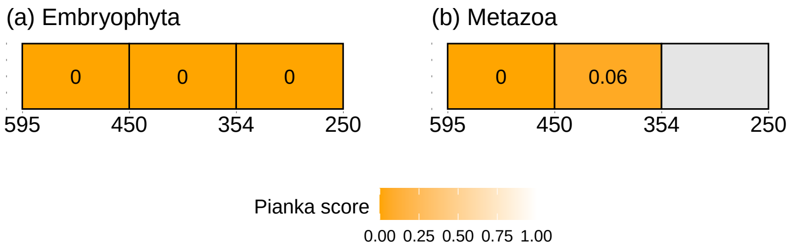


Figure 4.132: Pianka similarity scores between adjacent samples in ELF050 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

As expected from such low read counts, both Embryophyta and Metazoa have similarly low Pianka scores. There is not enough data to infer environmental change from the scores.

4.28.4 Summary

ELF050 shows a very small range of both terrestrial and aquatic taxa. There is too little data for further environmental reconstruction.

4.29 ELF053

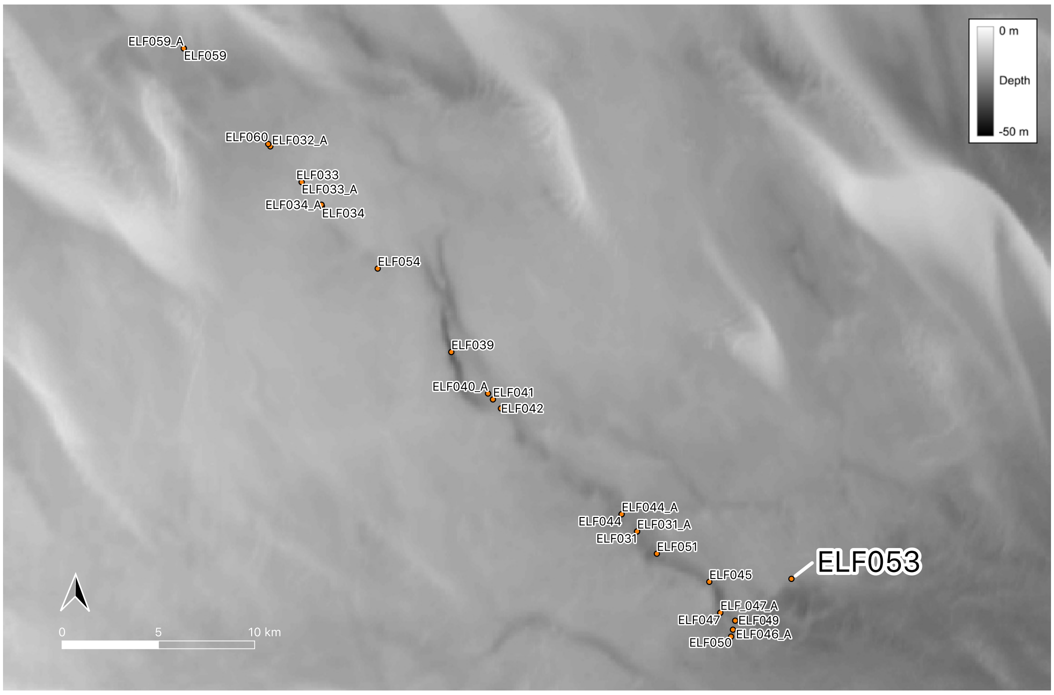


Figure 4.133: Bathymetry map of the palaeochannel transect highlighting ELF053. Darker grey indicates greater depth.

ELF053 is somewhat removed from the main transect. It was taken at the edge of the headland east of the palaeochannel mouth with no other cores in the immediate area.

4.29.1 Read counts

Read counts are very low, even for samples containing data from both sequencing runs. The results from the replicate of sample 275 (†) were discarded because of excessive contamination in the negative control. It is unlikely to have been a significant loss.

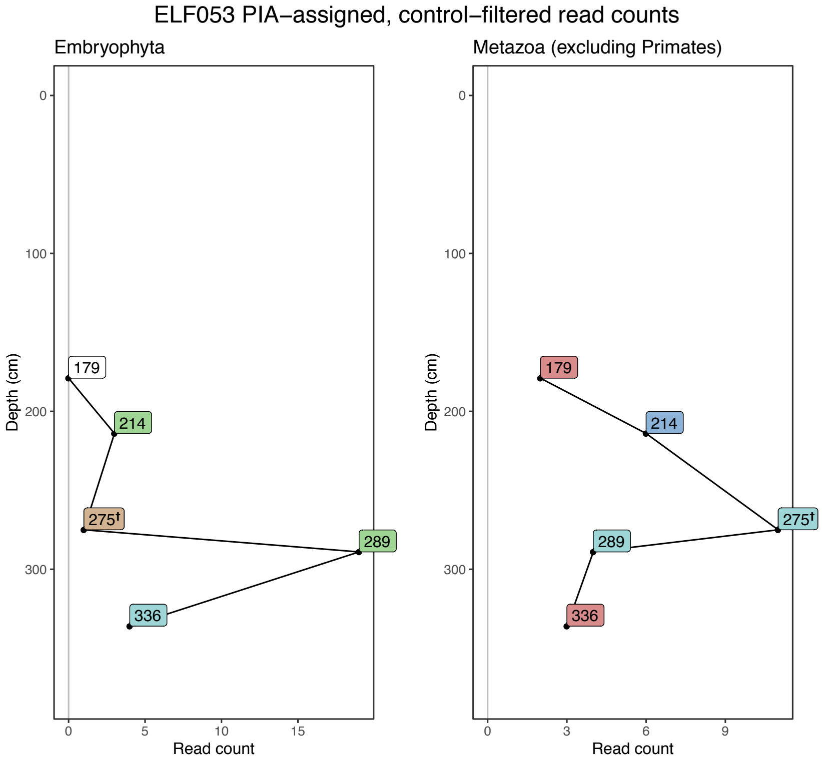
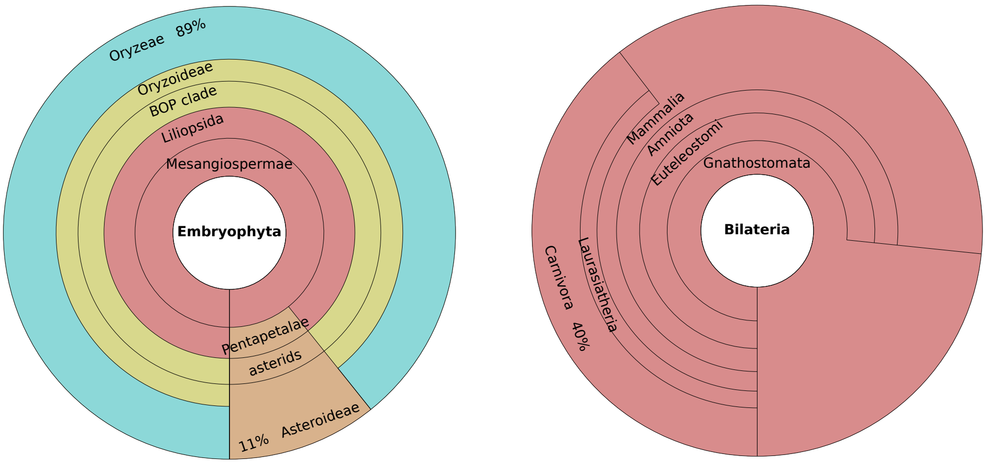


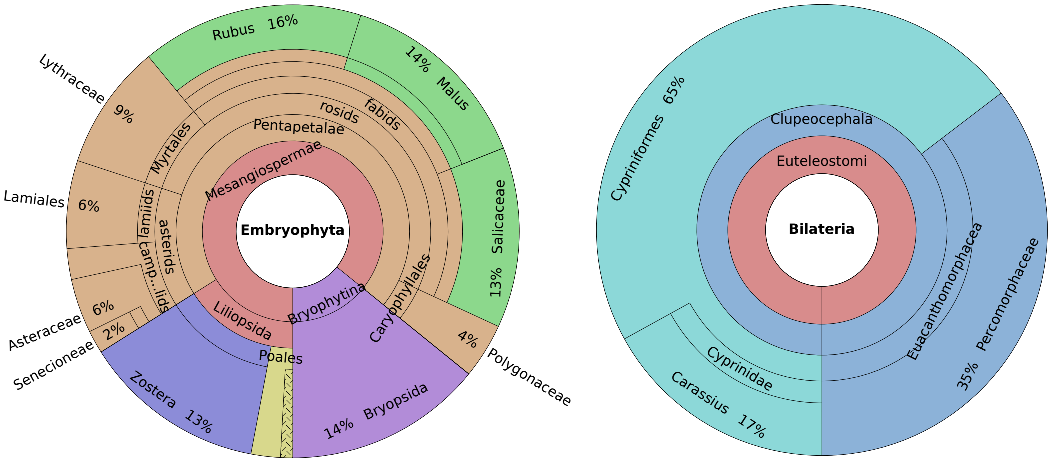
Figure 4.134: Read counts for Embryophyta and Metazoa (excluding Primates) for each sample (cm) in ELF053. Counts are post-PIA and filtered by negative controls, but may include reads assigned to non-European taxa. Sample 275 (†) does not include data from the replicate sequencing run because it was associated with excessive contamination. Sample labels are coloured by most frequent ecological category after GSA (according to results below; see key in figure 4.4).

4.29.2 Taxonomic profiles

ELF053_336



ELF053_289



ELF053_275

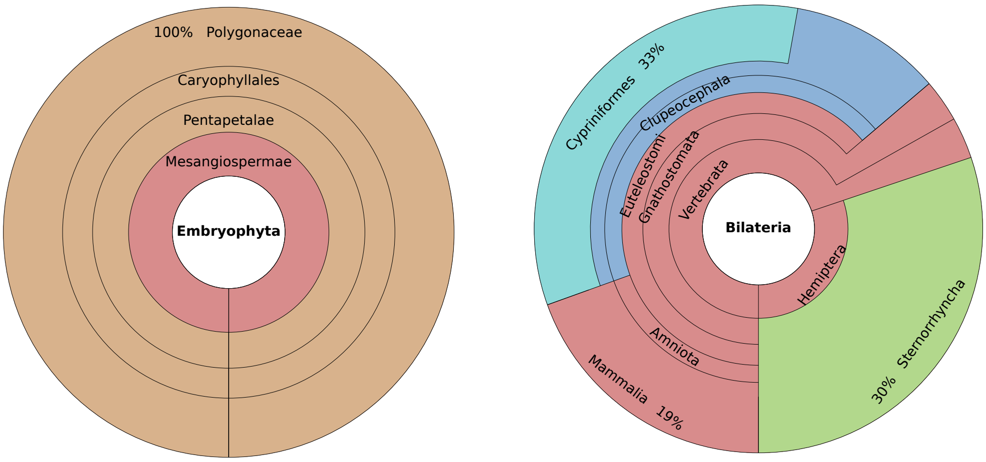
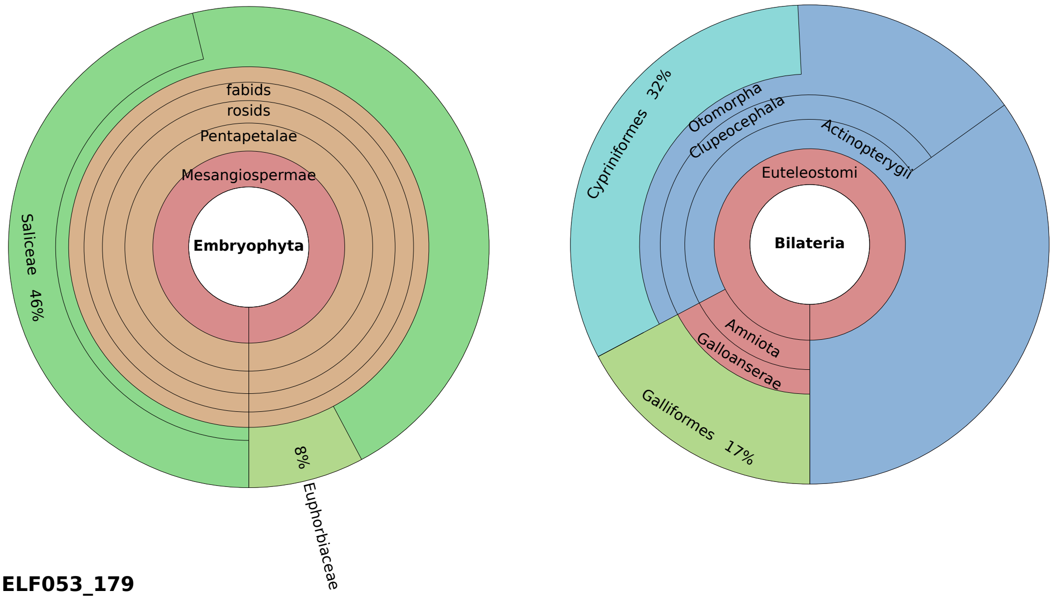


Figure 4.135: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF033, samples 336, 289, and 275. Sample 275 (†) does not include data from the replicate sequencing run because it was associated with excessive contamination. Continued in figure 4.136. See figure 4.4 for colour key.

ELF053_214



ELF053_179

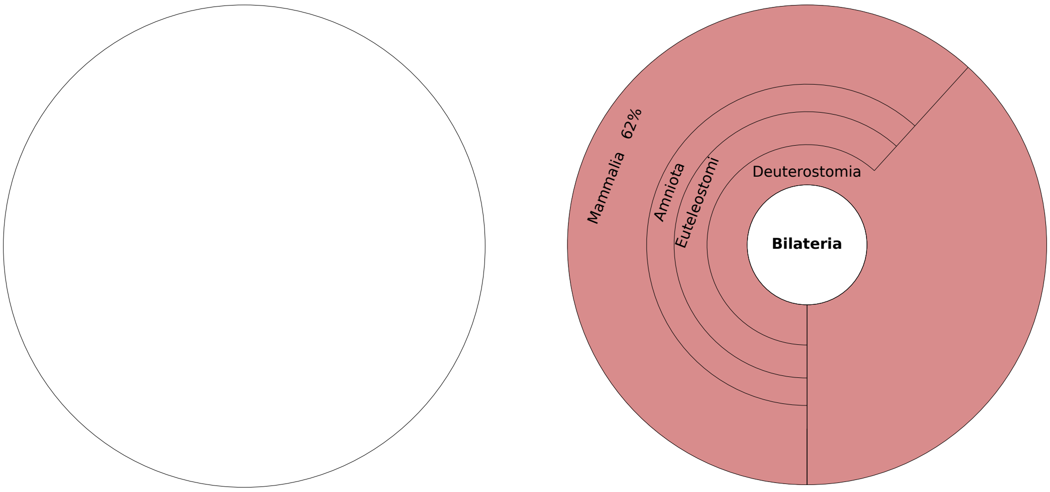


Figure 4.136: Most frequent European taxa from Embryophyta (left) and Metazoa (excluding Primates; right) in ELF053, samples 214 and 179. See figure 4.4 for colour key.

Table 4.55: Biogenomic masses of European taxa in Embryophyta from ELF053. Samples are in cm. Sample 275 (†) does not include data from the duplicate sequencing run because it was associated with excessive contamination.

Ecological category	Common name	Taxon	Total	336	289	275†	214	179
		Ingroup	24.72	3.16	17.08	0.55	3.93	0.00
Freshwater aquatics	Rice tribe	Oryzeae	2.08	2.08	0.00	0.00	0.00	0.00
Salt/brackish aquatics	Eelgrasses	Zostera	1.83	0.00	1.83	0.00	0.00	0.00
Trees/shrubs	Willow family	Salicaceae	3.61	0.00	1.81	0.00	1.81	0.00
Trees/shrubs	Willow tribe	Saliceae	1.82	0.00	0.00	0.00	1.82	0.00
Trees/shrubs	Apples	Malus	1.98	0.00	1.98	0.00	0.00	0.00
Trees/shrubs	Blackberry and relatives	Rubus	2.22	0.00	2.22	0.00	0.00	0.00
Herbs	Spurge family	Euphorbiaceae	0.31	0.00	0.00	0.00	0.31	0.00
Trees/shrubs and herbs		campanulids	0.31	0.00	0.31	0.00	0.00	0.00
Trees/shrubs and herbs	Daisy family	Asteraceae	0.54	0.00	0.54	0.00	0.00	0.00
Trees/shrubs and herbs		Asteroideae	0.25	0.25	0.00	0.00	0.00	0.00
Trees/shrubs and herbs	Ragwort tribe	Senecioneae	0.23	0.00	0.23	0.00	0.00	0.00
Trees/shrubs and herbs	Dead-nettle order	Lamiales	0.88	0.00	0.88	0.00	0.00	0.00
Trees/shrubs and herbs	Knotweed family	Polygonaceae	1.10	0.00	0.55	0.55	0.00	0.00
Trees/shrubs and herbs	Loosestrife family	Lythraceae	1.26	0.00	1.26	0.00	0.00	0.00
Grasses and relatives	Grass order	Poales	0.29	0.00	0.29	0.00	0.00	0.00
Grasses and relatives	Barley subtribe	Hordeinae	0.12	0.00	0.12	0.00	0.00	0.00
Bryophytes		Bryopsida	1.98	0.00	1.98	0.00	0.00	0.00

Table 4.56: Biogenomic masses of European taxa in Metazoa (excluding Primates) from ELF053. Samples are in cm. Sample 275 (†) does not include data from the duplicate sequencing run because it was associated with excessive contamination.

Ecological category	Group	Common name	Taxon	Total	336	289	275†	214	179
			Ingroup	15.18	0.84	2.98	6.40	4.45	0.51
Freshwater aquatics	Fishes	Carp order	Cypriniformes	4.98	0.00	1.42	2.13	1.42	0.00
Freshwater aquatics	Fishes	Crucian carps	Carassius	0.51	0.00	0.51	0.00	0.00	0.00
Mixed aquatics	Fishes	Ray-finned fishes	Actinopterygii	1.55	0.00	0.00	0.00	1.55	0.00
Mixed aquatics	Fishes		Percomorphaceae	1.06	0.00	1.06	0.00	0.00	0.00
Mixed aquatics	Fishes		Otomorpha	0.71	0.00	0.00	0.00	0.71	0.00
Mixed aquatics	Fishes	Carp subcohort	Ostariophysi	0.70	0.00	0.00	0.70	0.00	0.00
Terrestrial	Birds	Landfowl	Galliformes	0.77	0.00	0.00	0.00	0.77	0.00
Terrestrial	Hemipterans	Aphids and relatives	Sternorrhyncha	1.93	0.00	0.00	1.93	0.00	0.00
Mixed	Animals	Deuterostomes	Deuterostomia	0.19	0.00	0.00	0.00	0.00	0.19
Mixed	Vertebrates	Vertebrates	Vertebrata	0.19	0.00	0.00	0.19	0.00	0.00
Mixed	Vertebrates	Jawed vertebrates	Gnathostomata	0.39	0.20	0.00	0.20	0.00	0.00
Mixed	Mammals	Mammals	Mammalia	1.87	0.31	0.00	1.25	0.00	0.31
Mixed	Mammals	Carnivorans	Carnivora	0.33	0.33	0.00	0.00	0.00	0.00

Both the Embryophyta and Metazoa results show a scatter of aquatic and terrestrial taxa with no clear pattern across samples. The Oryzeae and Hordeinae in samples 336 and 289 do contain plausible British native grasses, but their proximity to over-represented domesticates casts some doubt on their assignment. Similarly, most of the Metazoa taxa are vertebrates, which may be more a reflection of the reference database than the environment. It is difficult to interpret such limited data.

4.29.3 Pianka scores

The Embryophyta Pianka scores are very low, as would be expected from such small read counts. The Metazoa scores are instead rather high, which could be explained by the inter-relatedness of most of the Metazoa taxa (fishes and higher vertebrate taxa), despite little direct overlap.

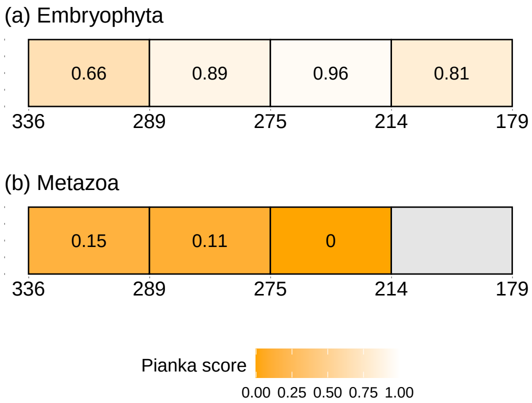


Figure 4.137: Pianka similarity scores between adjacent samples in ELF033 for (a) Embryophyta and (b) Metazoa. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores.

4.29.4 Summary

ELF053 was taken from a headland outside of the palaeochannel, a location where more terrestrial environments than in most other cores might be expected. Unfortunately, the low read counts in ELF053 makes reconstruction difficult, although a lack of data may itself be associated with a terrestrial environment.

4.30 Higher-level analyses

4.30.1 Data yield and dominant ecological category

It appears from the taxonomic profiles that samples with the largest data yields contain mostly reads from *Zostera* or higher taxa, although there are several high-yield samples dominated by grasses. The boxplot in figure 4.138 compares ingroup Embryophyta read counts per sample by dominant ecological category after GSA. These are the same numbers used in the read count plots for each core.

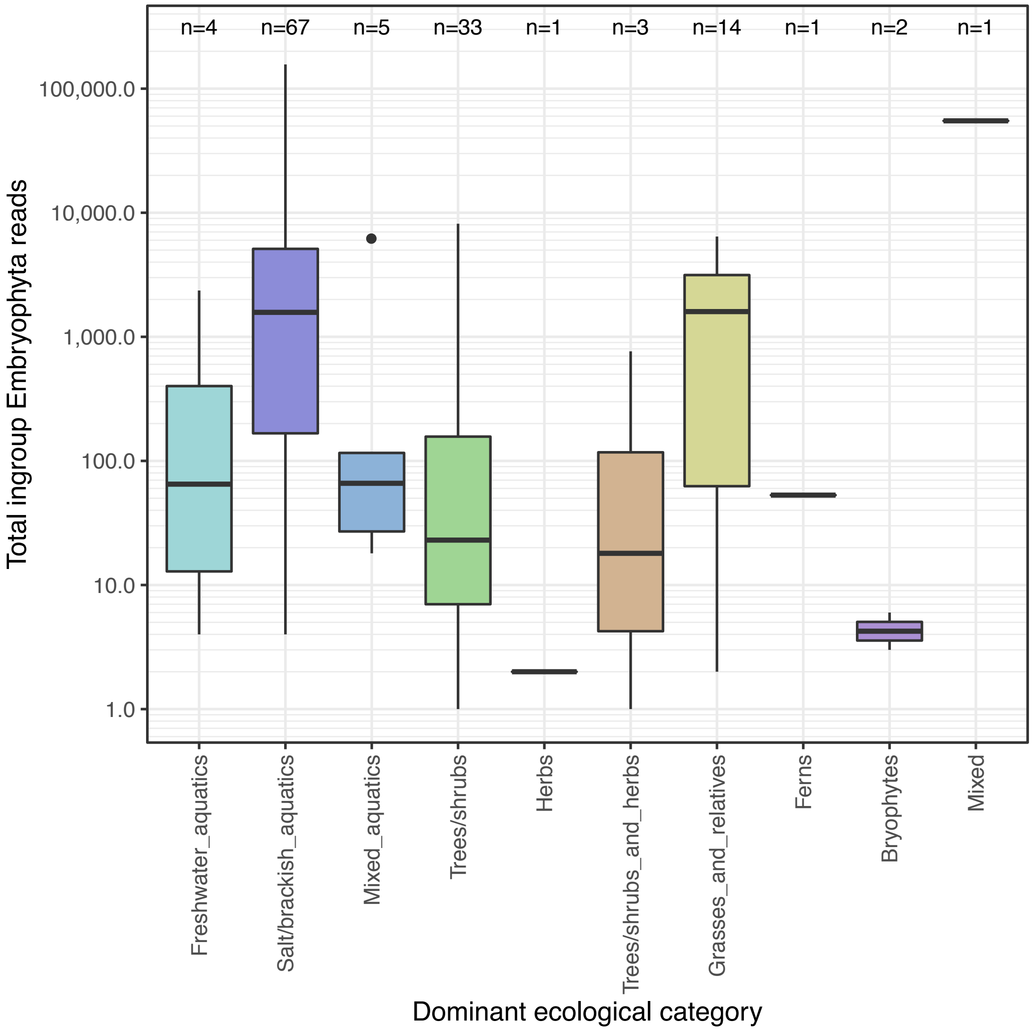


Figure 4.138: Boxplot of data yield by dominant ecological category. Boxes show the IQR and lines 1.5*IQR. Yield is total ingroup Embryophyta read counts per sample; note that the y-axis is in log base 10. Dominant ecological category is that with the highest biogenomic mass. Samples with 0 reads (therefore no category) are not shown. n is the number of samples in each group.

The samples with highest yields are indeed in the salt/brackish aquatics group, which almost entirely consists of *Zostera*. The mixed category is next, but this is not particularly informative as it contains mostly higher taxa and is only represented by a single sample. Quite some distance behind (note the log scale), the next three categories with high-yield samples are mixed aquatics, trees/shrubs, and grasses and relatives. However, the median yields for mixed aquatics and trees/shrubs are less than 100 reads. Instead, the most consistently high-yielding category after salt/brackish aquatics is grasses and relatives. The taxonomic profiles suggested that many of these grass-dominated samples would be best interpreted as reed beds, with an almost ubiquitous *Zostera* signal further suggesting saltmarsh reed beds. The association of high-yield samples with certain ecological categories could therefore be explained by the beneficial effect of salt on DNA preservation (Lindahl and Nyberg 1972, Kistler *et al.* 2017).

4.30.2 The ratio of higher to lower taxa and data yield

Another potential trend noted in the taxonomic profiles was of an increasingly large proportion of higher taxa in *Zostera*-dominated samples of greater data yield, such as ELF059A_250 and ELF060_350. These higher taxa (in particular Alismatales, Liliopsida, and Mesangiospermae) were rarely noted in other samples despite being equally related to many other common plants. Is this trend towards higher taxa genuine, and if so, is it linked to *Zostera* or more generally to data yield?

Taxa in the Embryophyta data were coded by rank (species, family, etc.) according to the NCBI taxonomy database (Federhen 2012). However, many taxa do not have assigned ranks. The following taxa found at high frequency in this data were therefore manually added to the rank hierarchy at the appropriate levels:

- Embryophyta
- Tracheophyta
- Euphyllophyta
- Spermatophyta
- Mesangiospermae
- Liliopsida
- PACMAD clade

Remaining unassigned taxa were coded "uranked". Samples with zero reads were excluded.

Figure 4.139 plots reads assigned to each rank as a proportion of the total. Samples are ordered from smallest to largest read count. The proportion of reads assigned to higher taxa (darker colours) does appear to increase with data yield. Lower-yield samples have high variation, but contain perhaps 50% genera until around the median sample (ELF039_384), where genera quickly shrink to ∼10% of reads and orders climb to over 75%. Orders are then gradually replaced by Liliopsida and Mesangiospermae, which together occupy around 50% of the highest-yield samples. Interestingly, the very highest taxa (Tracheophyta, Embryophyta) are rare across all samples, with if anything greater prevalence in those with low yields.

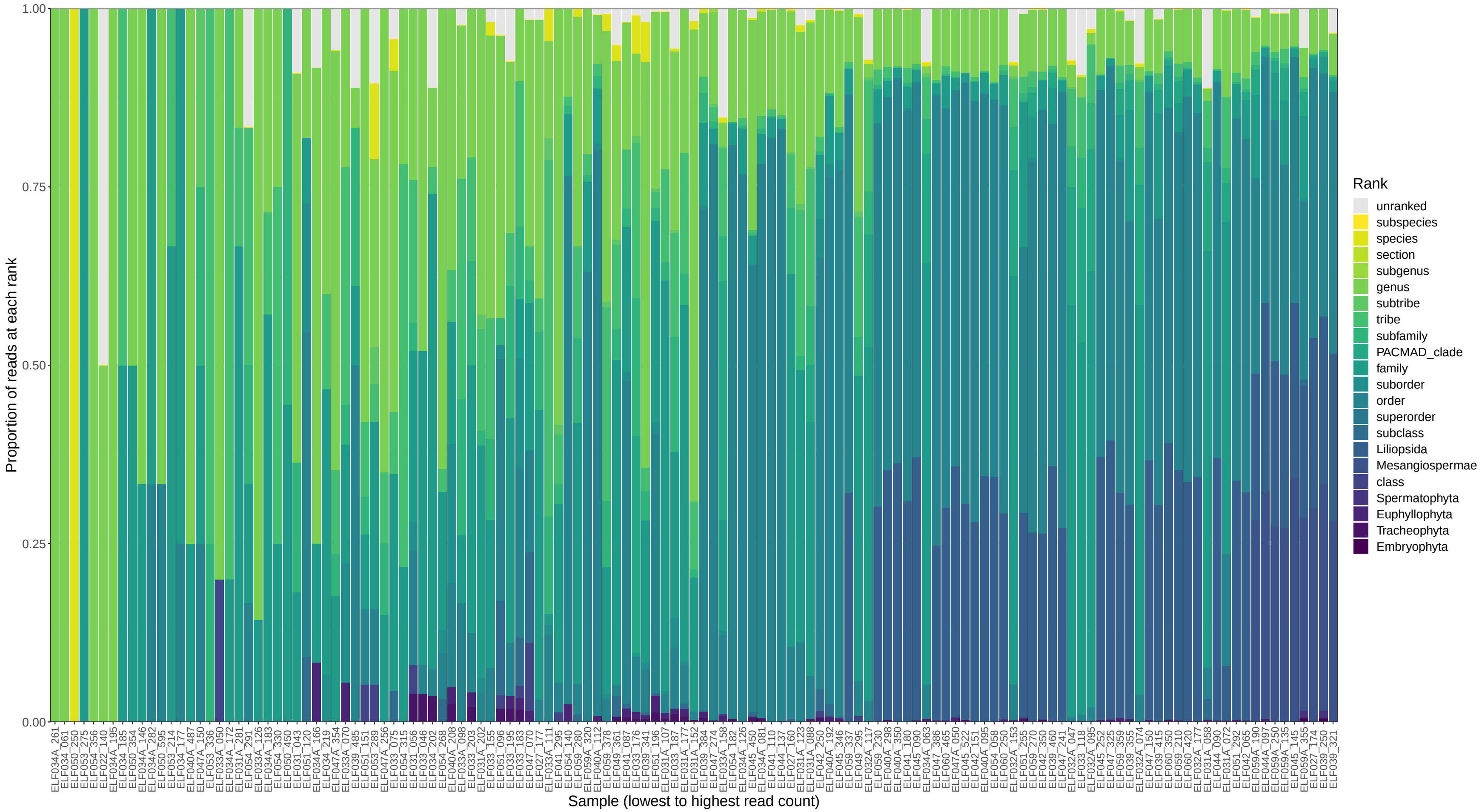


Figure 4.139: Proportions of Embryophyta reads by rank for all samples. Selected high-frequency unranked taxa were added to the hierarchy manually. Samples are ordered from smallest to largest read count.

This complicated picture was simplified by dividing ranks into two categories: "higher" (suborder and above) or "lower" (family and below). Family was chosen as the boundary because this was often the highest useful taxon for ecological reconstruction. Plotted in figure 4.141 (page 287), this data confirms the trend of figure 4.139: individual samples vary, but the proportion of higher taxa increases with data yield up to ∼90%.

This relationship was quantified by calculating the ratio of reads from higher to lower taxa for each sample. These ratios are plotted against total reads (natural-log-transformed) in figure 4.140. Many samples have very low ratios, with < 1 indicating more reads from lower taxa than higher. Nevertheless, those with higher ratios (with more higher taxa) show a positive relationship with read count. Pearson's test returned r=0.688 (95% CI [0.586, 0.768]; p<0.01), supporting a moderate positive correlation.

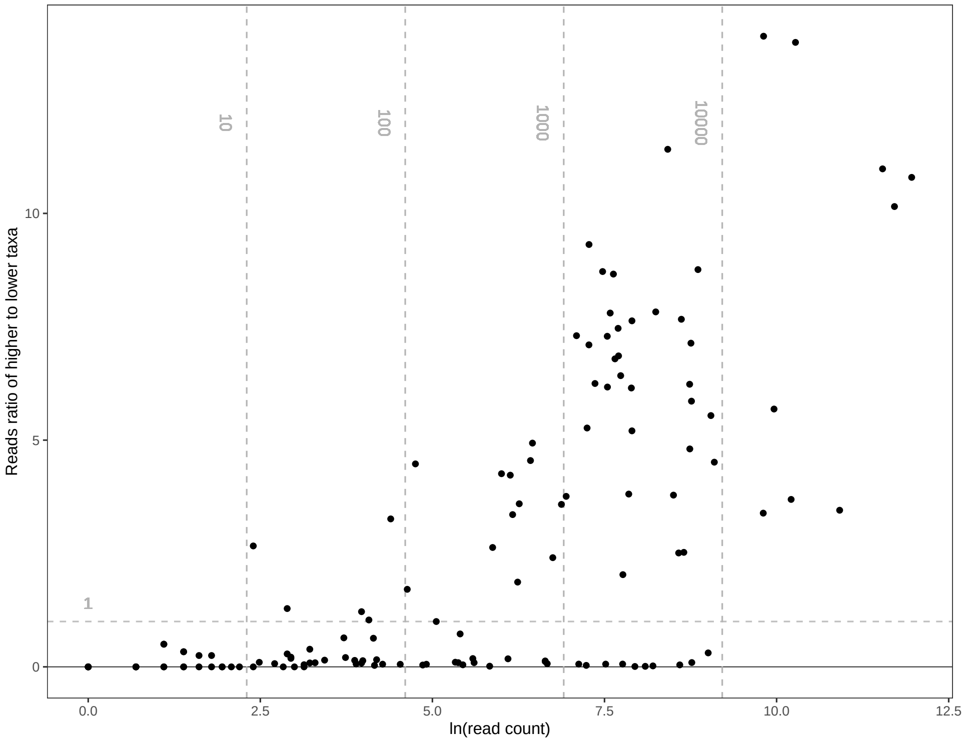


Figure 4.140: Ratios of higher to lower taxa against total Embryophyta reads. Total reads are natural-log transformed; vertical dashed lines mark raw masses for easier interpretation. The horizontal line marks a ratio of 1, above which samples have more reads assigned to higher taxa than to lower.

The pattern observed in the taxonomic profiles therefore seems to be genuine, at least for this data set. However, there are high-yield samples in figure 4.141 lacking higher taxa. These exceptions suggest that it may be a feature of *Zostera*-dominated samples rather than a high yield per se (subsection 4.30.1 above demonstrated that most high-yield samples contain mostly *Zostera*). Looking to the right of figure 4.141, we see that samples ELF031A_072 and 058; ELF032A_095, 047, and 153; and ELF033_118 have among the greatest data yields but very low proportions of higher taxa, breaking the pattern. What distinguishes these samples is that they contain mostly trees/shrubs (*Salix* and Salicaceae) or grasses (the PACMAD and BOP clades and Poaceae) instead of *Zostera*. If *Zostera* is minor, the pattern disappears.

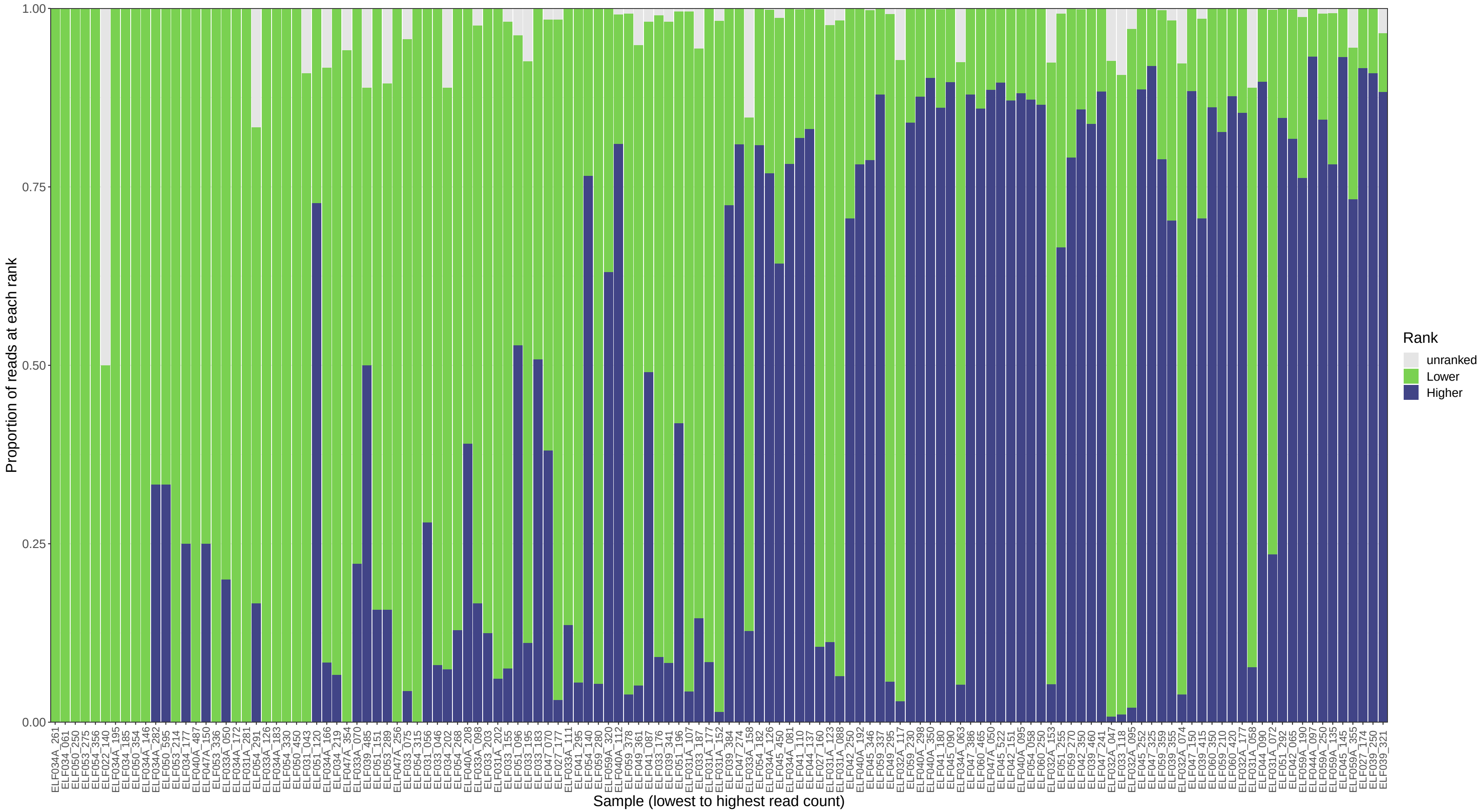


Figure 4.141: Proportions of Embryophyta reads by higher or lower rank for all samples. Lower ranks are family and below. Samples are ordered from smallest to largest read count.

Further evidence that the pattern is specific to *Zostera* comes from a repeat of the analysis using biogenomic mass instead of raw reads (figure 4.142). The pattern is still present, but weaker. See that the highest-yield samples only comprise ∼50% higher taxa instead of ∼90%. The Pearson's test correlating high-to-low rank ratios also returned a smaller *r* of 0.592 (95% CI [0.469, 0.692]; $p<0.01$). These results make sense if we assume the pattern is driven by *Zostera*-dominated samples: the estimated *Zostera* genome size is actually very small, at 0.55 pg DNA per haploid genome (appendix D; Leitch *et al.* 2019). Compare this with 0.33 for the famously-simple *Arabidopsis* and 11.25 for the heavyweight *Triticum*. This small size means that the proportions of *Zostera* and Zosteraceae (same estimate as *Zostera*) increase when reads are converted to biogenomic mass. All of the higher parent taxa, from Alismatales to Embryophyta, have much greater genome sizes (4.95-9.53, though variable), so their proportions *decrease* between reads and biogenomic mass. This explains the difference between the two analyses. Importantly, the few high-yield samples rich in *Salix* and grasses instead of *Zostera* gave much more similar results each time.

The specificity of the effect to *Zostera*-dominated samples suggests that many of the reads assigned to higher taxa have some connection to *Zostera* itself. The higher taxa most affected are parent taxa of *Zostera*; decreasing proximity to *Zostera* could explain why the very highest taxa, Tracheophyta and Embryophyta, do not follow the pattern. The simplest connection would be that the reads derive from *Zostera* but were difficult to assign. PIA is forced to assign to higher taxa if sequences are either conserved or otherwise present in many organisms (such as transposable elements), or not represented in the database. Either could result in diverse BLAST hits.

Is the *Zostera* genome rich in sequences shared by other taxa? Approximately 63% is composed of repeat elements, many of which are widely shared (Olsen *et al.* 2016). However, 63% is not unusually high for plants, and *Zostera* has also lost many conserved genes in its adaptation to a saltwater habitat. For example, the loss of stomata resulted in the loss of the many genes associated with their development and also the volatile chemicals and pathogens that would travel through them (Olsen *et al.* 2016). Recall that the estimated genome size of *Zostera* is really rather small (0.55), as are those of the other taxa in high-yield samples, the PACMAD clade (2.02) and *Salix* (0.58). Reads derived from any of these organisms, including *Zostera*, should be relatively straightforward to assign. The *Zostera* genome has not offered an explanation for the *Zostera*-associated pattern.

Instead, perhaps the issue is not that sequences are conservative, but unrepresented. Perhaps the many reads assigned to higher taxa do not derive from *Zostera*, but from taxa in the same environments that are not well represented in the database. *Zostera*-dominated samples, simple at first glance, may also be rich in taxa that present analyses cannot identify. However, this does not necessarily explain the increasing proportion of such taxa with data yield. For the purposes of this thesis, the high frequency of *Zostera* is enough to justify continuing interpreting these samples as salt/brackish environments, but the source of reads assigned to those higher taxa remains a mystery.

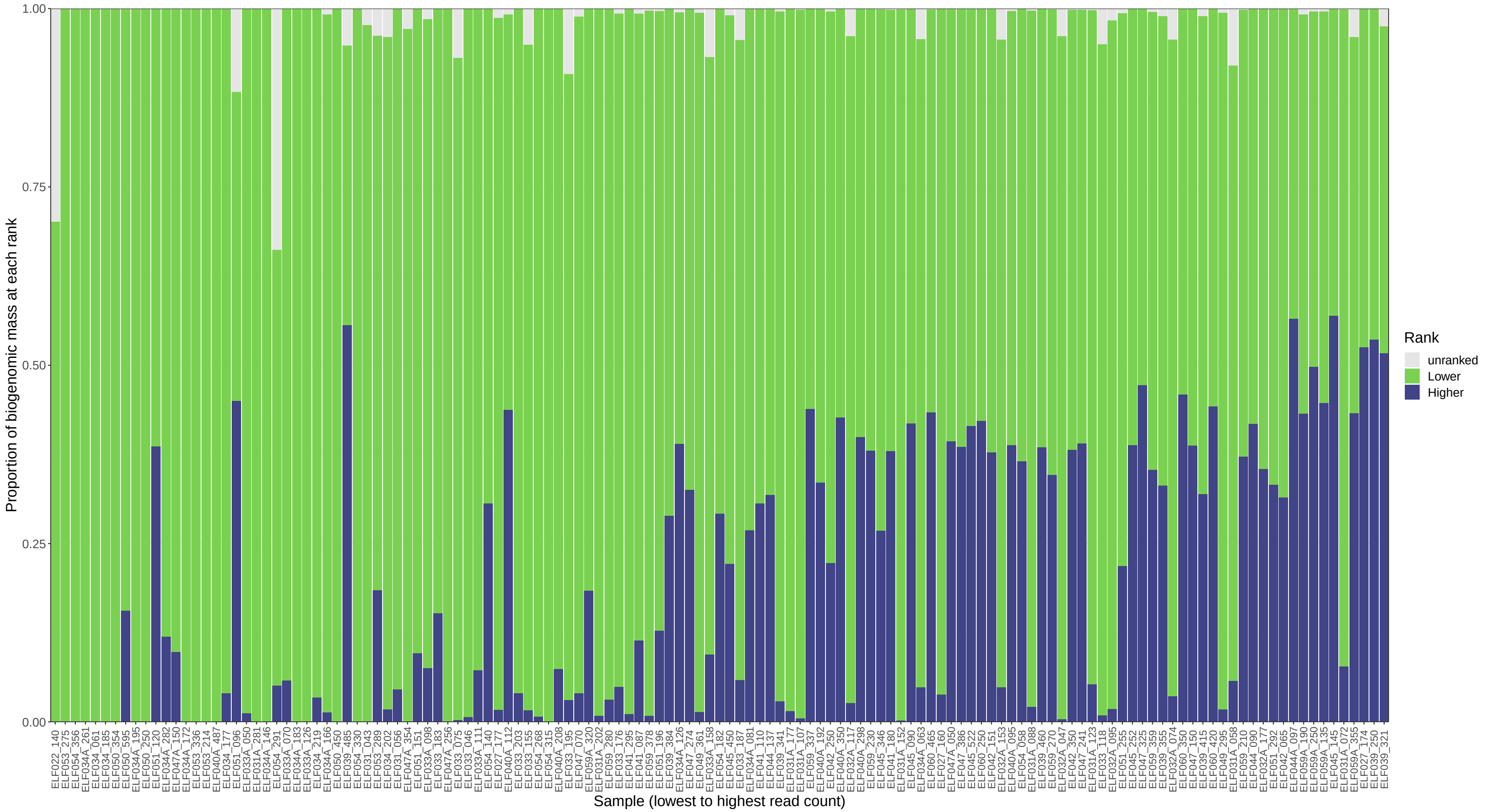


Figure 4.142: Proportions of Embryophyta biogenomic mass by higher or lower rank for all samples. Lower ranks are family and below. Samples are ordered from smallest to largest read count.

4.30.3 Cross-core Pianka analysis

An important question is whether sedaDNA can match similar strata across cores. How can we best summarise taxonomic profiles to determine similarity? The simplest approach would be to label samples by dominant ecological category. A more sensitive comparison was attempted with Pianka scores. Scores for Embryophyta were calculated between all samples with Embryophyta reads using the same methods as within-core Pianka analyses. A second scoring method, the Sørensen–Dice coefficient, was also investigated but appeared more sensitive to data yield than taxonomy. Pianka scores were then used to cluster samples, producing figure 4.143.

Figure 4.143 is symmetrical because Pianka scores are pairwise comparisons, so every score is shown twice; this results in a line of maximum scores through the centre of the heatmap where each sample is compared to itself. The generally bright colours show that most samples have at least a moderate overlap (which may be due to inflation from higher taxa), but there are clearly islands of even greater similarity. The main islands in the heatmap correspond to clusters in the tree. The tree annotation rows show three variables predicted to influence Pianka scores: dominant ecological category, core, and Embryophyta read count. Read counts represent data yield and are log-transformed to reduce the impact of very high outliers on the colour scale.

The variable that most closely follows the cluster pattern is ecological category. Cluster 5 is characterised by salt/brackish taxa and cluster 2 by grasses and relatives. The two samples in cluster 1 are an exception: their Pianka scores for each other are very low and, while one is predominantly bryophytes, the other is ferns. However, as both categories are very rare across the data, cluster 1 makes sense as an outgroup of two samples distinct from everything else. Less straightforward are clusters 3 and 4, which both contain mostly trees/shrubs-type samples. This cross-cluster similarity results in a lack of distinction in figure 4.143: unlike clusters 2 and 5, 3 and 4 do not form clear islands of similarity. Nevertheless, they are distinct from the other clusters, if not from each other.

Clustering by ecology is not surprising because most biogenomic mass is assigned to a small number of key taxa. Over 99% of salt/brackish reads, for example, are from *Zostera*. Pianka scores are based on taxonomy, but where taxonomy is essentially equivalent to ecological category, a close relationship is to be expected. The two mixed-aquatic-type samples in cluster 2 (mostly grasses and relatives) are a rare case of disagreement: much of their biogenomic mass is from reeds, which are mixed aquatics but taxonomically grasses. Pianka scores are not affected by the more artificial divisions of ecological category. Overall, however, scores and categories tell much the same story.

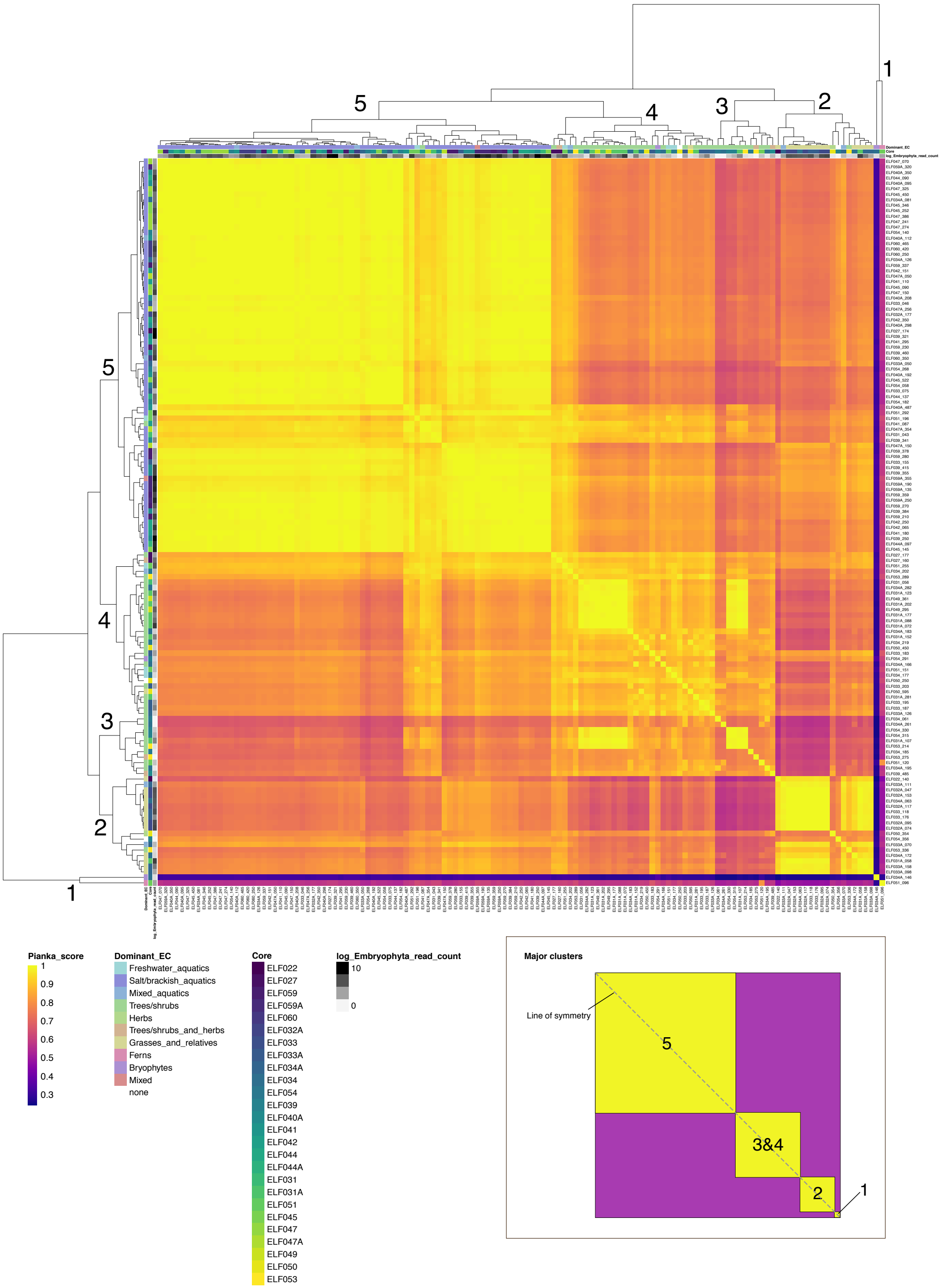


Figure 4.143: Pairwise Pianka similarity scores for Embryophyta between all samples with Embryophyta reads. Scores were calculated from biogenomic masses, counted cumulatively, as a proportion of the sample total. 0 indicates no overlap and 1 complete overlap, but note that cumulative counting can lead to inflated scores. The three annotations are "Dominant_EC" (the most frequent ecological category in that sample), core (in order of transect position), and Embryophyta read count (log-transformed to better distinguish colours). Produced using pheatmap v1.0.12 (Kolde 2018) with default cluster settings (complete linkage) and viridis colour palettes (Garnier *et al.* 2018). Inset explains the main clusters.

Core and data yield are less helpful for inferring ecology and, fortunately, appear less influential on Pianka score. Strings of samples from the same or neighbouring cores are mostly small and scattered. The only noticeable group contains nine of the eighteen samples in cluster 2. Their very high Pianka scores are reflected in the very short branch lengths in the tree. Data yield may have slightly more influence. The Pianka comparisons within cores were clearly affected by read count, with low-yield samples often appearing less similar, and yield is also partially correlated with dominant ecological category (subsection 4.30.1). Yield may be a separating factor for clusters 3 and 4, which were separated by Pianka score yet share ecological category and taxonomic similarity. The boxplot in figure 4.144 compares Embryophyta yield between clusters (note the log axis). Cluster 4 yields show more variation than 3, but are generally higher. High-yield samples are expected to contain more biogenomic mass from certain higher taxa, making high-yield samples appear more taxonomically similar to each other. However, there is clearly yield overlap between the clusters, so this does not fully explain their division.

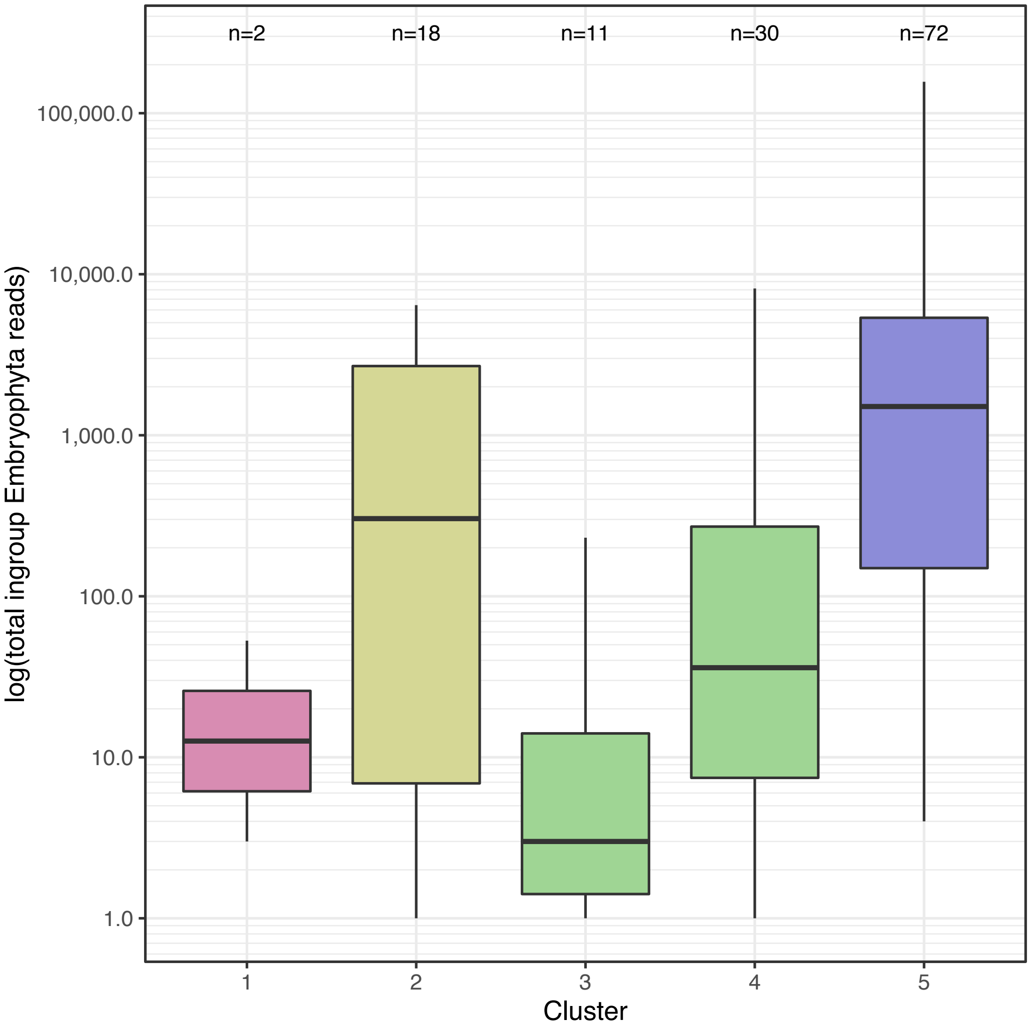


Figure 4.144: Boxplot of data yield by cluster, as defined in figure 4.143. Boxes show the IQR and lines 1.5*IQR. Yield is the total ingroup Embryophyta read counts per sample. Limited to samples with at least one ingroup Embryophyta read. Note that the y-axis is in log base 10.

In fact, further investigation casts doubt on the separation of clusters 3 and 4. Clustering was repeated with various subsets of clusters to check stability. For example, figure 4.145 was produced by omitting samples from clusters 4 and 5. These iterations often resulted in the movement of clusters and of samples within clusters (particularly when the overall sample number was small), but cluster membership was generally stable. There were only two exceptions. First, cluster 1 split when the analysis included clusters 1, 2, and 3 only (figure 4.145). These were the three smallest clusters, so the change may be explained by limited data, and the two samples in cluster 1 already seem to be grouped less by their shared features than their substantial differences to all other samples.

More interestingly, the boundaries of clusters 3 and 4 broke down in all subsets of samples. The fewer samples in the analysis, the more exchange of samples between clusters 3 and 4. For example, when only cluster 2 was omitted, the only movement was of sample ELF031A_202 from cluster 3 to 4. However, when cluster 5 (the largest) was omitted (figure 4.146), three samples moved from cluster 4 to 3 and five samples from 3 to 4. Cluster 2 forms a clear island of similarity, but high Pianka scores are shared by only a fraction of those in clusters 3 and 4.

It seems that the division of clusters 3 and 4 is driven more by relative similarity to other clusters, especially 5, than by differences within 3 and 4. Like cluster 1, this instability may be explained by a lack of data; figure 4.144 showed that samples in clusters 3 and 4 typically have far fewer reads than in clusters 2 and 5. Despite fewer similarities drawing those samples together, they find themselves grouped because they are excluded from better-defined clusters. The strongest divisions found by the Pianka analysis match dominant ecological category.

This is not to say that the Pianka clusters do not add any more information than ecological category, but the additional structure is not as well supported and depends on the influence of samples outside the clusters in question. I would be wary of over-interpreting the division between clusters 3 and 4, especially considering the generally low data yields in those samples. More broadly, the instability of clusters 3 and 4 and the different within-cluster topologies in different subset analyses may be a sign that the tree is difficult to resolve. Pianka scores were not designed for such incomplete data containing multiple taxonomic levels and are clearly limited when comparing these taxonomic profiles. Most samples share relatively high Pianka scores (the heatmaps are generally warm in colour), not necessarily because of similar ecologies, and clustering becomes less reliable when differences in scores are small. For further analyses, I will take dominant ecological category as the main source of taxonomic information instead of Pianka cluster.

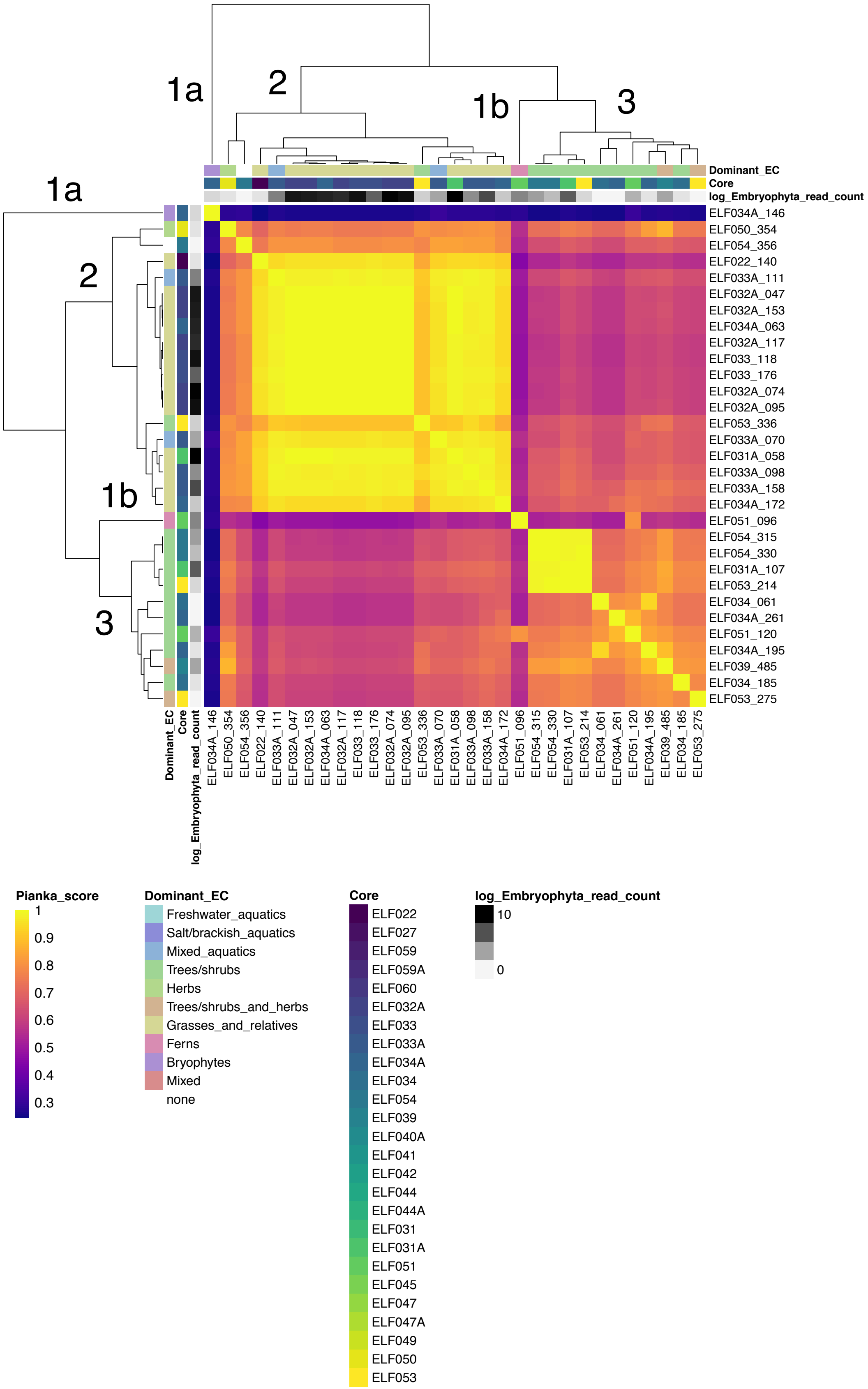


Figure 4.145: Pairwise Pianka similarity scores for Embryophyta between samples from clusters 1, 2, and 3. The separated samples from cluster 1 are labelled 1a and 1b; 1a is in the original position, but 1b is now the sister group to cluster 3. Produced using the same method as figure 4.143.

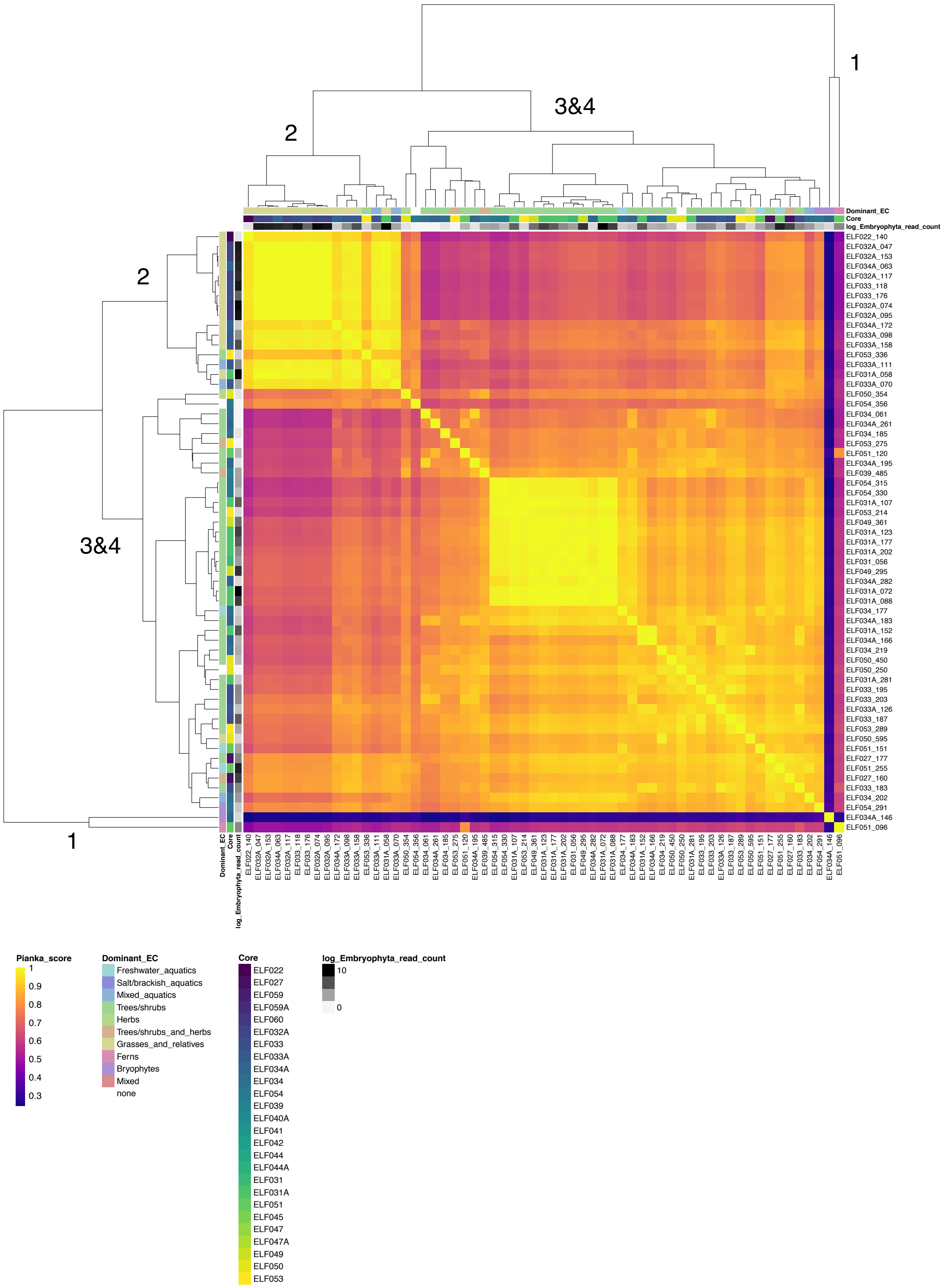


Figure 4.146: Pairwise Pianka similarity scores for Embryophyta between samples from clusters 1, 2, 3, and 4. Note that sample ELF031A_202 has moved from cluster 3 to 4. Produced using the same method as figure 4.143.

4.30.4 Spatial overview

Figures 4.147 and 4.148 plot summarised results onto core locations and starting depths respectively. Figure 4.147 shows few clear geographical patterns, although paired cores from the same locations often show similar profiles. Figure 4.148 may show some trends in depth, such saltwater samples generally appearing above non-saltwater in cores with environmental changes, as in ELF033, ELF033A, ELF054, ELF031. However, cores like ELF027 and ELF032A reverse this trend, showing the importance of interpreting each core individually.

There is little association between dominant ecological category and absolute depth; this was further investigated in a similar plot with samples at their actual depths, instead of represented by blocks, which is not readable at this scale. This is not surprising: while we may expect the same strata to share DNA profiles across cores, equal depth does not necessarily equal the same stratum. Many factors affect the pattern of sedimentation, and after sampling, and the depths of strata in a core may change because of changes in pressure and moisture. Ecological category across location and depth may have shown general trends, such as deeper strata appearing more terrestrial, but to attempt to match strata between cores, other proxies will be necessary, particularly dates. These will be examined in future work.

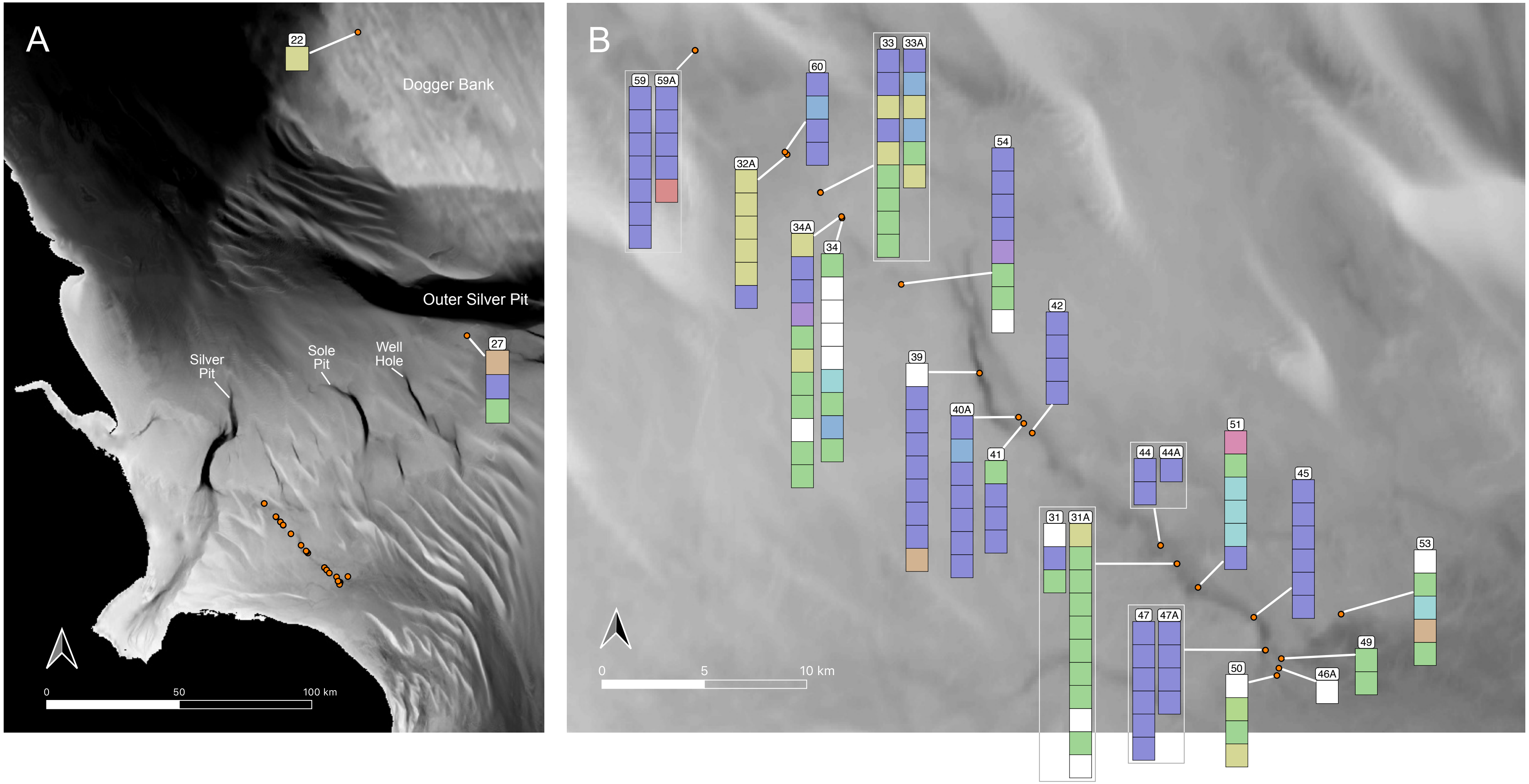


Figure 4.147: Figures 2.2 and 2.3 with the addition of summarised sedaDNA results. Samples within a core are not to scale, but are represented by blocks. Depth between samples is omitted. Blocks are coloured according to most frequent Embryophyta ecological category; see figure 4.4 for colour key. Cores from the same locations are boxed.

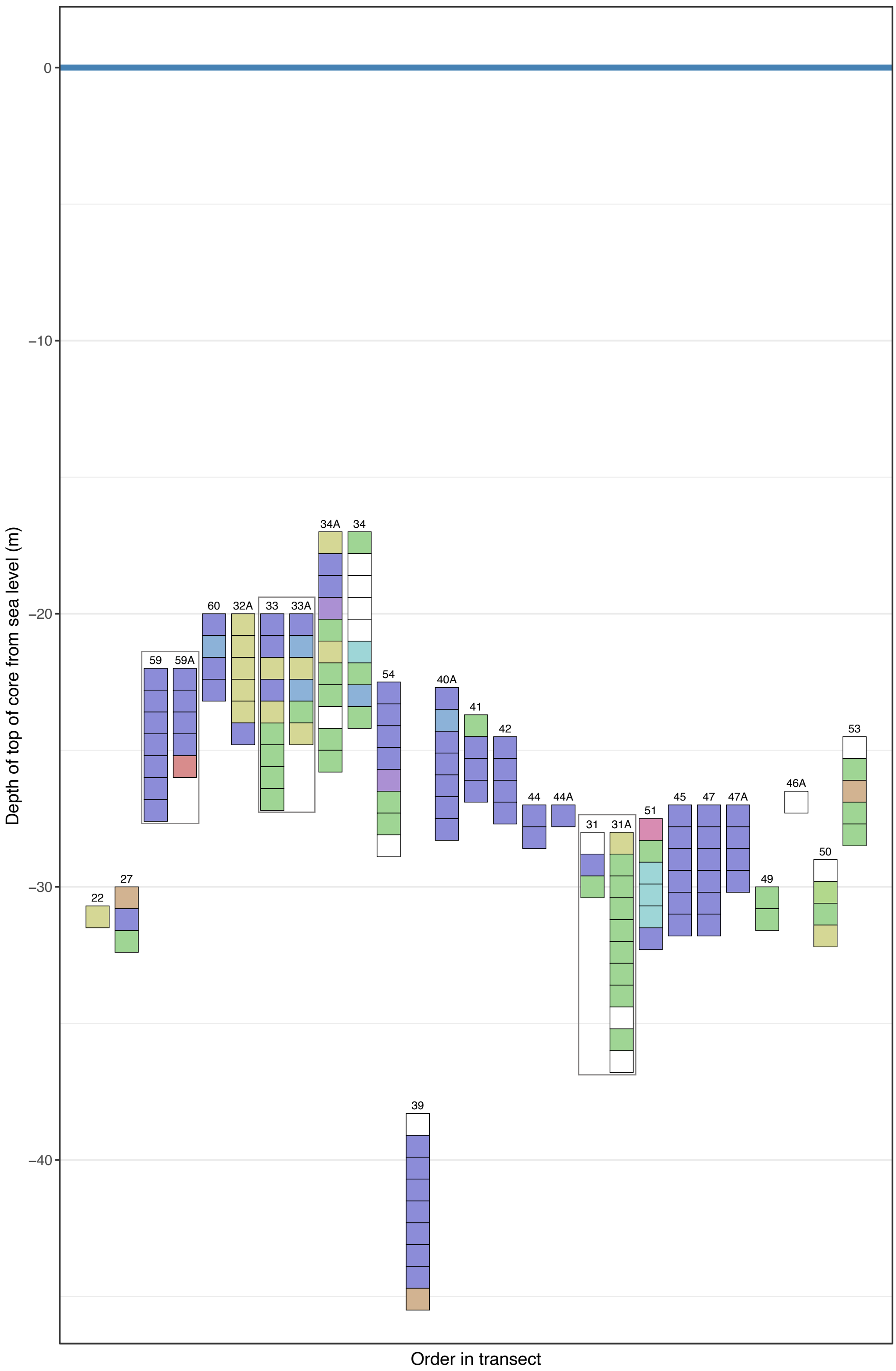


Figure 4.148: Cores by starting depth (m) below sea level (blue line at depth=0) with summarised sedaDNA results. Samples within a core are not to scale, but are represented by blocks, with the top block representing the top of the core. Depth between samples is omitted. Blocks are coloured according to most frequent Embryophyta ecological category for that sample; see figure 4.4 for colour key. Cores from the same locations are boxed.

4.31 Conclusion

The previous chapter outlined the main tool used for taxonomic assignment in this project, PIA, and justifications for its use. The beginning of this chapter revisited the most suitable existing alternative, MEGAN, for a broader benchmarking analysis. This confirmed that, for reads loosely assigned to Viridiplantae, MEGAN assigned with greater sensitivity but less accuracy. Although it was not feasible to measure accuracy for this much larger dataset, close examination of an outlier sample where MEGAN assigned many more reads than PIA revealed only that only 35% of MEGAN's assignments were correct, compared to 87% in the test data from Chapter 3. The incorrect assignments were mostly to over-represented taxa, suggesting an unusually high proportion of reads matching difficult areas of the database. PIA is designed to avoid assigning these reads, but MEGAN went right on ahead. This suggests that PIA is performing as expected, so we can assume the expected high accuracy for the PIA assignments upon which this chapter is built.

Table 4.57: Summary of taxonomic results by core.

Core	Embryophyta data yield	Most frequent ecological category	Environmental change?
ELF022	Low	Mixed	N/A (single sample)
ELF027	Variable	Salt/brackish aquatics	Trees/shrubs and herbs to salt/brackish to trees/shrubs
ELF059	High	Salt/brackish aquatics	N
ELF059A	High	Salt/brackish aquatics	N
ELF060	High	Salt/brackish aquatics	N
ELF032A	High	Grasses and relatives	Salt/brackish to grasses
ELF033	Variable	Grasses and relatives	Trees/shrubs to grasses to salt/brackish
ELF033A	Low	Salt/brackish aquatics	Trees/shrubs to salt/brackish
ELF034A	Variable	Grasses and relatives	Trees/shrubs to salt/brackish
ELF034	Low	Trees/shrubs	N
ELF054	Variable	Salt/brackish aquatics	Trees/shrubs to salt/brackish
ELF039	High	Salt/brackish aquatics	N
ELF040A	Variable	Salt/brackish aquatics	N
ELF041	Variable	Salt/brackish aquatics	N
ELF042	High	Salt/brackish aquatics	N
ELF044	High	Salt/brackish aquatics	N
ELF044A	High	Salt/brackish aquatics	N/A (single sample)
ELF031	Low	Trees/shrubs	N
ELF031A	Variable	Trees/shrubs	Trees/shrubs to grasses
ELF051	Variable	Salt/brackish aquatics	Salt/brackish to freshwater to terrestrial
ELF045	High	Salt/brackish aquatics	N
ELF047	Variable	Salt/brackish aquatics	N
ELF047A	Variable	Salt/brackish aquatics	N
ELF049	Variable	Trees/shrubs	N
ELF046A	N/A	N/A	N/A (single sample)
ELF050	Low	Trees/shrubs	N
ELF053	Low	Trees/shrubs	N

Table 4.57 summarises key sedaDNA results from the twenty-seven sediment cores analysed in this project. The dominant ecological category for most cores and samples was salt/brackish aquatics, overwhelmingly consisting of the seagrass *Zostera*, demonstrating a very clear marine signal. However, all of these samples also contained at least a small terrestrial signal, predominantly willow or poplar trees, and higher grass taxa which could often be interpreted as reeds. Together with the widespread but low-frequency *Ruppia*, a brackish seagrass, these placed most cores in a coastal setting.

Most taxa were consistent with a similar climate and biogeography to England today. However, certain taxa suggested otherwise. Highlighted in bold throughout this chapter were taxa native to Europe but not Great Britain, and in almost all cases their range extends through southern Europe only. Their presence in the Southern River therefore suggests a warmer climate. The most frequent of these mesophilic taxa are grasses in Panicoideae, present in 53% of samples, followed by Juglandaceae (the walnut family) and *Posidonia* (Neptune grasses). Overlapping with mesophiles are a second category of potentially human-associated plants, such as *Hordeum* (barleys) and *Triticum* (wheats). Juglandaceae could also fall into this group, having been transported by humans for thousands of years (de Rigo *et al.* 2016). These two groups, mesophiles and potentially human-associated taxa, are subject to age-authentication in Chapter 6 (Mesophilic taxa and human disturbance indicators) where data allows. Missing from Chapter 6 are a third group of surprising taxa, the cold-associated taxa noted in ELF049. While many of those taxa appeared in other cores, their concentration in ELF049 suggested a pattern. None had enough data for individual age-authentication, but the combined Embryophyta data for each sample passed (Chapter 5: Authentication results), so the cold signal therefore also appears genuine. Interpretation of the remaining unexpected taxa will be considered again in Chapter 6.

While data yield for Embryophyta was hugely variable, the number of Metazoa reads per sample was consistently low. This significantly limited interpretation. The difference may be partly due to the lower biomass expected of consumers than producers, but possibly also methodological differences, such as the impact of the predominantly human and animal laboratory contamination (Leonard *et al.* 2007). Potentially due to this contamination, there were also several human-associated Metazoa taxa, but none had sufficient data for age-authentication. The same was true for the potential mesophile, *Emys* (pond terrapins), so authentication of exotics in Chapter 6 deals exclusively with Embryophyta.

There is a clear relationship between ecological category and data yield, with salt/brackish and grasses and relatives (which includes brackish reeds) most clearly associated with high read counts. Greater DNA preservation is expected in high-salt environments (Lindahl and Nyberg 1972, Kistler *et al.* 2017). Less easily explained is the association in *Zostera*-dominated samples between high read counts and proportionally more assignments to higher taxa. This may be an unexpected result of how PIA assigns sequences. Further work on PIA, especially considering results from other projects now that it has been published, may improve our understanding of how it behaves with different taxa.

Cross-core Pianka analyses were evaluated as a possible method to identify ecologically similar samples. The results are generally very similar to simply grouping by dominant ecological category. The main exception, the division of most trees/shrubs-type samples into two groups, appeared less reliable than clusters in broad agreement with ecological category. Correlations between taxonomic profile and location or core starting depths are also weak, although sister cores generally appear similar.

Instead, perhaps the most interesting result from the taxonomic profiles is that eight cores appear to show environmental change. There is a potential wandering channel in ELF027, succession of reed beds or salt marsh in ELF032A, and either reversed inundation or an inverted core in ELF051. The remaining five show transition from terrestrial to either reeds or salt/brackish profiles. These may have successfully captured the inundation of Doggerland.

The next chapter considers the validity of these taxonomic interpretations by attempting to authenticate the data. Filtering out common contaminant and biogeographically implausible taxa is important, but there are two further tests for remaining data: age-related DNA damage and signs of movement between sediment strata.

Chapter 5

Authentication results

The previous chapter considered the assignment of DNA sequences to taxa. It assumed that all sequences were derived from endogenous DNA: DNA that had been deposited more or less simultaneously with the sediment and that was therefore of the same context as any other sedimentary environmental proxies, including dates. This chapter uses two methods to assess that assumption. The first, MetaDamage, looks for signs of age-related damage to exclude modern contamination. The second, stratification analysis, compares DNA profiles between adjacent samples to exclude DNA movement through the sediment, which would dissociate ancient sequences from their context. See Chapter 2 (Main materials and methods) for detailed methods.

5.1 MetaDamage analysis

The MetaDamage analysis should be more successful with a greater number of input reads. Therefore, the analysis was first performed for combined Embryophyta and Metazoa (no Primates) data from all cores, then for individual samples, and finally for subsets of interesting taxa.

5.1.1 Embryophyta across all samples

A key damage signature of ancient sequences is cytosine deamination, which causes C bases to appear as T on the forward DNA strand and G to appear as A on the reverse. Age-related damage such as this is more likely on the ends of molecules because this is where single-stranded overhangs occur, in which bases are more exposed (Kistler *et al.* 2017). Therefore, we would expect an increase in C-to-T mismatches at the 5' end of ancient sequences and G-to-A mismatches at the 3'. This is what the MetaDamage analysis searches for.

Figure 5.1 shows the MetaDamage plot from combined Embryophyta reads across all samples. Similar to the misincorporation plots produced by mapDamage2.0 (Jónsson *et al.* 2013), the y-axis measures the probability of a mismatch occurring (labelled "$\mathrm{P}_{\mathrm{substitution}}$" on the plots). Of positions containing a given base in the reference sequences, this is the proportion that had a different base in the sample sequences.

The twelve lines plotted each represent a different type of mismatch. C-to-T and G-to-A are highlighted in red and blue respectively as these are the most informative. Other mismatches are in grey and give an impression of the background rate, which we assume is more related to phylogenetic differences between the read and its reference than damage. The C-to-T and G-to-A lines also have error bars showing a 95% confidence interval for the mismatch rate at position zero. The position-zero value is a point estimate and the true value is likely to be within the interval, so the wider the interval, the more varied the data for this sample of reads, and the less confidence we can have in the position-zero estimate. Greater variance is expected from fewer reads.

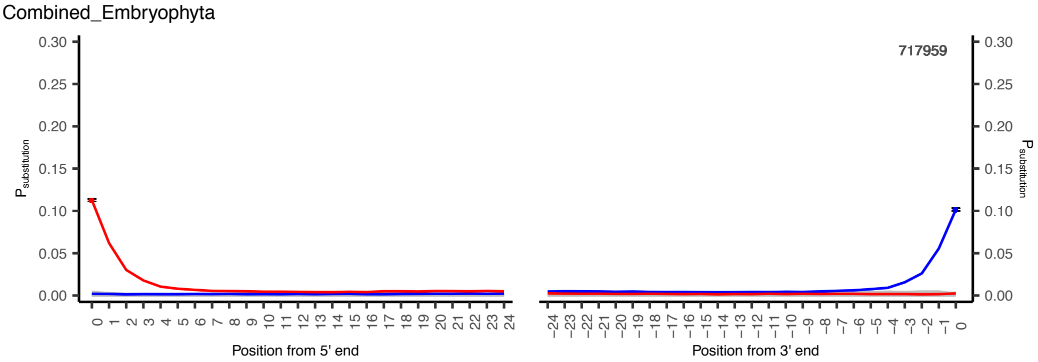


Figure 5.1: MetaDamage profile for Embryophyta reads across all samples. $P_{\text{substitution}}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right.

Figure 5.1 has a near-perfect ancient signal. The probabilities of most types of mismatch are nearly zero on both plots, with the exceptions of C-to-T at the 5' end and G-to-A at the 3', which show very clear upticks. Their values at position zero on the relevant plots are 0.113 and 0.102 respectively, far higher than other types of mismatch, and they become increasingly frequent in the end ten bases or so. Furthermore, both position-zero values have very narrow 95% confidence intervals, suggesting a highly consistent signal.

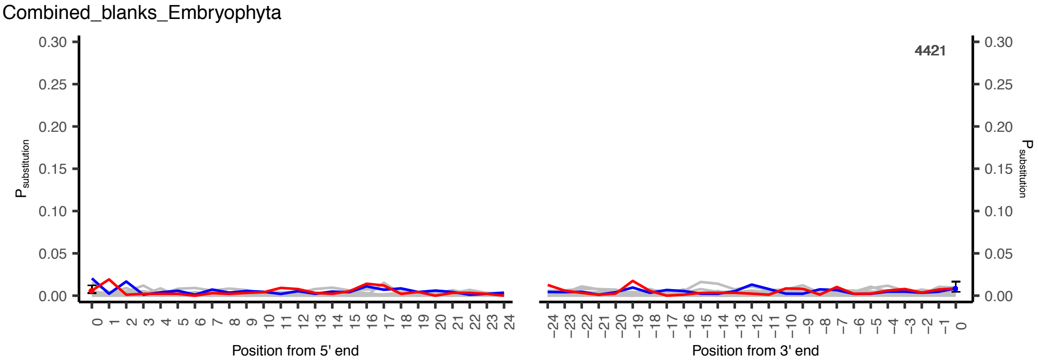


Figure 5.2: MetaDamage profile for Embryophyta negative controls. $P_{\text{substitution}}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right.

Contrast this with figure 5.2: the negative controls. There is generally more noise in all of the mismatch lines, which is to be expected from the less exceptionally high read count. The smaller amount of data also contributes to the wider confidence intervals, although they are still sufficiently narrow to show a consistent signal.

More importantly, even allowing for the variation, C-to-T or G-to-A upticks are absent. Their position-zero mismatch values are only 0.006 and 0.009; these are very low and also similar to those for other types of mismatch. Therefore, these sequences appear modern, as would be expected in a negative control. This result suggests a lack of cross-contamination through ancient sequences moving between samples.

A final analysis considers a subset of the Embryophyta data for which a damage signal may be unexpected. These reads were assigned to non-European taxa. They are not expected to represent genuine plant taxa from Holocene Doggerland, but instead may represent correctly-assigned modern contaminants or mis-assigned endogenous reads from other taxa. Figure 5.3 shows a clear damage signal, but also a high overall mismatch rate, especially for the large quantity of data (nearly 2,000 reads). This suggests limited similarity to the top BLAST hits for these sequences, supporting the explanation that they are mis-assigned ancient sequences.

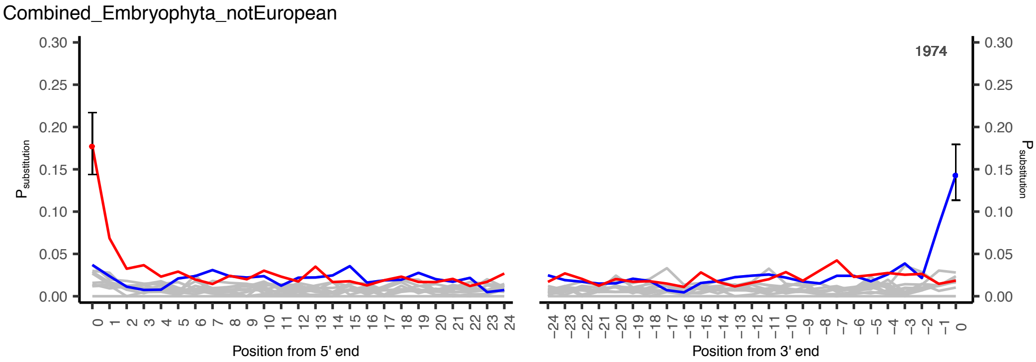


Figure 5.3: MetaDamage profile for non-European Embryophyta reads across the 101 samples in which they occurred. $P_{substitution}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right.

The ten most frequent non-European taxa are shown in table 5.1. All are domesticates or close parent taxa thereof, so have close links to heavily-studied taxa that are likely to be over-represented on GenBank. The reads must match their reference sequences relatively closely for a damage signal to be visible, and the MetaDamage BLAST only considers the first reasonable match found. This suggests that these sequences are either conservative or genuine close relatives of those taxa.

The identity of the original organisms was investigated by correlating the read counts of taxa in table 5.1 with those of all other taxa. If an unlikely taxon frequently co-occurs with a more likely relative, this could suggest the relative as a source of mis-assigned reads. The co-occurrence analysis was performed with the initial sequencing data from cores ELF001-60 (including from cores ELF001-20, analysed by Roselyn Ware) and involved all Viridiplantae taxa, European or not. Correlations between taxa

Table 5.1: The ten most frequent non-European Embryophyta taxa.

Common name	Taxon	Total
Bamboos	Bambusoideae	668
Rice genus	Oryza	380
Cultivated rice	Oryza sativa	101
Peanut genus	Arachis	69
Sugarcane	Saccharum hybrid cultivar	66
Maize subtribe	Tripsacinae	47
Sugarcane	Saccharum officinarum complex	40
Sugarcane subtribe	Saccharinae	35
Cacao subfamily	Byttnerioideae	25
Gourds	Cucurbita	25

were calculated as Pearson's *r*. The bamboo, rice, maize, and sugarcane taxa all show strong positive correlations with each other, other taxa in Panicoideae besides maize and sugarcane, and also *Phragmites* (common reed). *Phragmites* is a significant component of wetland systems in Northern Europe today (Packer *et al.* 2017), and featured in the post-inundation vegetation of the Doggerland core analysed by Wolters *et al.* (2010). It was found in many samples, but at relatively low frequency. This could be explained by its limited representation in the GenBank database used for this thesis (988 records and no full genome). *Phragmites* and Panicoideae (maize and sugarcane) are both in the PACMAD clade of grasses, which is sister to the BOP clade (bamboo and rice). Perhaps a portion of *Phragmites* reads matched reference sequences from these over-represented relatives well enough to be incorrectly assigned, but not for PIA to recognise the lack of diversity in the BLAST hits and discard the reads. Incidentally, this raises the possibility that *Phragmites* is also the source of the European Panicoideae reads, but these taxa will be investigated more thoroughly in Chapter 6 (Mesophilic taxa and human disturbance indicators).

Similar correlations may explain the remaining non-European taxa, although with lower read counts the analysis loses power. Byttnerioideae correlates highly with *Puccinellia* (saltmarsh grass), *Zostera*, and Alismatales, but also *Tilioideae*. The salt-water indicators are not taxonomically close, so this may instead be due to the greater data yields of samples high in salt/brackish aquatics (Chapter 4: Taxonomic results); samples rich in these taxa were more likely to contain rare taxa such as Byttnerioideae. However, *Tilioideae* is in the same family as Byttnerioideae so may be a more likely source. Likely European candidates are less forthcoming for *Arachis*, which correlates weakly with various taxa of polypod ferns, and *Cucurbita*, whose strong correlates include algae, the gymnosperms Cuppressaceae, and the over-represented *Brassica*. However, the co-occurrence analysis has suggested explanations for the more frequent non-European taxa and may be useful for future investigations into why PIA occasionally mis-assigns.

5.1.2 Metazoa across all samples

The Metazoa (no Primates) negative controls also present a reassuring MetaDamage profile (figure 5.4). This was produced using nearly 3,000 reads, yet neither plot shows an uptick in relevant mismatches.

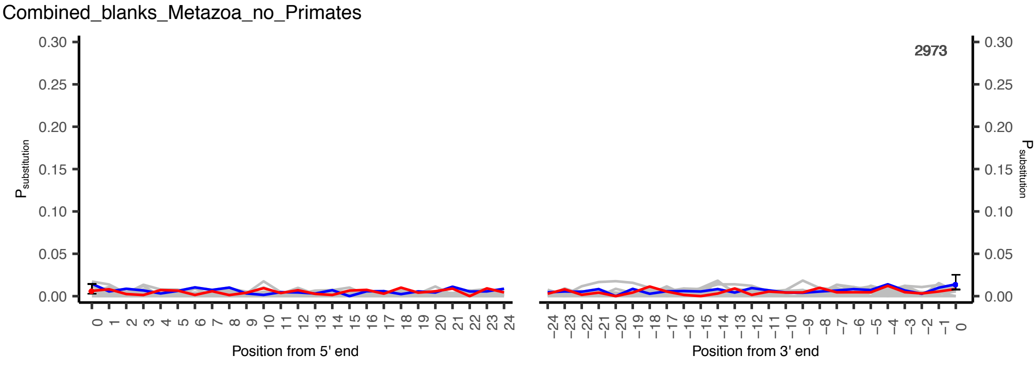


Figure 5.4: MetaDamage profile for Metazoa (no Primates) negative controls. $P_{substitution}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right.

However, the sample data for Metazoa is less promising. In figure 5.5, we see that the combined samples do show C-to-T and G-to-A upticks: the position-zero values are 0.069 and 0.075, which while below those of Embryophyta, are still relatively high compared to further along the molecules. The confidence intervals are rather wide, however, and any potential signal is overpowered by a high background level of mismatch, including possible upticks for several other mismatch types. This variation was surprising.

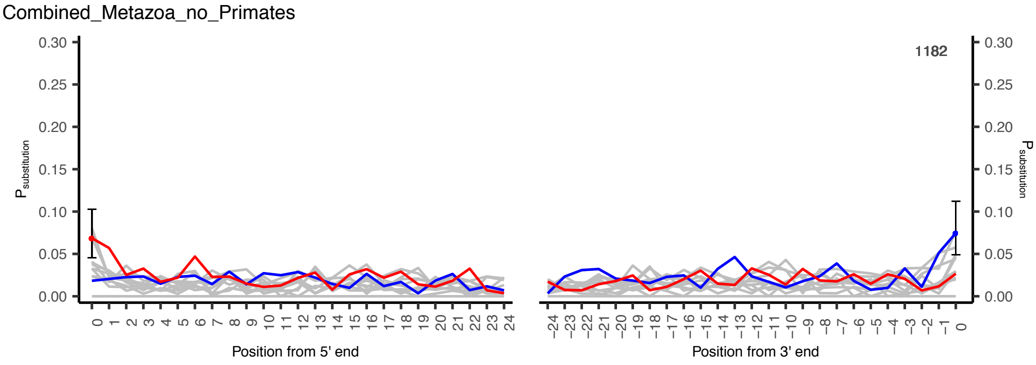


Figure 5.5: MetaDamage profile for Metazoa (no Primates) reads across all samples. $P_{substitution}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right.

The analysis involved fewer reads than previous plots shown, so more variation may be expected. However, many MetaDamage plots produced for the Embryophyta data of individual samples showed clear positive results with far fewer reads (see MetaDamage plots per sample in appendix F). This cannot be explained by lack of data alone.

Instead, a generally higher mismatch rate for Metazoa could suggest greater phylogenetic distance between these sample sequences and their references. Unlike PIA, MetaDamage only considers the first BLAST hit. A reference that matches perfectly is no use because it would not reveal age-associated mismatches. However, this means that MetaDamage cannot discount comparatively poor hits to over-represented taxa, unlike PIA. The high background mismatch rate seen for many Metazoa reads could be due to greater influence from over-represented taxa than for Embryophyta; at least in 2010, the top represented organisms in GenBank were mostly animals (Benson *et al.* 2011).

The high background rate may be obscuring an underlying damage signal. In an attempt to reduce it, the Metazoa (no Primates) data from the initial sequencing was combined with preliminary data from the deep sequencing runs, increasing the number of reads to 13,576. The resulting plot (figure 5.6) does show less variation, but still an elevated rate for most types of mismatch. There are C-to-T and G-to-A upticks with narrow 95% confidence intervals, but the point estimates are rather small at 0.040 and 0.045 respectively, compared to 0.113 and 0.102 for the Embryophyta initial sequencing results in figure 5.1. Only the G-to-A uptick has a confidence interval clear of non-age-associated mismatch types. The small upticks do not translate into a clear damage signal.

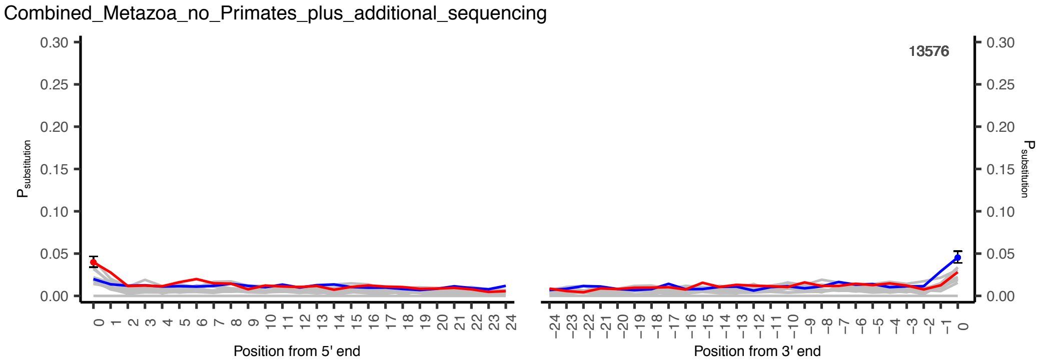


Figure 5.6: MetaDamage profile for Metazoa (no Primates) reads across all samples, including deep sequencing. $P_{\text{substitution}}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right.

The simplest explanation of an unexpectedly small damage signal is modern contamination. Laboratory contamination is predominantly animal, often human, and remains a risk despite best practice precautions (Fulton and Shapiro 2019, Leonard *et al.* 2007). The human content of these samples was roughly estimated by comparing the size of pre-PIA FASTA files (after initial taxonomic assignment by MEGAN) for Primates and for Metazoa without Primates from the first initial sequencing run (152 pairs). The mean ratio of Primates to other-Metazoa FASTAs was 21:1, although this varied greatly (n = 152, standard deviation = 33). A typical sample contained 21 times as much Primates data than other Metazoa. The Primates data is highly likely to be modern human contamination. Figure 5.7 shows the combined MetaDamage output from the pre-PIA Primates FASTAs from initial sequencing of cores ELF021-60: the age-associated mismatch lines are almost horizontal. The C-to-T and G-to-A position-zero mismatch values are only 0.004 95% CI [0.004, 0.005] and 0.005 95% CI [0.005, 0.006] respectively. It is possible that some of these overwhelmingly modern, human-derived reads were assigned to other animals or taxa above Primates that were not subsequently filtered out. Modern contamination from other animal sources, such as *Bos*, *Sus*, and *Gallus*, must also be considered (Leonard *et al.* 2007). The weak damage signal for the combined Metazoa (no Primates) data suggests that Metazoa reads, particularly if assigned to a likely contaminant taxon, should be taken with caution.

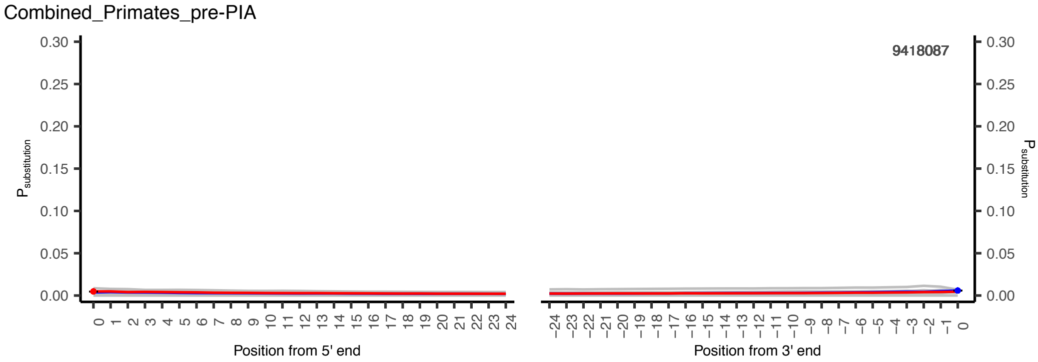


Figure 5.7: MetaDamage profile for pre-PIA Primates reads from initial sequencing of cores ELF021-60. $P_{\text{substitution}}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right.

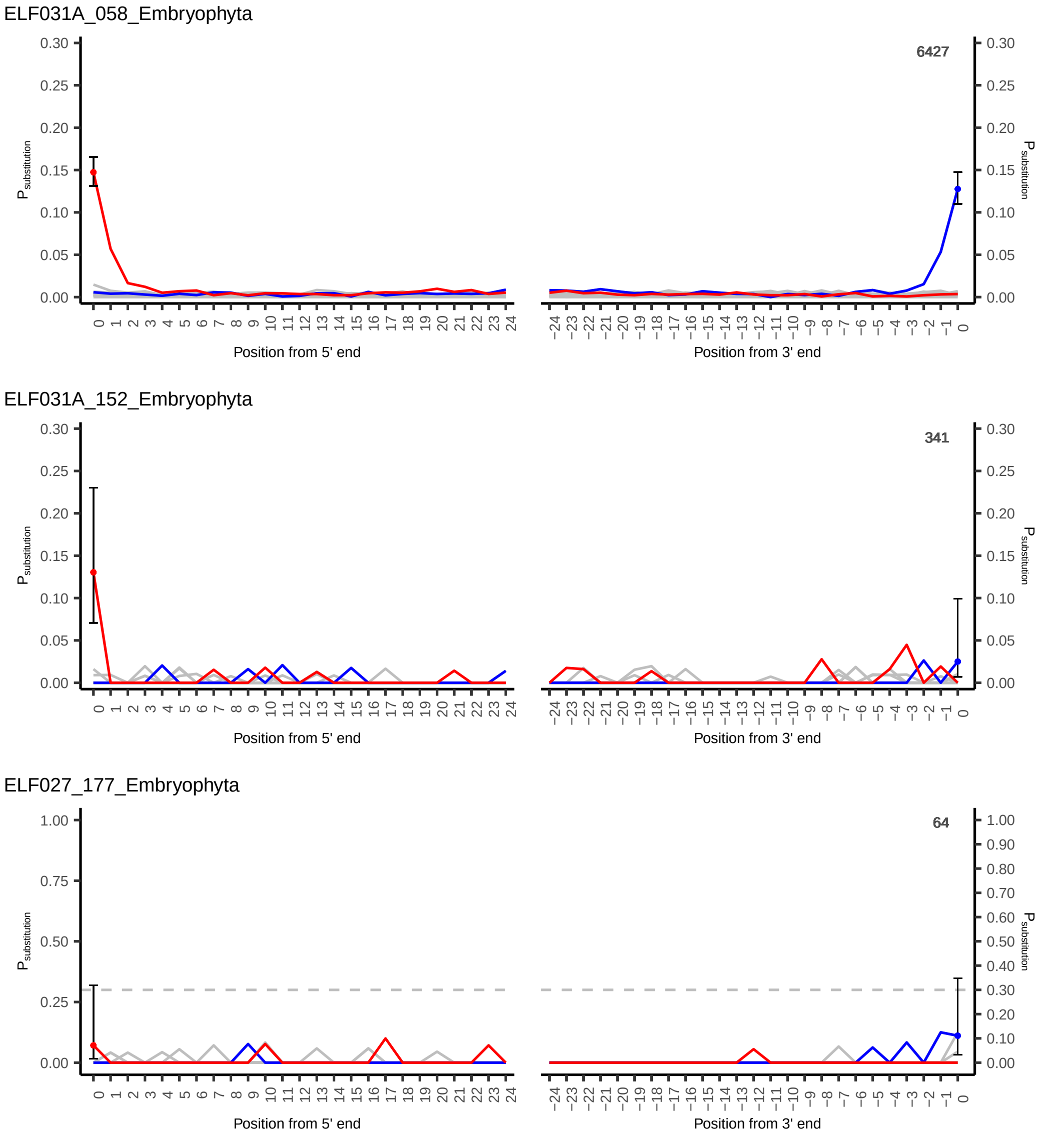
Unsurprisingly, the MetaDamage analyses for individual samples did not improve on this. Only two samples (ELF031A_058 and ELF059A_355) had the recommended minimum of 100-300 reads for a reliable result, and neither showed a strong signal. The MetaDamage results for individual samples have therefore been omitted or moved to appendix G for brevity, with further discussion limited to the Embryophyta data.

5.1.3 Embryophyta by sample

Table 5.2 (page 310) lists read count and MetaDamage output statistics for the Embryophyta data for each sample. The main output of MetaDamage is the plots. All 144 are given in appendix F, but as a summary, table 5.2 contains a manual interpretation of the signal in the plots as "strong", "weak", or "none". For a strong signal, the confidence interval must exclude zero and any mismatch value from other base positions, and this must occur on both the 5' and 3' plots. If this occurs only on one plot, the signal is "weak", and if this is absent from both plots, the signal is "none". The three MetaDamage plots in figure 5.8 demonstrate a strong, weak, and absent signal respectively.

Accompanying this manual classification are the properly quantitative output statistics from MetaDamage: the position-zero mismatch values and corresponding 95% confidence intervals for C-to-T on the 5' end and G-to-A on the 3'. These are coloured by value, with most saturation for the highest values.

Note that a position-zero value for two samples, ELF051_120 and ELF053_336, falls outside of its 95% confidence interval. This odd result can be attributed to the very low numbers of reads involved for these samples. For example, ELF051_120 had only four opportunities for a C-to-T mismatch, so the only possible position-zero values (proportion of successes) are 0, ¼, ½, ¾, and 1. The observed value is 1. The confidence interval, [0.158, 0.987], is derived from the beta distribution and is not limited to this discrete set of possible outcomes. As the number of reads involved increases, so does the number of opportunities for a given mismatch, so the possible outcomes become less limited, and are more likely to fall inside their 95% confidence intervals.



ELF031A_058_Embryophyta

ELF031A_152_Embryophyta

ELF027_177_Embryophyta

Figure 5.8: Example manual interpretations of MetaDamage profiles. ELF031A_058 shows a strong signal because it has a C-to-T uptick on the 5' end and a G-to-A on the 3', and neither confidence interval includes zero or mismatch values from other base positions. ELF031A_152 fulfils this for the 5' end, but the confidence interval of G-to-A mismatches on the 3' end includes values mismatch values from many other base positions; it is not a distinct uptick. A clear damage signal on only one end results in a "weak" classification. Finally, ELF027_177 has no visible signal because neither plot shows a distinct age-associated uptick.

We would expect the 5' and 3' measures to be roughly symmetrical with some additional variation in the 3' from the typically lower quality of reverse sequencing reads. However, a Pearson's correlation test of the position-zero values produced an *r* value of only 0.240. This indicates a weak association only, although it was statistically significant (p=0.005). The ends are therefore analysed separately. The symmetry may be obscured by the significant variation on both ends of the molecules in many analyses with little data (see below).

Table 5.2: Metadamage results by sample for reads assigned to Embryophyta, in depth order from the top of each core. Signal interpretation is determined manually from each plot. Position 0 values are the proportion of mismatches (labelled "$P_{\text{substitution}}$" on the plots) observed at the very ends of the molecules. Position 0 values are coloured by value, with most saturation for the highest values.

			C-to-T on 5' end		G-to-A on 3' end	
Sample name	Reads	Signal interpretation	Position 0	95% CI	Position 0	95% CI
ELF022_140	2	None	0.000	0.000, 0.975	0.000	0.000, 0.842
ELF027_160	764	Strong	0.082	0.051, 0.132	0.074	0.045, 0.121
ELF027_174	102692	Strong	0.038	0.035, 0.040	0.037	0.035, 0.039
ELF027_177	64	None	0.071	0.017, 0.319	0.111	0.032, 0.348
ELF031_033	0	NA	NA	NA, NA	NA	NA, NA
ELF031_043	11	None	0.333	0.068, 0.806	0.000	0.000, 0.708
ELF031_056	25	None	0.000	0.000, 0.247	0.000	0.000, 0.842
ELF031A_058	6427	Strong	0.147	0.131, 0.165	0.128	0.110, 0.148
ELF031A_072	8155	Strong	0.245	0.225, 0.265	0.209	0.187, 0.232
ELF031A_088	787	Strong	0.219	0.160, 0.291	0.333	0.251, 0.427
ELF031A_107	231	Strong	0.239	0.139, 0.380	0.356	0.194, 0.561
ELF031A_123	764	Strong	0.139	0.093, 0.203	0.243	0.175, 0.328
ELF031A_152	341	Weak	0.130	0.071, 0.230	0.025	0.007, 0.099
ELF031A_177	272	Strong	0.362	0.250, 0.491	0.172	0.092, 0.303
ELF031A_202	49	Weak	0.750	0.400, 0.925	0.222	0.055, 0.626
ELF031A_219	0	NA	NA	NA, NA	NA	NA, NA
ELF031A_281	6	None	0.000	0.000, 0.975	0.000	0.000, 0.708
ELF031A_310	0	NA	NA	NA, NA	NA	NA, NA
ELF032A_047	2812	Strong	0.193	0.166, 0.222	0.149	0.121, 0.183
ELF032A_074	5396	Strong	0.102	0.087, 0.119	0.095	0.081, 0.112
ELF032A_095	3654	Strong	0.211	0.185, 0.238	0.216	0.187, 0.249
ELF032A_117	1387	Strong	0.136	0.103, 0.176	0.120	0.089, 0.161
ELF032A_153	2353	Strong	0.195	0.164, 0.230	0.185	0.152, 0.225
ELF032A_177	6396	Strong	0.180	0.160, 0.202	0.173	0.151, 0.198
ELF033_046	25	None	0.000	0.000, 0.410	0.250	0.039, 0.823
ELF033_075	23	None	0.333	0.099, 0.710	0.000	0.000, 0.369
ELF033_118	3268	Strong	0.137	0.116, 0.161	0.113	0.091, 0.140
ELF033_155	53	None	0.000	0.000, 0.195	0.000	0.000, 0.308
ELF033_176	207	Weak	0.122	0.058, 0.243	0.134	0.067, 0.254
ELF033_183	59	None	0.154	0.063, 0.337	0.000	0.000, 0.336
ELF033_187	267	Strong	0.333	0.229, 0.457	0.190	0.101, 0.329
ELF033_195	54	Weak	0.500	0.278, 0.722	0.333	0.099, 0.710
ELF033_203	48	Weak	0.333	0.139, 0.614	0.231	0.053, 0.665
ELF033A_050	5	None	0.000	0.000, 0.708	0.000	0.000, 0.708
ELF033A_070	18	None	0.333	0.068, 0.806	0.125	0.023, 0.629
ELF033A_098	42	None	0.200	0.043, 0.641	0.167	0.042, 0.524
ELF033A_111	66	Weak	0.308	0.128, 0.581	0.158	0.056, 0.386
ELF033A_126	7	None	0.000	0.000, 0.975	0.333	0.040, 0.913
ELF033A_158	445	Strong	0.276	0.197, 0.371	0.337	0.242, 0.447
ELF034_061	1	None	0.000	0.000, 0.975	0.000	0.000, 0.842
ELF034_079	0	NA	NA	NA, NA	NA	NA, NA
ELF034_094	0	NA	NA	NA, NA	NA	NA, NA
ELF034_132	0	NA	NA	NA, NA	NA	NA, NA
ELF034_157	0	NA	NA	NA, NA	NA	NA, NA
ELF034_177	4	None	0.000	0.000, 0.975	0.000	0.000, 0.975
ELF034_185	2	None	0.000	0.000, 0.842	0.000	0.000, 0.842
ELF034_202	27	None	0.000	0.000, 0.522	0.111	0.020, 0.869
ELF034_219	15	None	0.000	0.000, 0.842	0.000	0.000, 0.708
ELF034A_063	1839	Strong	0.173	0.142, 0.210	0.144	0.111, 0.186
ELF034A_081	524	Strong	0.221	0.155, 0.306	0.217	0.137, 0.328
ELF034A_126	476	Weak	0.078	0.042, 0.142	0.044	0.019, 0.101
ELF034A_146	3	None	0.000	0.000, 0.842	0.000	0.000, 0.975
ELF034A_166	12	None	0.000	0.000, 0.369	0.000	0.000, 0.708
ELF034A_172	5	None	0.500	0.094, 0.906	0.000	0.000, 0.975
ELF034A_183	7	None	1.000	0.158, 0.987	0.000	0.000, 0.842
ELF034A_195	2	None	0.000	0.000, 0.975	0.000	0.000, 0.975
ELF034A_225	0	NA	NA	NA, NA	NA	NA, NA
ELF034A_261	1	None	0.000	0.000, 0.975	0.000	0.000, 0.975
ELF034A_282	3	None	0.000	0.000, 0.975	0.500	0.025, 0.975
ELF039_145	0	NA	NA	NA, NA	NA	NA, NA
ELF039_250	122077	Strong	0.154	0.149, 0.158	0.136	0.132, 0.141
ELF039_321	156752	Strong	0.119	0.115, 0.122	0.102	0.098, 0.106
ELF039_341	216	None	0.133	0.040, 0.383	0.000	0.000, 0.067

Continued on next page

Table 5.2 continued

			C-to-T on 5' end		G-to-A on 3' end	
Sample name	Reads	Signal interpretation	Position 0	95% CI	Position 0	95% CI
ELF039_355	5311	Strong	0.070	0.057, 0.086	0.065	0.052, 0.080
ELF039_384	356	Weak	0.044	0.018, 0.108	0.101	0.055, 0.181
ELF039_415	5725	Strong	0.041	0.031, 0.053	0.053	0.042, 0.067
ELF039_460	2694	Strong	0.036	0.024, 0.055	0.041	0.029, 0.058
ELF039_485	18	None	0.000	0.000, 0.602	0.000	0.000, 0.522
ELF040A_095	2203	Strong	0.066	0.047, 0.091	0.058	0.041, 0.081
ELF040A_112	116	None	0.115	0.042, 0.292	0.056	0.016, 0.190
ELF040A_192	967	Strong	0.088	0.057, 0.133	0.103	0.070, 0.151
ELF040A_208	41	Weak	0.273	0.099, 0.572	0.300	0.092, 0.662
ELF040A_298	1442	Strong	0.114	0.083, 0.156	0.170	0.131, 0.218
ELF040A_350	1444	Strong	0.143	0.109, 0.187	0.082	0.051, 0.130
ELF040A_487	4	None	0.000	0.000, 0.842	0.000	0.000, 0.842
ELF041_087	157	None	0.100	0.036, 0.258	0.038	0.009, 0.199
ELF041_110	617	Strong	0.133	0.084, 0.206	0.216	0.149, 0.305
ELF041_180	1575	Strong	0.143	0.109, 0.186	0.198	0.155, 0.249
ELF041_295	72	None	0.000	0.000, 0.247	0.200	0.054, 0.555
ELF042_065	8916	Strong	0.147	0.131, 0.164	0.145	0.128, 0.164
ELF042_151	2105	Strong	0.152	0.122, 0.189	0.152	0.116, 0.196
ELF042_250	853	Strong	0.126	0.087, 0.181	0.163	0.114, 0.229
ELF042_350	2671	Strong	0.160	0.132, 0.193	0.152	0.122, 0.187
ELF044_090	7029	Strong	0.108	0.093, 0.125	0.087	0.072, 0.104
ELF044_137	635	Strong	0.043	0.020, 0.090	0.052	0.027, 0.100
ELF044A_097	18233	Strong	0.124	0.114, 0.135	0.117	0.107, 0.127
ELF045_090	1759	Strong	0.105	0.079, 0.140	0.096	0.070, 0.132
ELF045_145	28984	Strong	0.058	0.052, 0.064	0.055	0.050, 0.060
ELF045_252	3806	Strong	0.189	0.163, 0.217	0.155	0.128, 0.186
ELF045_346	1036	Strong	0.170	0.128, 0.222	0.165	0.115, 0.230
ELF045_450	512	Strong	0.160	0.103, 0.242	0.131	0.080, 0.209
ELF045_522	2058	Strong	0.159	0.128, 0.196	0.214	0.173, 0.263
ELF046A_270	0	NA	NA	NA, NA	NA	NA, NA
ELF047_070	63	None	0.071	0.017, 0.319	0.125	0.035, 0.396
ELF047_150	5522	Strong	0.116	0.099, 0.136	0.124	0.105, 0.145
ELF047_241	2694	Strong	0.206	0.176, 0.241	0.166	0.131, 0.209
ELF047_274	405	Strong	0.135	0.081, 0.218	0.191	0.110, 0.313
ELF047_325	4531	Strong	0.040	0.029, 0.054	0.038	0.029, 0.050
ELF047_386	1882	Strong	0.057	0.038, 0.085	0.038	0.025, 0.057
ELF047A_050	1964	Strong	0.199	0.163, 0.242	0.129	0.095, 0.174
ELF047A_150	4	None	0.000	0.000, 0.842	0.000	0.000, 0.975
ELF047A_256	20	None	0.000	0.000, 0.459	0.667	0.194, 0.932
ELF047A_354	17	None	0.500	0.094, 0.906	0.000	0.000, 0.602
ELF049_295	1244	Strong	0.165	0.126, 0.215	0.157	0.116, 0.211
ELF049_361	136	Strong	0.207	0.099, 0.386	0.217	0.087, 0.458
ELF050_250	1	None	0.000	0.000, 0.975	0.000	0.000, 0.975
ELF050_354	2	None	0.000	0.000, 0.842	0.000	0.000, 0.842
ELF050_450	9	None	0.000	0.000, 0.842	0.000	0.000, 0.602
ELF050_595	3	None	0.000	0.000, 0.975	0.000	0.000, 0.975
ELF051_096	53	None	0.063	0.015, 0.287	0.111	0.020, 0.618
ELF051_120	11	Weak	1.000	0.158, 0.987	0.500	0.147, 0.853
ELF051_151	19	None	0.333	0.099, 0.710	0.400	0.048, 0.924
ELF051_196	222	Weak	0.113	0.056, 0.216	0.140	0.067, 0.274
ELF051_255	2362	Strong	0.216	0.183, 0.253	0.210	0.170, 0.258
ELF051_292	8483	Strong	0.084	0.073, 0.098	0.087	0.075, 0.101
ELF053_179	0	NA	NA	NA, NA	NA	NA, NA
ELF053_214	3	None	0.000	0.000, 0.708	0.000	0.000, 0.975
ELF053_275	1	None	0.000	0.000, 0.975	0.000	0.000, 0.975
ELF053_289	19	Weak	0.000	0.000, 0.602	0.200	0.039, 0.683
ELF053_336	4	None	0.000	0.000, 0.602	1.000	0.025, 0.975
ELF054_058	2217	Strong	0.061	0.043, 0.087	0.092	0.071, 0.119
ELF054_140	81	None	0.111	0.034, 0.331	0.077	0.018, 0.326
ELF054_182	460	Weak	0.071	0.036, 0.140	0.039	0.017, 0.090
ELF054_268	31	None	0.000	0.000, 0.410	0.286	0.091, 0.630
ELF054_291	6	None	0.000	0.000, 0.522	0.000	0.000, 0.975
ELF054_315	23	None	0.167	0.037, 0.579	0.200	0.039, 0.683
ELF054_330	8	None	0.000	0.000, 0.522	0.000	0.000, 0.842
ELF054_356	1	None	0.000	0.000, 0.975	0.000	0.000, 0.975
ELF059_210	6243	Strong	0.097	0.083, 0.114	0.089	0.073, 0.107
ELF059_230	1404	Strong	0.069	0.045, 0.104	0.054	0.034, 0.084
ELF059_270	2572	Strong	0.091	0.071, 0.117	0.094	0.070, 0.124
ELF059_280	93	None	0.037	0.009, 0.183	0.000	0.000, 0.132
ELF059_337	1204	Strong	0.038	0.021, 0.069	0.063	0.041, 0.095
ELF059_359	4925	Strong	0.131	0.112, 0.153	0.156	0.133, 0.181
ELF059_378	129	None	0.115	0.042, 0.292	0.132	0.056, 0.287
ELF059A_135	27206	Strong	0.131	0.122, 0.140	0.103	0.095, 0.112
ELF059A_190	18149	Strong	0.113	0.104, 0.124	0.099	0.089, 0.110
ELF059A_250	21222	Strong	0.050	0.044, 0.057	0.048	0.042, 0.054
ELF059A_320	103	None	0.000	0.000, 0.206	0.100	0.040, 0.238
ELF059A_355	55046	Strong	0.150	0.144, 0.157	0.131	0.125, 0.138
ELF060_250	2287	Strong	0.048	0.033, 0.070	0.065	0.048, 0.088
ELF060_350	6228	Strong	0.092	0.078, 0.108	0.066	0.054, 0.080
ELF060_420	6348	Strong	0.068	0.056, 0.082	0.069	0.057, 0.082
ELF060_465	1887	Strong	0.068	0.047, 0.098	0.098	0.074, 0.130

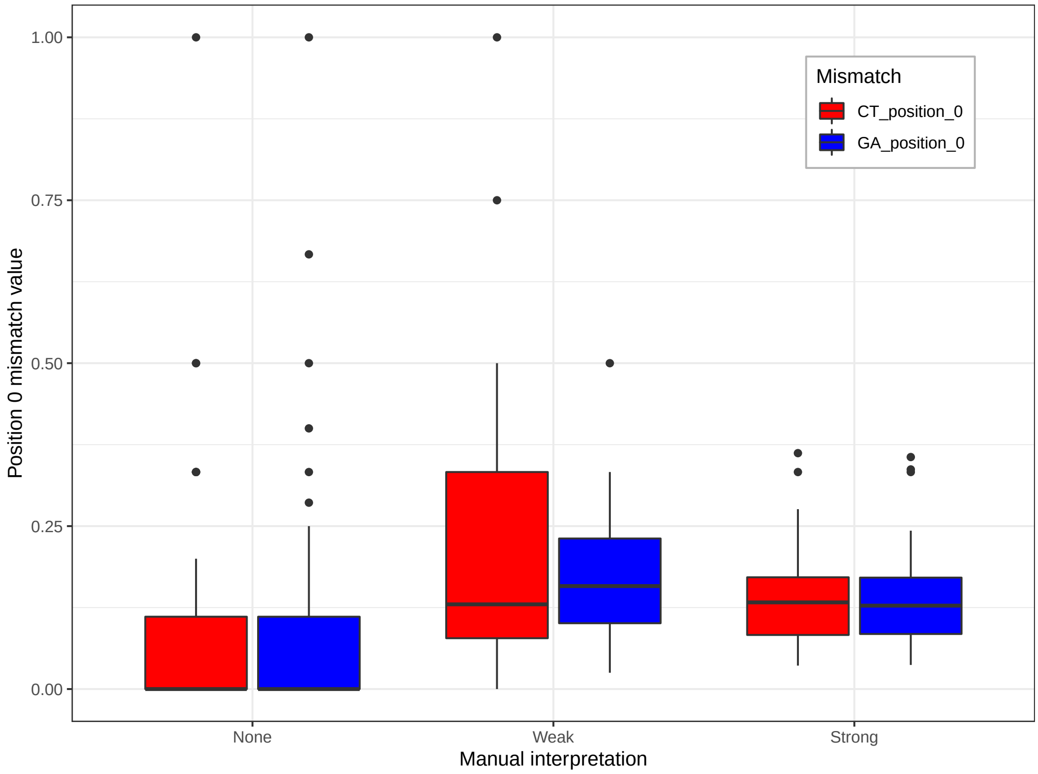


Figure 5.9: Comparing Embryophyta position-zero values and manual interpretation. Central lines show the median, boxes extend to the first and third quartiles, and whiskers to 1.5 times the inter-quartile range.

The first analysis from this data evaluates the agreement between the position-zero mismatch values and the signal interpreted from the plots. Figure 5.9 plots them against each other. C-to-T and G-to-A position-zero values show similar patterns. Most samples interpreted as having no signal have position-zero values close to zero, as expected with no visible upticks, but outliers extend all the way to a probability of one. Samples with a weak signal also have an outlier at one and show a greater general spread of data, but have a higher median that suggests upticks did occur. Finally, samples with a strong signal have the same higher median, but with a narrow inter-quartile range and few outliers. There is also even less difference between C-to-T and G-to-A than for absent or weak signals.

In this data, a strong signal can be reliably characterised as having a position-zero value of approximately 0.08-0.17 for both relevant types of mismatch. However, this range is also occupied by many samples interpreted as having weak signals, and some samples with no signal at all. There is too much variation in position-zero value among weak and absent signals for it to reliably capture signal strength alone.

The variation in position-zero values can largely be explained by read count, which is known to affect MetaDamage output. Figure 5.10 plots position-zero mismatch values against natural-log-transformed read count (dotted lines mark raw read counts). More extreme mismatch values are associated with lower counts. This is expected: if there is only one sequence in the MetaDamage analysis, the estimated probability of a given mismatch can either be 0 or 1. With two sequences, the probability could be 0.5.

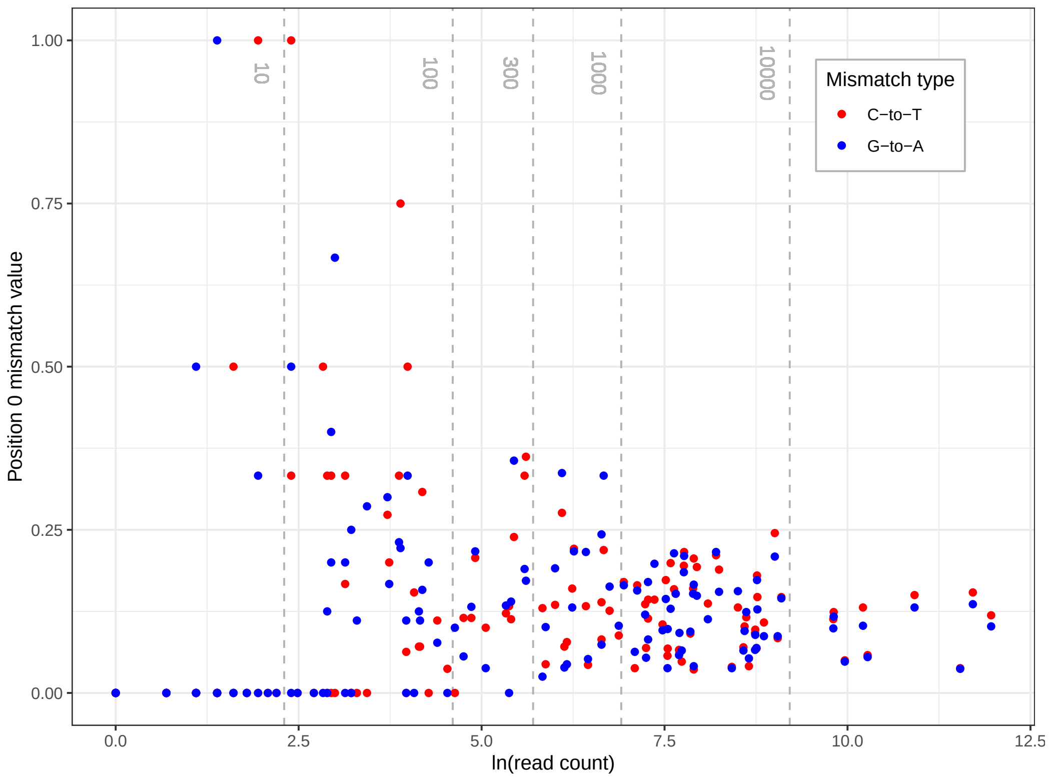


Figure 5.10: Comparing Embryophyta read count and position-zero mismatch values. Read counts are natural-log-transformed, but dotted lines mark raw values to aid interpretation. The recommended minimum read count is 100-300.

As read count increases, so can the precision of the mismatch value. Sensitivity tests performed by Rosie Everett suggest using at least 100-300 sequences; MetaDamage failed to find a reliable signal from 100 sequences, but 300 was sufficient. 300 reads is marked on figure 5.10 along with orders of magnitude from 10. All extremely high position-zero mismatch values occur at fewer than 100 reads, and only values two of zero (which are more likely to be genuine than high values) subsequently occur before 300 reads.

This suggests that combining read count with position-zero values, perhaps as a filter excluding samples with <300 reads, may produce a suitable quantification of MetaDamage output. However, while the position-zero values can offer a route to quantification, their primary function is to enable the calculation of 95% confidence intervals, which give a measure of variation in the data. The intervals are as important for interpreting the plots as the upticks themselves.

The manual interpretation of plots is also affected by read count but, perhaps because it takes the confidence intervals into account, the relationship is more straightforward. Figure 5.11 shows that apparent signal strength broadly increases with read count. All of the samples with no signal have <300 reads and the vast majority have <100. Samples with weak signals could have up to around 500 reads, but again, most have <300. Finally, although strong signals were shown by samples with a very wide range of read counts, samples with >1000 reads appeared exclusively strong and no strong signal was found with <100 reads. Together, these categories generally concur

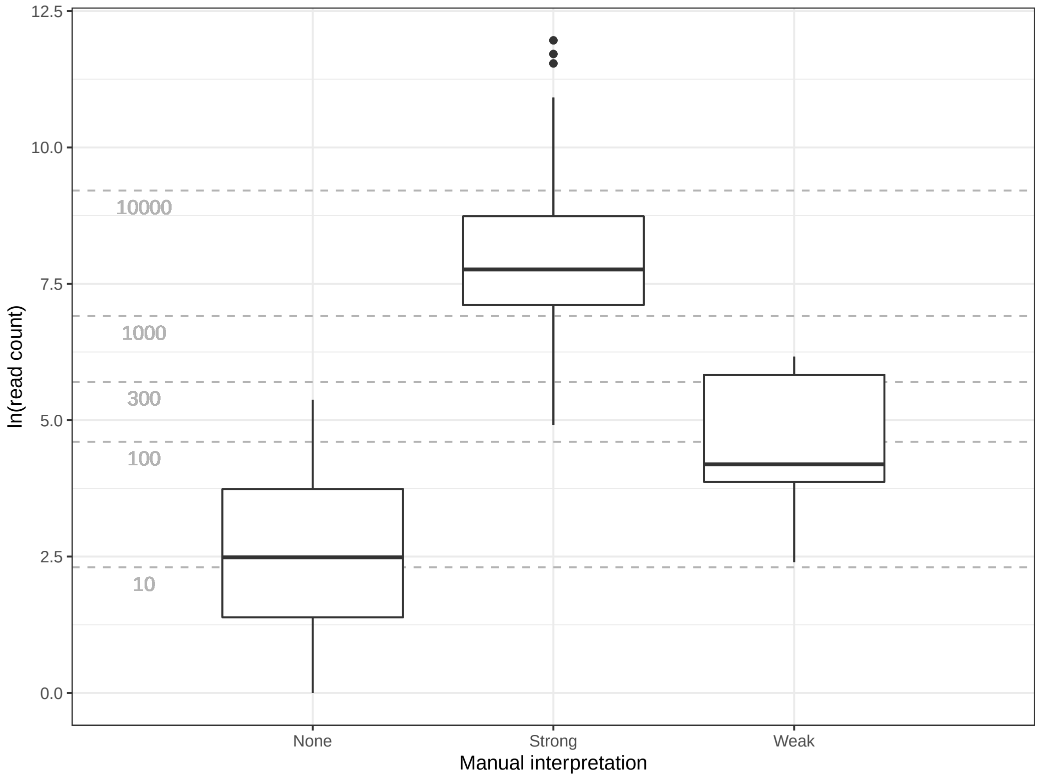


Figure 5.11: Comparing Embryophyta read count and manual interpretation. Read counts are natural-log-transformed, but dotted lines mark raw values to aid interpretation. The recommended minimum read count is 100-300. Central lines show the median, boxes extend to the first and third quartiles, and whiskers to 1.5 times the inter-quartile range.

with the higher recommended minimum of 300 reads. Stronger signals may occur with less data than this, but all samples with at least 300 reads showed at least some signal.

The presence of a MetaDamage signal in all samples with sufficient reads is an important result. Because the absence or ambiguity of signal in samples with fewer reads can be explained by read count, there is no evidence of modern contamination in Embryophyta. Higher-yield samples present a convincing damage signal. Results for lower-yield samples are inconclusive, but in general, these samples also lacked data for reliable ecological reconstruction. Higher-yield samples show both damage signal and, generally, are useful for reconstruction. Therefore, plant communities inferred with confidence in Chapter 4 (Taxonomic Results) can be taken as authentic.

5.1.4 Summary of MetaDamage authentication

Damage-based authentication was performed using the MetaDamage analysis, which is designed for metagenomic samples but requires sufficient reads. Therefore, while Embryophyta data overall produced a strong signal of age-associated DNA damage, many samples did not have enough Embryophyta data for a conclusive result. However, data quantity also limits environmental reconstruction. Samples with enough data for confident reconstructions in Chapter 4 (Taxonomic Results) also passed damage-based authentication. Reconstructions based on smaller read counts may be inconclusive, but the consistent signal from higher-yield samples suggests that modern contamination is unlikely to be a significant issue.

The Metazoa reads are universally few in number at sample level. Combining data from across samples, including with preliminary deep-sequencing data, revealed a potential damage signal obscured by a high background mismatch rate. These excessive mismatches may be due to a greater influence of over-represented taxa when finding reference sequences for Metazoa reads. However, it is also quite possible that the damage signal has also been directly reduced by modern contamination, particularly human. On the rare occasions that Metazoa reads could inform environmental reconstruction, they should be taken with additional care, especially if the reads have been assigned to likely contaminant taxa.

Finally, this section has mainly focused on metagenomic analysis, evaluating all Embryophyta or Metazoa reads per sample. MetaDamage will be used again to evaluate more specific Embryophyta taxa that may influence interpretations of climate or human activity in Chapter 6 (Mesophilic taxa and human disturbance indicators). Here, we move on to a second means of authentication: stratification.

5.2 Stratification analysis

The stratification analysis is a new approach that searches for positive evidence of DNA stratification - a lack of movement between strata - between adjacent samples in a core. It uses *proportional* read counts (in an attempt to account for data yield) to calculate an index of change and a corresponding p-value for each taxon for each pair of samples. This chapter presents and interprets those measures; see chapter 2 (Main materials and methods) for detail of the process. Results for each core are first presented individually, followed by validation against data from other environmental proxies in subsection 5.2.22, and a summary of results in subsection 5.2.23.

A combination of highly significant increases and decreases indicates that the two samples have different DNA profiles. This is consistent with a lack of DNA movement. However, DNA movement is not the only process that could produce similar DNA assemblages. They could also arise from a genuinely similar ancient plant community. Rather than being discounted, cores with a high degree of similarity between DNA profiles are therefore simply flagged for further investigation. Note also that a significant change in the DNA profile of a given taxon does not necessarily mean a significant change in biomass. Unlike in Chapter 4 (Taxonomic results), these results will not be used to infer changes in plant community.

There are two further limitations to consider when interpreting stratification analysis output. First, though the minimum read limit excludes most rare taxa, some remain. Rare taxa often manifest as repeated appearances and disappearances that produce over-weighted indices of change (infinity and negative infinity respectively). These are exaggerated differences between DNA profiles. Fortunately, associated p-values are usually not significant, so their impact is limited.

The second limitation may also result in over-interpretation of some changes. Chapter 4 (Taxonomic results) discussed a trend among *Zostera*-dominated samples where, as data yield increases, a greater proportion of reads are assigned to higher taxa and fewer to *Zostera* and Zosteraceae. Although the cause for this pattern is not entirely clear, it appears to be an artefact of sequence assignment. Even with cumulative counting, where reads assigned to child taxa are also counted for parent taxa, it appeared to generate changes in proportional read counts that have more to do with data yield than DNA profile. Reciprocal changes between *Zostera*/Zosteraceae and Alismatales associated with large corresponding increases or decreases in read count are therefore discounted (parent taxa above Alismatales had already been removed).

Results from each sediment core are presented using a set of three tables (e.g. figure 5.12). All rows refer to taxa, with taxon information shown on the left in the same format as the taxonomic profiles (Chapter 4). The first table shows cumulative read counts per sample. Recall that in cumulative counting, reads assigned to a taxon also count towards any of its parent taxa in the analysis. In the second and third tables, columns refer not to samples, but to comparisons between pairs of samples, with the deepest on the left.

Cells in the second table contain an index of change for that comparison, where negative indices (blue) show a proportional decrease for that taxon, and positive indices (red) show an increase.

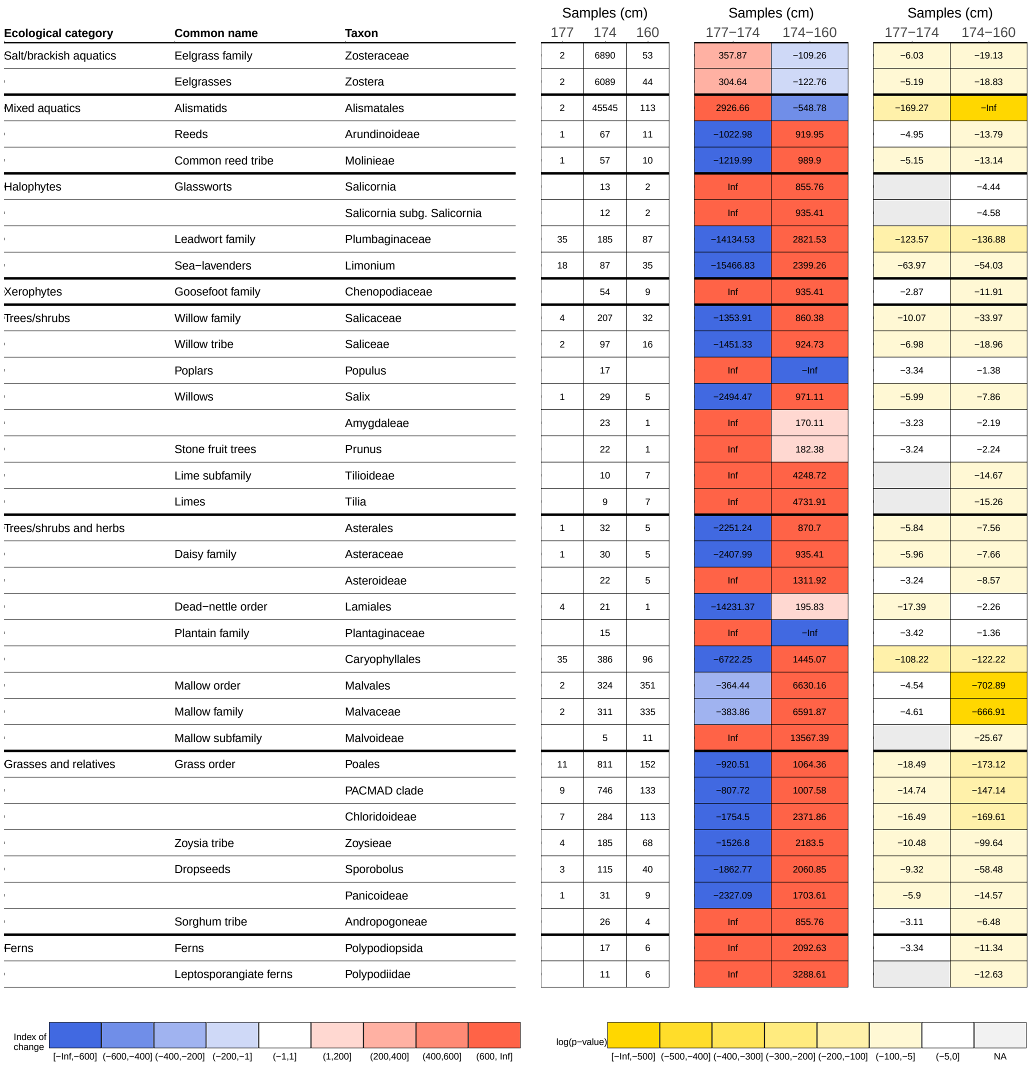
The third table contains corresponding p-values. These are natural-log-transformed to emphasise differences between very significant values. ln(p-values) of -5 and below are coloured in shades of gold; ln(0.01) = -4.61, so all coloured tiles are statistically significant to a reasonable degree. Grey tiles ("NA") indicate comparisons involving fewer than the minimum of 14 reads, so no p-value was calculated. Taxa for which no comparisons had sufficient data are omitted.

The stratification analysis was performed on European Embryophyta read counts, counted cumulatively. Very high taxa were omitted (see list in Chapter 2). Cores ELF022 and ELF044A were excluded because they contain only a single sample, and cores ELF034, ELF031, ELF046A, ELF050, and ELF053 were excluded because they contain insufficient data.

5.2.1 ELF027

DNA profiles in ELF027 show among the most changes across all cores. There are only three samples, but the taxonomic profile (Chapter 4: Taxonomic results) suggests an environmental change between each one, from terrestrial to brackish and back to terrestrial. Data yield also varies dramatically, with an extremely high peak in the middle sample (174). This can explain the many appearances ("Inf" indices of change) of rare taxa between samples 177 and 174. We would not expect to observe these taxa in lower-yield samples and their read counts in 174 are proportionally very small, so although the indices of change are of the highest magnitude, none comes with a significant p-value. These taxa do not demonstrate stratification.

Instead, we must look to more frequent taxa, most of which also show dramatic indices of change but now with significant p-values. In fact, 76% of p-values in ELF027 are significant. This among the highest rate of all cores. Recall from Chapter 2 (Main materials and methods) that we can summarise the completeness of the stratification signal across the core by checking for a significant increase and decrease between each pair of samples. A complete signal shows at least one instance of stratification in both directions between every pair. This is the case here: both sample comparisons have at least one significant increase and decrease, giving a signal completeness of 100%. The evidence for stratification is particularly strong in this core.



			Samples (cm)			Samples (cm)		Samples (cm)	
Ecological category	**Common name**	**Taxon**	177	174	160	177−174	174−160	177−174	174−160
Salt/brackish aquatics	Eelgrass family	Zosteraceae	2	6890	53	357.87	−109.26	−6.03	−19.13
	Eelgrasses	Zostera	2	6089	44	304.64	−122.76	−5.19	−18.83
Mixed aquatics	Alismatids	Alismatales	2	45545	113	2926.66	−548.78	−169.27	−Inf
	Reeds	Arundinoideae	1	67	11	−1022.98	919.95	−4.95	−13.79
	Common reed tribe	Molinieae	1	57	10	−1219.99	989.9	−5.15	−13.14
Halophytes	Glassworts	Salicornia		13	2	Inf	855.76		−4.44
		Salicornia subg. Salicornia		12	2	Inf	935.41		−4.58
	Leadwort family	Plumbaginaceae	35	185	87	−14134.53	2821.53	−123.57	−136.88
	Sea−lavenders	Limonium	18	87	35	−15466.83	2399.26	−63.97	−54.03
Xerophytes	Goosefoot family	Chenopodiaceae		54	9	Inf	935.41	−2.87	−11.91
Trees/shrubs	Willow family	Salicaceae	4	207	32	−1353.91	860.38	−10.07	−33.97
	Willow tribe	Saliceae	2	97	16	−1451.33	924.73	−6.98	−18.96
	Poplars	Populus		17		Inf	−Inf	−3.34	−1.38
	Willows	Salix	1	29	5	−2494.47	971.11	−5.99	−7.86
		Amygdaleae		23	1	Inf	170.11	−3.23	−2.19
	Stone fruit trees	Prunus		22	1	Inf	182.38	−3.24	−2.24
	Lime subfamily	Tilioideae		10	7	Inf	4248.72		−14.67
	Limes	Tilia		9	7	Inf	4731.91		−15.26
Trees/shrubs and herbs		Asterales	1	32	5	−2251.24	870.7	−5.84	−7.56
	Daisy family	Asteraceae	1	30	5	−2407.99	935.41	−5.96	−7.66
		Asteroideae		22	5	Inf	1311.92	−3.24	−8.57
	Dead−nettle order	Lamiales	4	21	1	−14231.37	195.83	−17.39	−2.26
	Plantain family	Plantaginaceae		15		Inf	−Inf	−3.42	−1.36
		Caryophyllales	35	386	96	−6722.25	1445.07	−108.22	−122.22
	Mallow order	Malvales	2	324	351	−364.44	6630.16	−4.54	−702.89
	Mallow family	Malvaceae	2	311	335	−383.86	6591.87	−4.61	−666.91
	Mallow subfamily	Malvoideae		5	11	Inf	13567.39		−25.67
Grasses and relatives	Grass order	Poales	11	811	152	−920.51	1064.36	−18.49	−173.12
		PACMAD clade	9	746	133	−807.72	1007.58	−14.74	−147.14
		Chloridoideae	7	284	113	−1754.5	2371.86	−16.49	−169.61
	Zoysia tribe	Zoysieae	4	185	68	−1526.8	2183.5	−10.48	−99.64
	Dropseeds	Sporobolus	3	115	40	−1862.77	2060.85	−9.32	−58.48
		Panicoideae	1	31	9	−2327.09	1703.61	−5.9	−14.57
	Sorghum tribe	Andropogoneae		26	4	Inf	855.76	−3.11	−6.48
Ferns	Ferns	Polypodiopsida		17	6	Inf	2092.63	−3.34	−11.34
	Leptosporangiate ferns	Polypodiidae		11	6	Inf	3288.61		−12.63

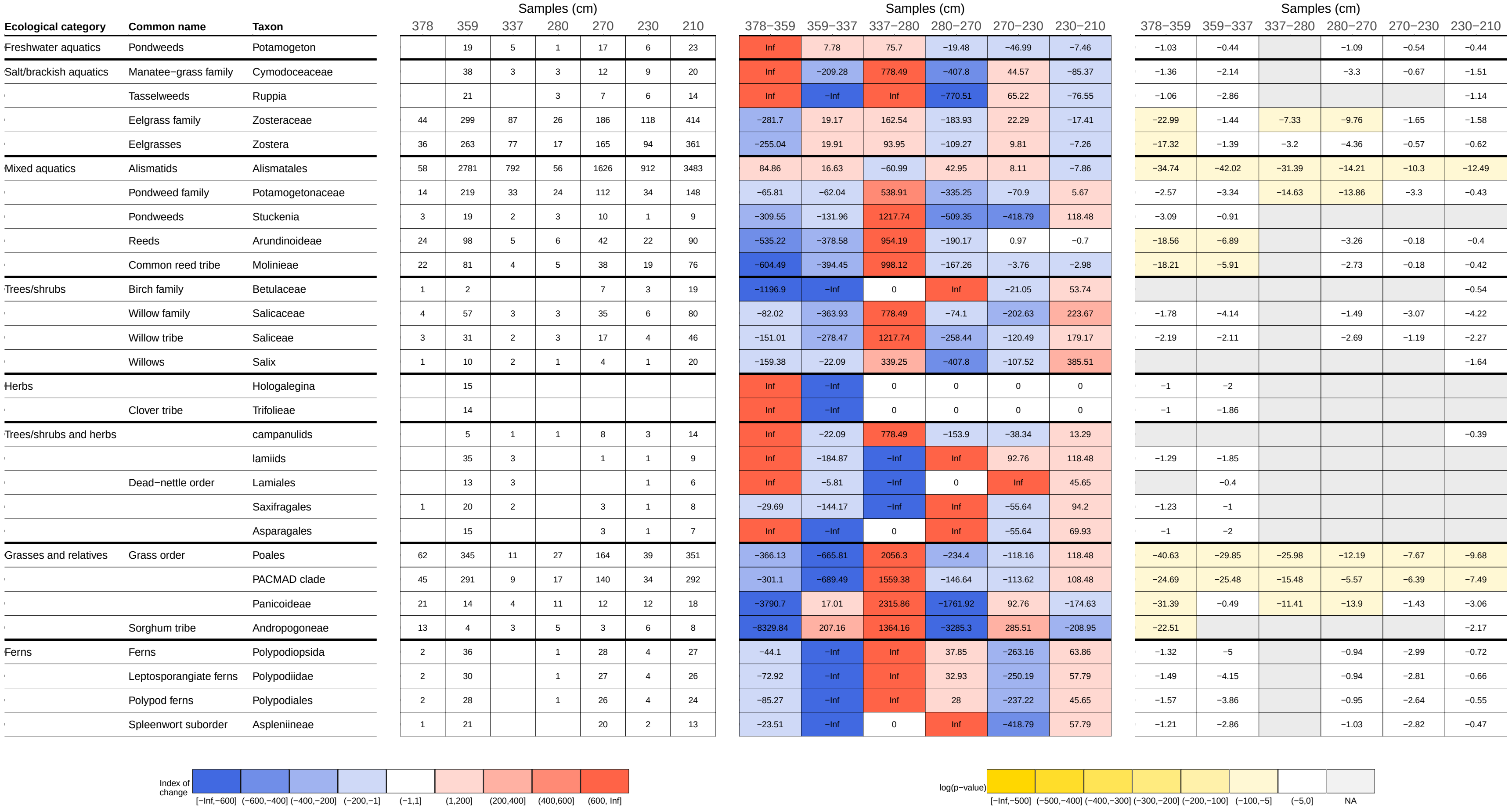
Figure 5.12: Stratification analysis for ELF027: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 100%.

5.2.2 ELF059

The seven samples in ELF059 appear taxonomically similar: mostly *Zostera* with diverse low-frequency terrestrial taxa. However, like in ELF027, there are many dramatic indices of change for rare taxa that can be explained by variable data yield. For example, the string of appearances and disappearances in the first three comparisons are probably due to the particularly high yield in sample 359 and low yield in 337. Again, these changes rarely involve enough data to calculate p-values and those that exist are not significant.

The significant changes in ELF059 are instead found mostly among aquatics and grasses, although some changes in *Zostera*, Zosteraceae, and Alismatales may also have been influenced by the variable read count. For example, from sample 378 to 359, read count increases substantially from 129 to nearly 500. As expected from the trend described at the end of Chapter 4 (Taxonomic results), we also see a significant decrease in *Zostera* and Zosteraceae but an increase in Alismatales (a higher taxon). These three changes may not actually reflect meaningful changes in the DNA profile.

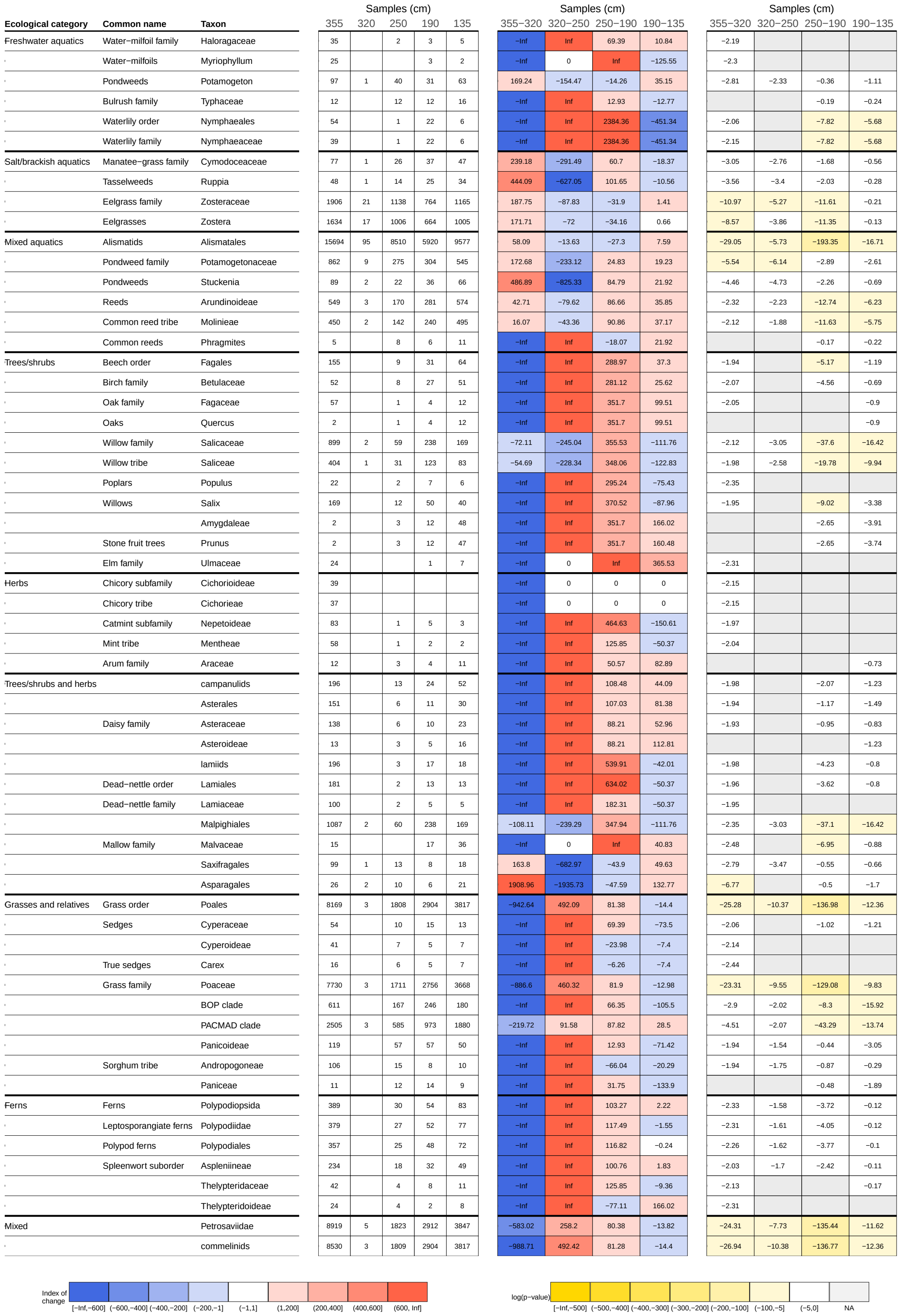
However, there are also significant changes in several other taxa across the core. Unlike in ELF027, there are no particularly low p-values (below ln(-100); beyond the first coloured bin), but very low p-values are rare across the cores. Even discounting the suspect aquatics, nearly every comparison in ELF059 has at least one significant increase and decrease, resulting in a signal completeness of 83.33%. There is therefore still good evidence of stratification.



Ecological category	Common name	Taxon	Samples (cm)							Samples (cm)						Samples (cm)					
			378	359	337	280	270	230	210	378−359	359−337	337−280	280−270	270−230	230−210	378−359	359−337	337−280	280−270	270−230	230−210
Freshwater aquatics	Pondweeds	Potamogeton		19	5	1	17	6	23	Inf	7.78	75.7	−19.48	−46.99	−7.46	−1.03	−0.44		−1.09	−0.54	−0.44
Salt/brackish aquatics	Manatee−grass family	Cymodoceaceae		38	3	3	12	9	20	Inf	−209.28	778.49	−407.8	44.57	−85.37	−1.36	−2.14		−3.3	−0.67	−1.51
	Tasselweeds	Ruppia		21		3	7	6	14	Inf	−Inf	Inf	−770.51	65.22	−76.55	−1.06	−2.86				−1.14
	Eelgrass family	Zosteraceae	44	299	87	26	186	118	414	−281.7	19.17	162.54	−183.93	22.29	−17.41	−22.99	−1.44	−7.33	−9.76	−1.65	−1.58
	Eelgrasses	Zostera	36	263	77	17	165	94	361	−255.04	19.91	93.95	−109.27	9.81	−7.26	−17.32	−1.39	−3.2	−4.36	−0.57	−0.62
Mixed aquatics	Alismatids	Alismatales	58	2781	792	56	1626	912	3483	84.86	16.63	−60.99	42.95	8.11	−7.86	−34.74	−42.02	−31.39	−14.21	−10.3	−12.49
	Pondweed family	Potamogetonaceae	14	219	33	24	112	34	148	−65.81	−62.04	538.91	−335.25	−70.9	5.67	−2.57	−3.34	−14.63	−13.86	−3.3	−0.43
	Pondweeds	Stuckenia	3	19	2	3	10	1	9	−309.55	−131.96	1217.74	−509.35	−418.79	118.48	−3.09	−0.91				
	Reeds	Arundinoideae	24	98	5	6	42	22	90	−535.22	−378.58	954.19	−190.17	0.97	−0.7	−18.56	−6.89		−3.26	−0.18	−0.4
	Common reed tribe	Molinieae	22	81	4	5	38	19	76	−604.49	−394.45	998.12	−167.26	−3.76	−2.98	−18.21	−5.91		−2.73	−0.18	−0.42
Trees/shrubs	Birch family	Betulaceae	1	2			7	3	19	−1196.9	−Inf	0	Inf	−21.05	53.74						−0.54
	Willow family	Salicaceae	4	57	3	3	35	6	80	−82.02	−363.93	778.49	−74.1	−202.63	223.67	−1.78	−4.14		−1.49	−3.07	−4.22
	Willow tribe	Saliceae	3	31	2	3	17	4	46	−151.01	−278.47	1217.74	−258.44	−120.49	179.17	−2.19	−2.11		−2.69	−1.19	−2.27
	Willows	Salix	1	10	2	1	4	1	20	−159.38	−22.09	339.25	−407.8	−107.52	385.51						−1.64
Herbs		Hologalegina		15						Inf	−Inf	0	0	0	0	−1	−2				
	Clover tribe	Trifolieae		14						Inf	−Inf	0	0	0	0	−1	−1.86				
Trees/shrubs and herbs		campanulids		5	1	1	8	3	14	Inf	−22.09	778.49	−153.9	−38.34	13.29						−0.39
		lamiids		35	3		1	1	9	Inf	−184.87	−Inf	Inf	92.76	118.48	−1.29	−1.85				
	Dead−nettle order	Lamiales		13	3			1	6	Inf	−5.81	−Inf	0	Inf	45.65		−0.4				
		Saxifragales	1	20	2		3	1	8	−29.69	−144.17	−Inf	Inf	−55.64	94.2	−1.23	−1				
		Asparagales		15			3	1	7	Inf	−Inf	0	Inf	−55.64	69.93	−1	−2				
Grasses and relatives	Grass order	Poales	62	345	11	27	164	39	351	−366.13	−665.81	2056.3	−234.4	−118.16	118.48	−40.63	−29.85	−25.98	−12.19	−7.67	−9.68
		PACMAD clade	45	291	9	17	140	34	292	−301.1	−689.49	1559.38	−146.64	−113.62	108.48	−24.69	−25.48	−15.48	−5.57	−6.39	−7.49
		Panicoideae	21	14	4	11	12	12	18	−3790.7	17.01	2315.86	−1761.92	92.76	−174.63	−31.39	−0.49	−11.41	−13.9	−1.43	−3.06
	Sorghum tribe	Andropogoneae	13	4	3	5	3	6	8	−8329.84	207.16	1364.16	−3285.3	285.51	−208.95	−22.51					−2.17
Ferns	Ferns	Polypodiopsida	2	36		1	28	4	27	−44.1	−Inf	Inf	37.85	−263.16	63.86	−1.32	−5		−0.94	−2.99	−0.72
	Leptosporangiate ferns	Polypodiidae	2	30		1	27	4	26	−72.92	−Inf	Inf	32.93	−250.19	57.79	−1.49	−4.15		−0.94	−2.81	−0.66
	Polypod ferns	Polypodiales	2	28		1	26	4	24	−85.27	−Inf	Inf	28	−237.22	45.65	−1.57	−3.86		−0.95	−2.64	−0.55
	Spleenwort suborder	Aspleniineae	1	21			20	2	13	−23.51	−Inf	0	Inf	−418.79	57.79	−1.21	−2.86		−1.03	−2.82	−0.47

Figure 5.13: Stratification analysis for ELF059: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 83.33%.

5.2.3 ELF059A



			Samples (cm)					Samples (cm)				Samples (cm)			
Ecological category	**Common name**	**Taxon**	355	320	250	190	135	355−320	320−250	250−190	190−135	355−320	320−250	250−190	190−135
Freshwater aquatics	Water−milfoil family	Haloragaceae	35		2	3	5	−Inf	Inf	69.39	10.84	−2.19			
	Water−milfoils	Myriophyllum	25			3	2	−Inf	0	Inf	−125.55	−2.3			
	Pondweeds	Potamogeton	97	1	40	31	63	169.24	−154.47	−14.26	35.15	−2.81	−2.33	−0.36	−1.11
	Bulrush family	Typhaceae	12		12	12	16	−Inf	Inf	12.93	−12.77			−0.19	−0.24
	Waterlily order	Nymphaeales	54		1	22	6	−Inf	Inf	2384.36	−451.34	−2.06		−7.82	−5.68
	Waterlily family	Nymphaeaceae	39		1	22	6	−Inf	Inf	2384.36	−451.34	−2.15		−7.82	−5.68
Salt/brackish aquatics	Manatee−grass family	Cymodoceaceae	77	1	26	37	47	239.18	−291.49	60.7	−18.37	−3.05	−2.76	−1.68	−0.56
	Tasselweeds	Ruppia	48	1	14	25	34	444.09	−627.05	101.65	−10.56	−3.56	−3.4	−2.03	−0.28
	Eelgrass family	Zosteraceae	1906	21	1138	764	1165	187.75	−87.83	−31.9	1.41	−10.97	−5.27	−11.61	−0.21
	Eelgrasses	Zostera	1634	17	1006	664	1005	171.71	−72	−34.16	0.66	−8.57	−3.86	−11.35	−0.13
Mixed aquatics	Alismatids	Alismatales	15694	95	8510	5920	9577	58.09	−13.63	−27.3	7.59	−29.05	−5.73	−193.35	−16.71
	Pondweed family	Potamogetonaceae	862	9	275	304	545	172.68	−233.12	24.83	19.23	−5.54	−6.14	−2.89	−2.61
	Pondweeds	Stuckenia	89	2	22	36	66	486.89	−825.33	84.79	21.92	−4.46	−4.73	−2.26	−0.69
	Reeds	Arundinoideae	549	3	170	281	574	42.71	−79.62	86.66	35.85	−2.32	−2.23	−12.74	−6.23
	Common reed tribe	Molinieae	450	2	142	240	495	16.07	−43.36	90.86	37.17	−2.12	−1.88	−11.63	−5.75
	Common reeds	Phragmites	5		8	6	11	−Inf	Inf	−18.07	21.92			−0.17	−0.22
Trees/shrubs	Beech order	Fagales	155		9	31	64	−Inf	Inf	288.97	37.3	−1.94		−5.17	−1.19
	Birch family	Betulaceae	52		8	27	51	−Inf	Inf	281.12	25.62	−2.07		−4.56	−0.69
	Oak family	Fagaceae	57		1	4	12	−Inf	Inf	351.7	99.51	−2.05			−0.9
	Oaks	Quercus	2		1	4	12	−Inf	Inf	351.7	99.51				−0.9
	Willow family	Salicaceae	899	2	59	238	169	−72.11	−245.04	355.53	−111.76	−2.12	−3.05	−37.6	−16.42
	Willow tribe	Saliceae	404	1	31	123	83	−54.69	−228.34	348.06	−122.83	−1.98	−2.58	−19.78	−9.94
	Poplars	Populus	22		2	7	6	−Inf	Inf	295.24	−75.43	−2.35			
	Willows	Salix	169		12	50	40	−Inf	Inf	370.52	−87.96	−1.95		−9.02	−3.38
		Amygdaleae	2		3	12	48	−Inf	Inf	351.7	166.02			−2.65	−3.91
	Stone fruit trees	Prunus	2		3	12	47	−Inf	Inf	351.7	160.48			−2.65	−3.74
	Elm family	Ulmaceae	24			1	7	−Inf	0	Inf	365.53	−2.31			
Herbs	Chicory subfamily	Cichorioideae	39					−Inf	0	0	0	−2.15			
	Chicory tribe	Cichorieae	37					−Inf	0	0	0	−2.15			
	Catmint subfamily	Nepetoideae	83		1	5	3	−Inf	Inf	464.63	−150.61	−1.97			
	Mint tribe	Mentheae	58		1	2	2	−Inf	Inf	125.85	−50.37	−2.04			
	Arum family	Araceae	12		3	4	11	−Inf	Inf	50.57	82.89				−0.73
Trees/shrubs and herbs		campanulids	196		13	24	52	−Inf	Inf	108.48	44.09	−1.98		−2.07	−1.23
		Asterales	151		6	11	30	−Inf	Inf	107.03	81.38	−1.94		−1.17	−1.49
	Daisy family	Asteraceae	138		6	10	23	−Inf	Inf	88.21	52.96	−1.93		−0.95	−0.83
		Asteroideae	13		3	5	16	−Inf	Inf	88.21	112.81				−1.23
		lamiids	196		3	17	18	−Inf	Inf	539.91	−42.01	−1.98		−4.23	−0.8
	Dead−nettle order	Lamiales	181		2	13	13	−Inf	Inf	634.02	−50.37	−1.96		−3.62	−0.8
	Dead−nettle family	Lamiaceae	100		2	5	5	−Inf	Inf	182.31	−50.37	−1.95			
		Malpighiales	1087	2	60	238	169	−108.11	−239.29	347.94	−111.76	−2.35	−3.03	−37.1	−16.42
	Mallow family	Malvaceae	15			17	36	−Inf	0	Inf	40.83	−2.48		−6.95	−0.88
		Saxifragales	99	1	13	8	18	163.8	−682.97	−43.9	49.63	−2.79	−3.47	−0.55	−0.66
		Asparagales	26	2	10	6	21	1908.96	−1935.73	−47.59	132.77	−6.77		−0.5	−1.7
Grasses and relatives	Grass order	Poales	8169	3	1808	2904	3817	−942.64	492.09	81.38	−14.4	−25.28	−10.37	−136.98	−12.36
	Sedges	Cyperaceae	54		10	15	13	−Inf	Inf	69.39	−73.5	−2.06		−1.02	−1.21
		Cyperoideae	41		7	5	7	−Inf	Inf	−23.98	−7.4	−2.14			
	True sedges	Carex	16		6	5	7	−Inf	Inf	−6.26	−7.4	−2.44			
	Grass family	Poaceae	7730	3	1711	2756	3668	−886.6	460.32	81.9	−12.98	−23.31	−9.55	−129.08	−9.83
		BOP clade	611		167	246	180	−Inf	Inf	66.35	−105.5	−2.9	−2.02	−8.3	−15.92
		PACMAD clade	2505	3	585	973	1880	−219.72	91.58	87.82	28.5	−4.51	−2.07	−43.29	−13.74
		Panicoideae	119		57	57	50	−Inf	Inf	12.93	−71.42	−1.94	−1.54	−0.44	−3.05
	Sorghum tribe	Andropogoneae	106		15	8	10	−Inf	Inf	−66.04	−20.29	−1.94	−1.75	−0.87	−0.29
		Paniceae	11		12	14	9	−Inf	Inf	31.75	−133.9			−0.48	−1.89
Ferns	Ferns	Polypodiopsida	389		30	54	83	−Inf	Inf	103.27	2.22	−2.33	−1.58	−3.72	−0.12
	Leptosporangiate ferns	Polypodiidae	379		27	52	77	−Inf	Inf	117.49	−1.55	−2.31	−1.61	−4.05	−0.12
	Polypod ferns	Polypodiales	357		25	48	72	−Inf	Inf	116.82	−0.24	−2.26	−1.62	−3.77	−0.1
	Spleenwort suborder	Aspleniineae	234		18	32	49	−Inf	Inf	100.76	1.83	−2.03	−1.7	−2.42	−0.11
		Thelypteridaceae	42		4	8	11	−Inf	Inf	125.85	−9.36	−2.13			−0.17
		Thelypteridoideae	24		4	2	8	−Inf	Inf	−77.11	166.02	−2.31			
Mixed		Petrosaviidae	8919	5	1823	2912	3847	−583.02	258.2	80.38	−13.82	−24.31	−7.73	−135.44	−11.62
		commelinids	8530	3	1809	2904	3817	−988.71	492.42	81.28	−14.4	−26.94	−10.38	−136.77	−12.36

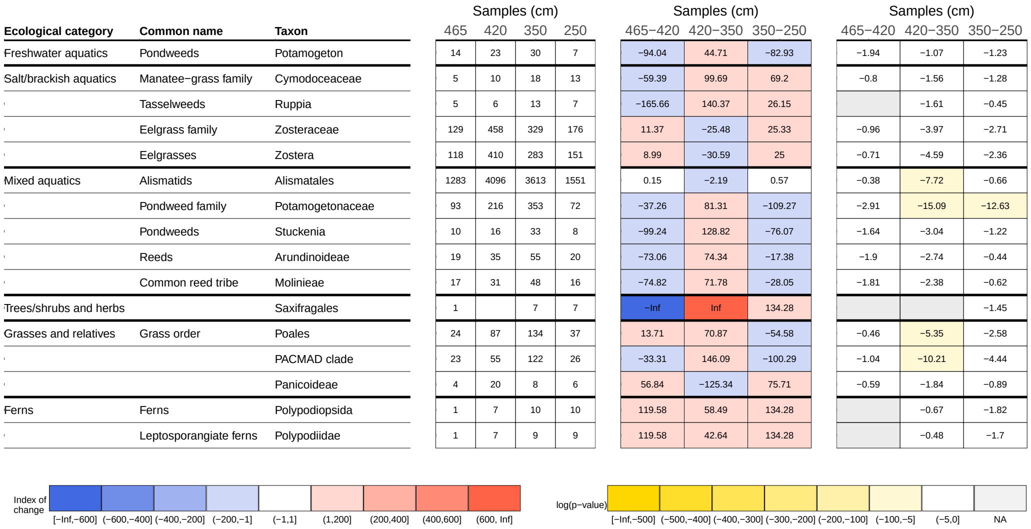
Index of change	[−Inf,−600]	(−600,−400]	(−400,−200]	(−200,−1]	(−1,1]	(1,200]	(200,400]	(400,600]	(600, Inf]

log(p−value)	[−Inf,−500]	(−500,−400]	(−400,−300]	(−300,−200]	(−200,−100]	(−100,−5]	(−5,0]	NA

Figure 5.14: Stratification analysis for ELF059A: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 100%.

ELF059A is similar to ELF059 but with generally more data, although still significant disparity in yield. Sample 320 in particular has far fewer taxa than its neighbours. All samples in this core are *Zostera*-dominated with small freshwater and woodland signals, and there is little evidence of ecological change. Nevertheless, there are significant changes to be found - 29.07% of the many comparisons with sufficient reads for a p-value - and enough to demonstrate complete stratification between all samples.

5.2.4 ELF060



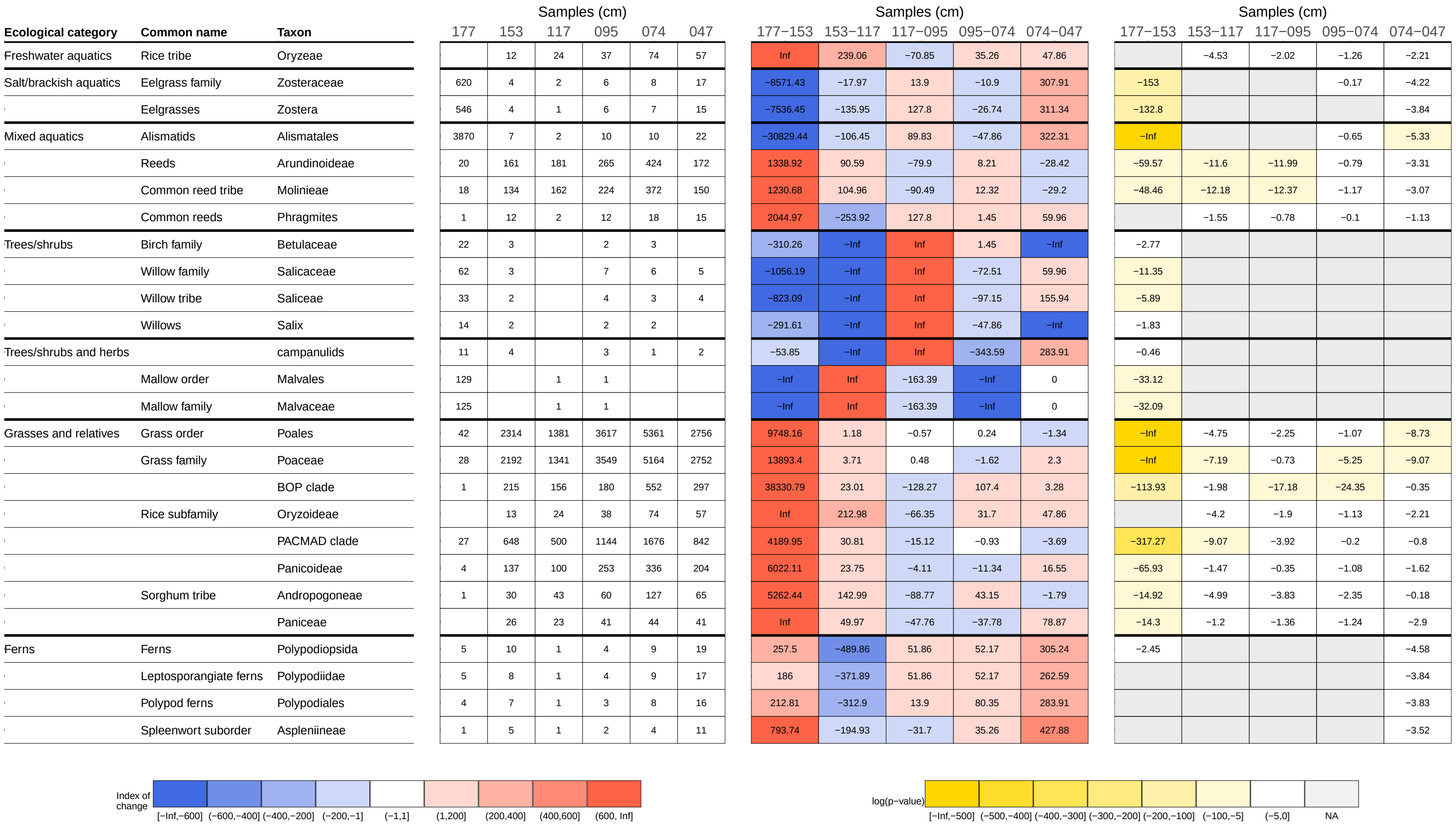
			Samples (cm)				Samples (cm)			Samples (cm)		
Ecological category	**Common name**	**Taxon**	465	420	350	250	465−420	420−350	350−250	465−420	420−350	350−250
Freshwater aquatics	Pondweeds	Potamogeton	14	23	30	7	−94.04	44.71	−82.93	−1.94	−1.07	−1.23
Salt/brackish aquatics	Manatee−grass family	Cymodoceaceae	5	10	18	13	−59.39	99.69	69.2	−0.8	−1.56	−1.28
	Tasselweeds	Ruppia	5	6	13	7	−165.66	140.37	26.15		−1.61	−0.45
	Eelgrass family	Zosteraceae	129	458	329	176	11.37	−25.48	25.33	−0.96	−3.97	−2.71
	Eelgrasses	Zostera	118	410	283	151	8.99	−30.59	25	−0.71	−4.59	−2.36
Mixed aquatics	Alismatids	Alismatales	1283	4096	3613	1551	0.15	−2.19	0.57	−0.38	−7.72	−0.66
	Pondweed family	Potamogetonaceae	93	216	353	72	−37.26	81.31	−109.27	−2.91	−15.09	−12.63
	Pondweeds	Stuckenia	10	16	33	8	−99.24	128.82	−76.07	−1.64	−3.04	−1.22
	Reeds	Arundinoideae	19	35	55	20	−73.06	74.34	−17.38	−1.9	−2.74	−0.44
	Common reed tribe	Molinieae	17	31	48	16	−74.82	71.78	−28.05	−1.81	−2.38	−0.62
Trees/shrubs and herbs		Saxifragales	1		7	7	−Inf	Inf	134.28			−1.45
Grasses and relatives	Grass order	Poales	24	87	134	37	13.71	70.87	−54.58	−0.46	−5.35	−2.58
		PACMAD clade	23	55	122	26	−33.31	146.09	−100.29	−1.04	−10.21	−4.44
		Panicoideae	4	20	8	6	56.84	−125.34	75.71	−0.59	−1.84	−0.89
Ferns	Ferns	Polypodiopsida	1	7	10	10	119.58	58.49	134.28		−0.67	−1.82
	Leptosporangiate ferns	Polypodiidae	1	7	9	9	119.58	42.64	134.28		−0.48	−1.7

Figure 5.15: Stratification analysis for ELF060: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 50%.

This is a particularly marine core, and particularly homogeneous in data yield and taxonomic profile. Accordingly, almost all indices of change are small and only 11.63% of p-values are significant, although these few can demonstrate stratification in 50% of instances, especially between samples 420 and 350. There is some evidence of stratification.

5.2.5 ELF032A

The deepest sample in ELF032A has a typical *Zostera*-dominated taxonomic profile, but all others resemble reed bed or marsh. Data yield is fairly consistent. In figure 5.16, the distinction between sample 177 and the others is clear. The 177-153 comparison contains many more significant p-values than later comparisons and most of the large indices of change. *Zostera*, trees/shrubs, and trees/shrubs and herbs are greatly reduced after sample 177. However, samples further up the core do show some small but significant differences, particularly among reeds and other grasses that appeared constant in the taxonomic profile. Only 38.96% of p-values are significant, but the stratification signal completeness is 80%. There is good evidence of stratification of sample 177 and a reasonable case for the rest of the core.

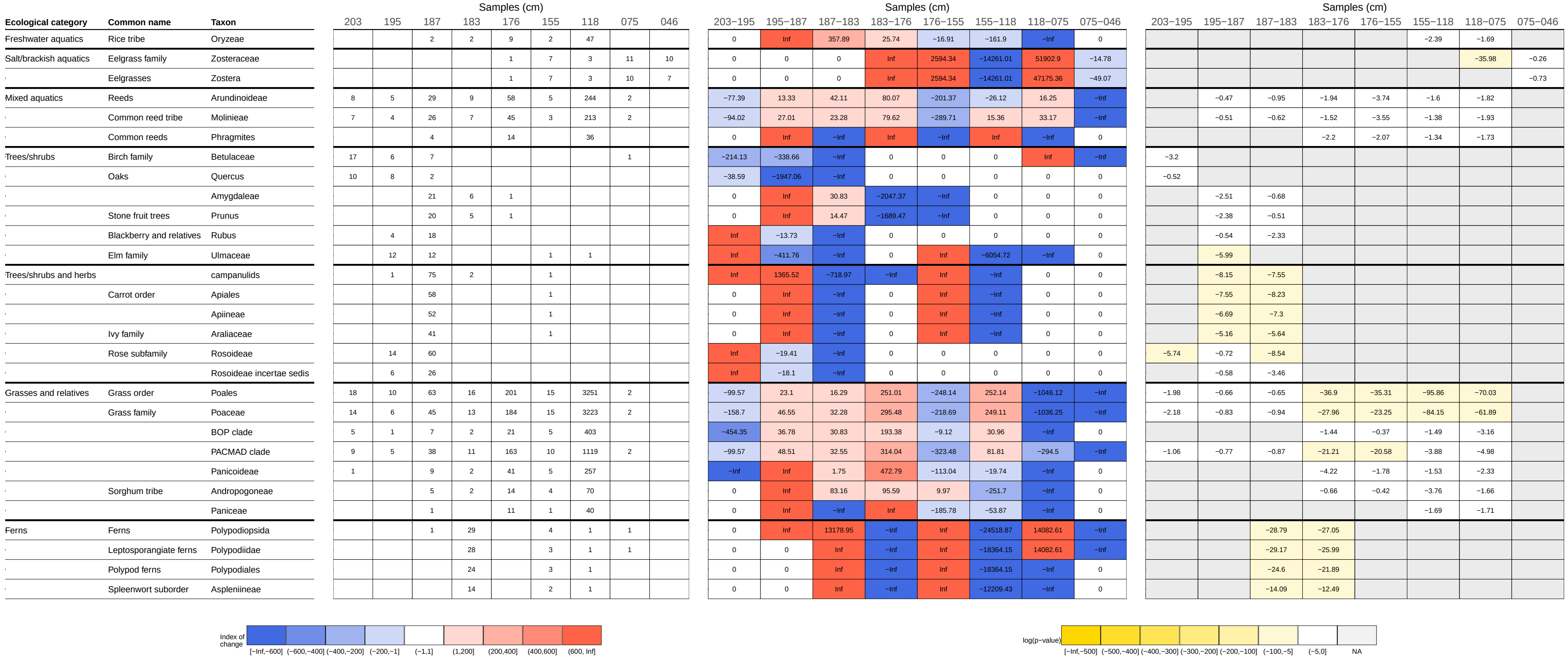


			Samples (cm)						Samples (cm)					Samples (cm)				
Ecological category	**Common name**	**Taxon**	177	153	117	095	074	047	177−153	153−117	117−095	095−074	074−047	177−153	153−117	117−095	095−074	074−047
Freshwater aquatics	Rice tribe	Oryzeae		12	24	37	74	57	Inf	239.06	−70.85	35.26	47.86		−4.53	−2.02	−1.26	−2.21
Salt/brackish aquatics	Eelgrass family	Zosteraceae	620	4	2	6	8	17	−8571.43	−17.97	13.9	−10.9	307.91	−153			−0.17	−4.22
	Eelgrasses	Zostera	546	4	1	6	7	15	−7536.45	−135.95	127.8	−26.74	311.34	−132.8				−3.84
Mixed aquatics	Alismatids	Alismatales	3870	7	2	10	10	22	−30829.44	−106.45	89.83	−47.86	322.31	−Inf			−0.65	−5.33
	Reeds	Arundinoideae	20	161	181	265	424	172	1338.92	90.59	−79.9	8.21	−28.42	−59.57	−11.6	−11.99	−0.79	−3.31
	Common reed tribe	Molinieae	18	134	162	224	372	150	1230.68	104.96	−90.49	12.32	−29.2	−48.46	−12.18	−12.37	−1.17	−3.07
	Common reeds	Phragmites	1	12	2	12	18	15	2044.97	−253.92	127.8	1.45	59.96		−1.55	−0.78	−0.1	−1.13
Trees/shrubs	Birch family	Betulaceae	22	3		2	3		−310.26	−Inf	Inf	1.45	−Inf	−2.77				
	Willow family	Salicaceae	62	3		7	6	5	−1056.19	−Inf	Inf	−72.51	59.96	−11.35				
	Willow tribe	Saliceae	33	2		4	3	4	−823.09	−Inf	Inf	−97.15	155.94	−5.89				
	Willows	Salix	14	2		2	2		−291.61	−Inf	Inf	−47.86	−Inf	−1.83				
Trees/shrubs and herbs		campanulids	11	4		3	1	2	−53.85	−Inf	Inf	−343.59	283.91	−0.46				
	Mallow order	Malvales	129		1	1			−Inf	Inf	−163.39	−Inf	0	−33.12				
	Mallow family	Malvaceae	125		1	1			−Inf	Inf	−163.39	−Inf	0	−32.09				
Grasses and relatives	Grass order	Poales	42	2314	1381	3617	5361	2756	9748.16	1.18	−0.57	0.24	−1.34	−Inf	−4.75	−2.25	−1.07	−8.73
	Grass family	Poaceae	28	2192	1341	3549	5164	2752	13893.4	3.71	0.48	−1.62	2.3	−Inf	−7.19	−0.73	−5.25	−9.07
		BOP clade	1	215	156	180	552	297	38330.79	23.01	−128.27	107.4	3.28	−113.93	−1.98	−17.18	−24.35	−0.35
	Rice subfamily	Oryzoideae		13	24	38	74	57	Inf	212.98	−66.35	31.7	47.86		−4.2	−1.9	−1.13	−2.21
		PACMAD clade	27	648	500	1144	1676	842	4189.95	30.81	−15.12	−0.93	−3.69	−317.27	−9.07	−3.92	−0.2	−0.8
		Panicoideae	4	137	100	253	336	204	6022.11	23.75	−4.11	−11.34	16.55	−65.93	−1.47	−0.35	−1.08	−1.62
	Sorghum tribe	Andropogoneae	1	30	43	60	127	65	5262.44	142.99	−88.77	43.15	−1.79	−14.92	−4.99	−3.83	−2.35	−0.18
		Paniceae		26	23	41	44	41	Inf	49.97	−47.76	−37.78	78.87	−14.3	−1.2	−1.36	−1.24	−2.9
Ferns	Ferns	Polypodiopsida	5	10	1	4	9	19	257.5	−489.86	51.86	52.17	305.24	−2.45				−4.58
	Leptosporangiate ferns	Polypodiidae	5	8	1	4	9	17	186	−371.89	51.86	52.17	262.59					−3.84
	Polypod ferns	Polypodiales	4	7	1	3	8	16	212.81	−312.9	13.9	80.35	283.91					−3.83
	Spleenwort suborder	Aspleniineae	1	5	1	2	4	11	793.74	−194.93	−31.7	35.26	427.88					−3.52

Figure 5.16: Stratification analysis for ELF032A: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 80%.

5.2.6 ELF033

The DNA profiles in ELF033 appear very different at first glance, but the lack of data complicates matters. The taxonomic profile shows transitions from woodland to grasses, which may represent reeds, in samples 183-176, and then from grasses to estuarine and back again in remaining samples. Accordingly, there are many dramatic indices of change. However, most comparisons lack sufficient reads for p-values, and of those that could be calculated, only 36.05% are significant. Nevertheless, the significant changes support stratification along most of the core, within and between the different leading ecological categories, with a reasonable signal completeness of 68.75%. Overall, the evidence for stratification is widespread but limited in strength.



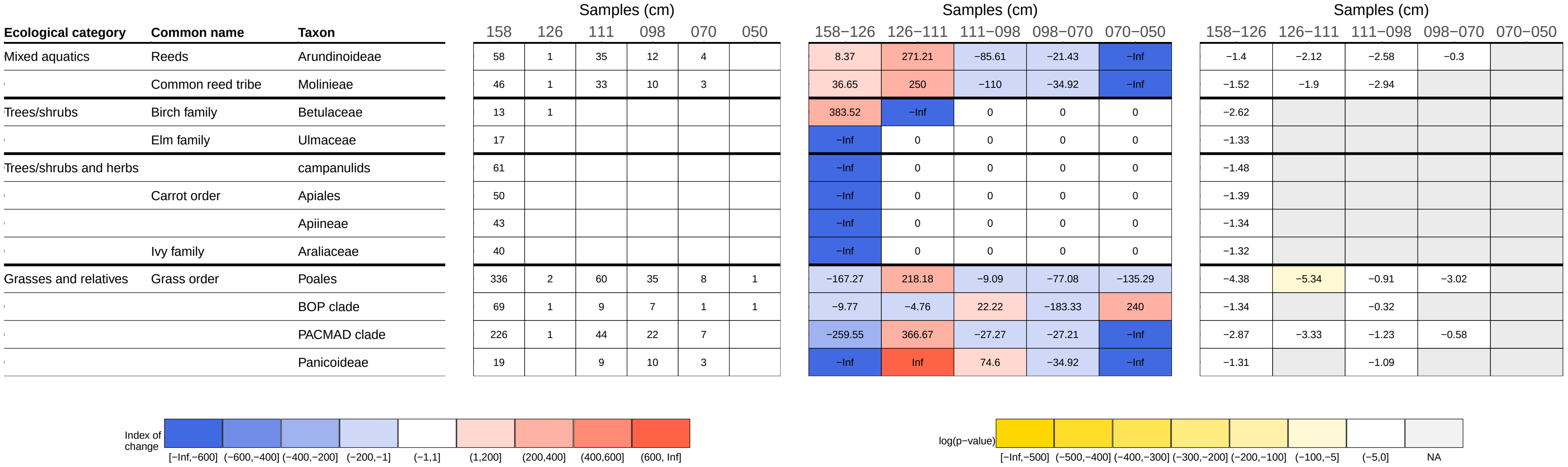
Ecological category	Common name	Taxon	203	195	187	183	176	155	118	075	046
			Samples (cm)								
Freshwater aquatics	Rice tribe	Oryzeae			2	2	9	2	47		
Salt/brackish aquatics	Eelgrass family	Zosteraceae					1	7	3	11	10
	Eelgrasses	Zostera					1	7	3	10	7
Mixed aquatics	Reeds	Arundinoideae	8	5	29	9	58	5	244	2	
	Common reed tribe	Molinieae	7	4	26	7	45	3	213	2	
	Common reeds	Phragmites			4		14		36		
Trees/shrubs	Birch family	Betulaceae	17	6	7					1	
	Oaks	Quercus	10	8	2						
		Amygdaleae			21	6	1				
	Stone fruit trees	Prunus			20	5	1				
	Blackberry and relatives	Rubus		4	18						
	Elm family	Ulmaceae		12	12			1	1		
Trees/shrubs and herbs		campanulids		1	75	2		1			
	Carrot order	Apiales			58			1			
		Apiineae			52			1			
	Ivy family	Araliaceae			41			1			
	Rose subfamily	Rosoideae		14	60						
		Rosoideae incertae sedis		6	26						
Grasses and relatives	Grass order	Poales	18	10	63	16	201	15	3251	2	
	Grass family	Poaceae	14	6	45	13	184	15	3223	2	
		BOP clade	5	1	7	2	21	5	403		
		PACMAD clade	9	5	38	11	163	10	1119	2	
		Panicoideae	1		9	2	41	5	257		
	Sorghum tribe	Andropogoneae			5	2	14	4	70		
		Paniceae			1		11	1	40		
Ferns	Ferns	Polypodiopsida			1	29		4	1	1	
	Leptosporangiate ferns	Polypodiidae				28		3	1	1	
	Polypod ferns	Polypodiales				24		3	1		
	Spleenwort suborder	Aspleniineae				14		2	1		

Taxon	203−195	195−187	187−183	183−176	176−155	155−118	118−075	075−046
	Samples (cm)							
Oryzeae	0	Inf	357.89	25.74	−16.91	−161.9	−Inf	0
Zosteraceae	0	0	0	Inf	2594.34	−14261.01	51902.9	−14.78
Zostera	0	0	0	Inf	2594.34	−14261.01	47175.36	−49.07
Arundinoideae	−77.39	13.33	42.11	80.07	−201.37	−26.12	16.25	−Inf
Molinieae	−94.02	27.01	23.28	79.62	−289.71	15.36	33.17	−Inf
Phragmites	0	Inf	−Inf	Inf	−Inf	Inf	−Inf	0
Betulaceae	−214.13	−338.66	−Inf	0	0	0	Inf	−Inf
Quercus	−38.59	−1947.06	−Inf	0	0	0	0	0
Amygdaleae	0	Inf	30.83	−2047.37	−Inf	0	0	0
Prunus	0	Inf	14.47	−1689.47	−Inf	0	0	0
Rubus	Inf	−13.73	−Inf	0	0	0	0	0
Ulmaceae	Inf	−411.76	−Inf	0	Inf	−6054.72	−Inf	0
campanulids	Inf	1365.52	−718.97	−Inf	Inf	−Inf	0	0
Apiales	0	Inf	−Inf	0	Inf	−Inf	0	0
Apiineae	0	Inf	−Inf	0	Inf	−Inf	0	0
Araliaceae	0	Inf	−Inf	0	Inf	−Inf	0	0
Rosoideae	Inf	−19.41	−Inf	0	0	0	0	0
Rosoideae incertae sedis	Inf	−18.1	−Inf	0	0	0	0	0
Poales	−99.57	23.1	16.29	251.01	−248.14	252.14	−1046.12	−Inf
Poaceae	−158.7	46.55	32.28	295.48	−218.69	249.11	−1036.25	−Inf
BOP clade	−454.35	36.78	30.83	193.38	−9.12	30.96	−Inf	0
PACMAD clade	−99.57	48.51	32.55	314.04	−323.48	81.81	−294.5	−Inf
Panicoideae	−Inf	Inf	1.75	472.79	−113.04	−19.74	−Inf	0
Andropogoneae	0	Inf	83.16	95.59	9.97	−251.7	−Inf	0
Paniceae	0	Inf	−Inf	Inf	−185.78	−53.87	−Inf	0
Polypodiopsida	0	Inf	13178.95	−Inf	Inf	−24518.87	14082.61	−Inf
Polypodiidae	0	0	Inf	−Inf	Inf	−18364.15	14082.61	−Inf
Polypodiales	0	0	Inf	−Inf	Inf	−18364.15	−Inf	0
Aspleniineae	0	0	Inf	−Inf	Inf	−12209.43	−Inf	0

Taxon	203−195	195−187	187−183	183−176	176−155	155−118	118−075	075−046
	Samples (cm)							
Oryzeae						−2.39	−1.69	
Zosteraceae							−35.98	−0.26
Zostera								−0.73
Arundinoideae		−0.47	−0.95	−1.94	−3.74	−1.6	−1.82	
Molinieae		−0.51	−0.62	−1.52	−3.55	−1.38	−1.93	
Phragmites				−2.2	−2.07	−1.34	−1.73	
Betulaceae	−3.2							
Quercus	−0.52							
Amygdaleae		−2.51	−0.68					
Prunus		−2.38	−0.51					
Rubus		−0.54	−2.33					
Ulmaceae		−5.99						
campanulids		−8.15	−7.55					
Apiales		−7.55	−8.23					
Apiineae		−6.69	−7.3					
Araliaceae		−5.16	−5.64					
Rosoideae	−5.74	−0.72	−8.54					
Rosoideae incertae sedis		−0.58	−3.46					
Poales	−1.98	−0.66	−0.65	−36.9	−35.31	−95.86	−70.03	
Poaceae	−2.18	−0.83	−0.94	−27.96	−23.25	−84.15	−61.89	
BOP clade				−1.44	−0.37	−1.49	−3.16	
PACMAD clade	−1.06	−0.77	−0.87	−21.21	−20.58	−3.88	−4.98	
Panicoideae				−4.22	−1.78	−1.53	−2.33	
Andropogoneae				−0.66	−0.42	−3.76	−1.66	
Paniceae						−1.69	−1.71	
Polypodiopsida			−28.79	−27.05				
Polypodiidae			−29.17	−25.99				
Polypodiales			−24.6	−21.89				
Aspleniineae			−14.09	−12.49				

Figure 5.17: Stratification analysis for ELF033: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 68.75%.

5.2.7 ELF033A



Ecological category	Common name	Taxon	158	126	111	098	070	050
Mixed aquatics	Reeds	Arundinoideae	58	1	35	12	4	
	Common reed tribe	Molinieae	46	1	33	10	3	
Trees/shrubs	Birch family	Betulaceae	13	1				
	Elm family	Ulmaceae	17					
Trees/shrubs and herbs		campanulids	61					
	Carrot order	Apiales	50					
		Apiineae	43					
	Ivy family	Araliaceae	40					
Grasses and relatives	Grass order	Poales	336	2	60	35	8	1
		BOP clade	69	1	9	7	1	1
		PACMAD clade	226	1	44	22	7	
		Panicoideae	19		9	10	3	

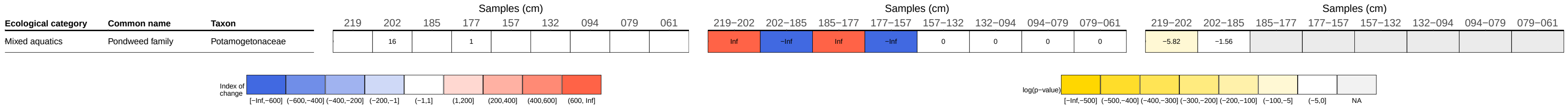
Ecological category	Common name	Taxon	158−126	126−111	111−098	098−070	070−050
Mixed aquatics	Reeds	Arundinoideae	8.37	271.21	−85.61	−21.43	−Inf
	Common reed tribe	Molinieae	36.65	250	−110	−34.92	−Inf
Trees/shrubs	Birch family	Betulaceae	383.52	−Inf	0	0	0
	Elm family	Ulmaceae	−Inf	0	0	0	0
Trees/shrubs and herbs		campanulids	−Inf	0	0	0	0
	Carrot order	Apiales	−Inf	0	0	0	0
		Apiineae	−Inf	0	0	0	0
	Ivy family	Araliaceae	−Inf	0	0	0	0
Grasses and relatives	Grass order	Poales	−167.27	218.18	−9.09	−77.08	−135.29
		BOP clade	−9.77	−4.76	22.22	−183.33	240
		PACMAD clade	−259.55	366.67	−27.27	−27.21	−Inf
		Panicoideae	−Inf	Inf	74.6	−34.92	−Inf

Ecological category	Common name	Taxon	158−126	126−111	111−098	098−070	070−050
Mixed aquatics	Reeds	Arundinoideae	−1.4	−2.12	−2.58	−0.3	
	Common reed tribe	Molinieae	−1.52	−1.9	−2.94		
Trees/shrubs	Birch family	Betulaceae	−2.62				
	Elm family	Ulmaceae	−1.33				
Trees/shrubs and herbs		campanulids	−1.48				
	Carrot order	Apiales	−1.39				
		Apiineae	−1.34				
	Ivy family	Araliaceae	−1.32				
Grasses and relatives	Grass order	Poales	−4.38	−5.34	−0.91	−3.02	
		BOP clade	−1.34		−0.32		
		PACMAD clade	−2.87	−3.33	−1.23	−0.58	
		Panicoideae	−1.31		−1.09		

Figure 5.18: Stratification analysis for ELF033A: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. ”N/A” indicates insufficient data. Stratification signal completeness: 10%.

ELF033A is similar to ELF033 in its variable taxonomic profile, but has even less data. Most comparisons only show p-values for a handful of taxa, and only one is significant (4%). There is not good evidence for stratification in this core.

5.2.8 ELF034



Samples (cm)

Ecological category	Common name	Taxon	219	202	185	177	157	132	094	079	061
Mixed aquatics	Pondweed family	Potamogetonaceae		16		1					

Samples (cm)

Ecological category	Common name	Taxon	219−202	202−185	185−177	177−157	157−132	132−094	094−079	079−061
Mixed aquatics	Pondweed family	Potamogetonaceae	Inf	−Inf	Inf	−Inf	0	0	0	0

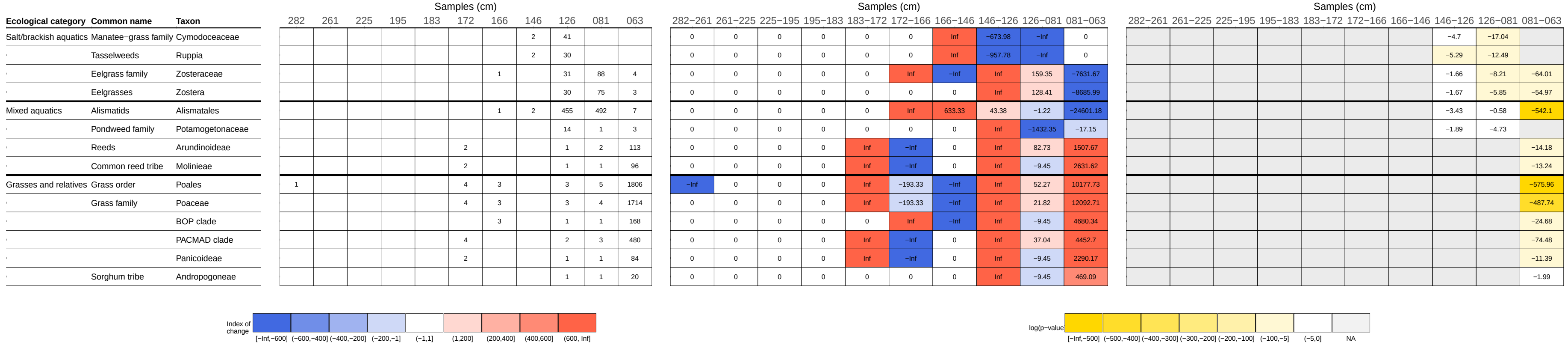
Samples (cm)

Ecological category	Common name	Taxon	219−202	202−185	185−177	177−157	157−132	132−094	094−079	079−061
Mixed aquatics	Pondweed family	Potamogetonaceae	−5.82	−1.56						

Figure 5.19: Stratification analysis for ELF034: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. ”N/A” indicates insufficient data. Stratification signal completeness: 6.25%.

The little data returned by ELF034 suggest a woody environment with some freshwater influence. There were only enough reads to compare a single taxon, Potamogetonaceae, across the deepest three samples. Only the deepest two showed a significant change.

5.2.9 ELF034A



Samples (cm)

Ecological category	Common name	Taxon	282	261	225	195	183	172	166	146	126	081	063
Salt/brackish aquatics	Manatee−grass family	Cymodoceaceae								2	41		
	Tasselweeds	Ruppia								2	30		
	Eelgrass family	Zosteraceae							1		31	88	4
	Eelgrasses	Zostera									30	75	3
Mixed aquatics	Alismatids	Alismatales							1	2	455	492	7
	Pondweed family	Potamogetonaceae									14	1	3
	Reeds	Arundinoideae						2			1	2	113
	Common reed tribe	Molinieae						2			1	1	96
Grasses and relatives	Grass order	Poales	1					4	3		3	5	1806
	Grass family	Poaceae						4	3		3	4	1714
		BOP clade							3		1	1	168
		PACMAD clade						4			2	3	480
		Panicoideae						2			1	1	84
	Sorghum tribe	Andropogoneae									1	1	20

Samples (cm)

Taxon	282−261	261−225	225−195	195−183	183−172	172−166	166−146	146−126	126−081	081−063
Cymodoceaceae	0	0	0	0	0	0	Inf	−673.98	−Inf	0
Ruppia	0	0	0	0	0	0	Inf	−957.78	−Inf	0
Zosteraceae	0	0	0	0	0	Inf	−Inf	Inf	159.35	−7631.67
Zostera	0	0	0	0	0	0	0	Inf	128.41	−8685.99
Alismatales	0	0	0	0	0	Inf	633.33	43.38	−1.22	−24601.18
Potamogetonaceae	0	0	0	0	0	0	0	Inf	−1432.35	−17.15
Arundinoideae	0	0	0	0	Inf	−Inf	0	Inf	82.73	1507.67
Molinieae	0	0	0	0	Inf	−Inf	0	Inf	−9.45	2631.62
Poales	−Inf	0	0	0	Inf	−193.33	−Inf	Inf	52.27	10177.73
Poaceae	0	0	0	0	Inf	−193.33	−Inf	Inf	21.82	12092.71
BOP clade	0	0	0	0	0	Inf	−Inf	Inf	−9.45	4680.34
PACMAD clade	0	0	0	0	Inf	−Inf	0	Inf	37.04	4452.7
Panicoideae	0	0	0	0	Inf	−Inf	0	Inf	−9.45	2290.17
Andropogoneae	0	0	0	0	0	0	0	Inf	−9.45	469.09

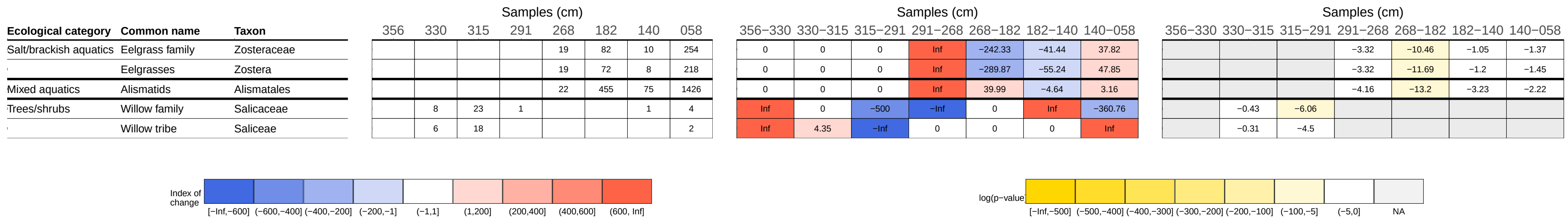
Samples (cm)

Taxon	282−261	261−225	225−195	195−183	183−172	172−166	166−146	146−126	126−081	081−063
Cymodoceaceae								−4.7	−17.04	
Ruppia								−5.29	−12.49	
Zosteraceae								−1.66	−8.21	−64.01
Zostera								−1.67	−5.85	−54.97
Alismatales								−3.43	−0.58	−542.1
Potamogetonaceae								−1.89	−4.73	
Arundinoideae										−14.18
Molinieae										−13.24
Poales										−575.96
Poaceae										−487.74
BOP clade										−24.68
PACMAD clade										−74.48
Panicoideae										−11.39
Andropogoneae										−1.99

Figure 5.20: Stratification analysis for ELF034A: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 25%.

The taxonomic profile of ELF034A has three distinct phases. The first eight samples resemble woodland, the next two are estuarine, and the top sample has mostly grasses or reeds. Data yield is very low for the woodland samples but reasonable for the top three. Accordingly, only comparisons involving the top three samples have enough data for p-values. Most are significant (73.91%), and they provide evidence against upward and downward movement between the top three samples and upward from the final woodland sample. The two estuarine samples therefore appear stratified despite their similar taxonomic profiles, a reminder that DNA profile and biogenomic mass are not directly comparable. However, the absence of data prevents assessment of stratification lower down the core. Overall stratification signal completeness is only 25%.

5.2.10 ELF054



Ecological category	Common name	Taxon	356	330	315	291	268	182	140	058
Salt/brackish aquatics	Eelgrass family	Zosteraceae					19	82	10	254
	Eelgrasses	Zostera					19	72	8	218
Mixed aquatics	Alismatids	Alismatales					22	455	75	1426
Trees/shrubs	Willow family	Salicaceae		8	23	1			1	4
	Willow tribe	Saliceae		6	18					2

Taxon	356−330	330−315	315−291	291−268	268−182	182−140	140−058
Zosteraceae	0	0	0	Inf	−242.33	−41.44	37.82
Zostera	0	0	0	Inf	−289.87	−55.24	47.85
Alismatales	0	0	0	Inf	39.99	−4.64	3.16
Salicaceae	Inf	0	−500	−Inf	0	Inf	−360.76
Saliceae	Inf	4.35	−Inf	0	0	0	Inf

Taxon	356−330	330−315	315−291	291−268	268−182	182−140	140−058
Zosteraceae				−3.32	−10.46	−1.05	−1.37
Zostera				−3.32	−11.69	−1.2	−1.45
Alismatales				−4.16	−13.2	−3.23	−2.22
Salicaceae		−0.43	−6.06				
Saliceae		−0.31	−4.5				

Figure 5.21: Stratification analysis for ELF054: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 7.14%.

The taxonomic profile for ELF054 again shows a woodland to estuarine transition, possibly with reeds in sample 291 (middle), although the low data yield prohibits p-values to test most of these changes. There are only four statistically significant comparisons, and three of these may be unduly influenced by read count: from sample 268 to 182, we see a decrease in *Zostera* and Zosteraceae and increase in Alismatales that can be explained by the rise in read count. Stratification signal completeness is only 7.14%. There is little evidence of stratification in ELF054.

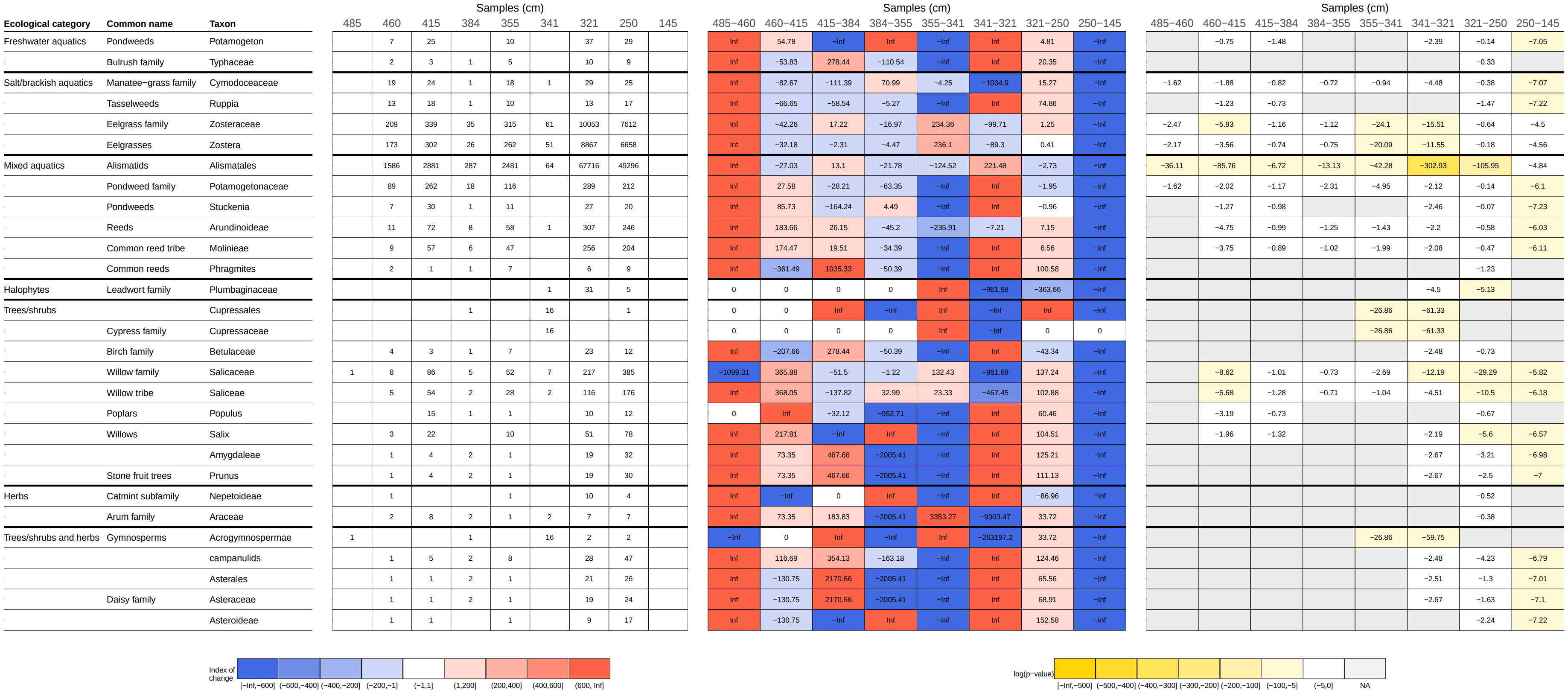
5.2.11 ELF039

Note that the stratification plot for ELF039 has been split across two pages (figures 5.22 and 5.23). The taxonomic profile is quite consistent, with all but the deepest sample containing mostly *Zostera* plus small woody and grass signals. However, the data yield is highly variable. The very top and bottom samples are essentially empty, explaining the many Inf/-Inf indices of change, while samples 321 and 250 have extraordinarily high read counts. Intermediate samples also vary. This creates a complicated signal.

Both figures show a mixture of indices of change and, as expected from a high-yield core with many rare taxa, most comparisons do not have sufficient data for p-values. The yield disparity also casts doubt on six statistically significant comparisons involving *Zostera*/Zosteraceae and Alismatales: there is a relative decrease in Alismatales from sample 355 to 341, associated with a decrease in overall read count, and a relative increase from 341 to 321 associated with a very large increase in read count. The Alismatales increase from 341 to 321 has a particularly low p-value. However, all could be explained by the change in data yield.

Nevertheless, there are many other significant changes across most comparisons. Signal completeness is high at 81.25%. There is good evidence for stratification in ELF039.

Part 2 (figure 5.23) has a similar pattern. The many rare taxa are not especially informative, and the significant p-values associated with commelinids (mixed) could be an artefact of commelinids being a higher taxon. However, overall, there are enough comparisons with sufficient data to show significant changes in DNA profile between most samples. Signal completeness is high at 81.25%. There is good evidence for stratification in ELF039.

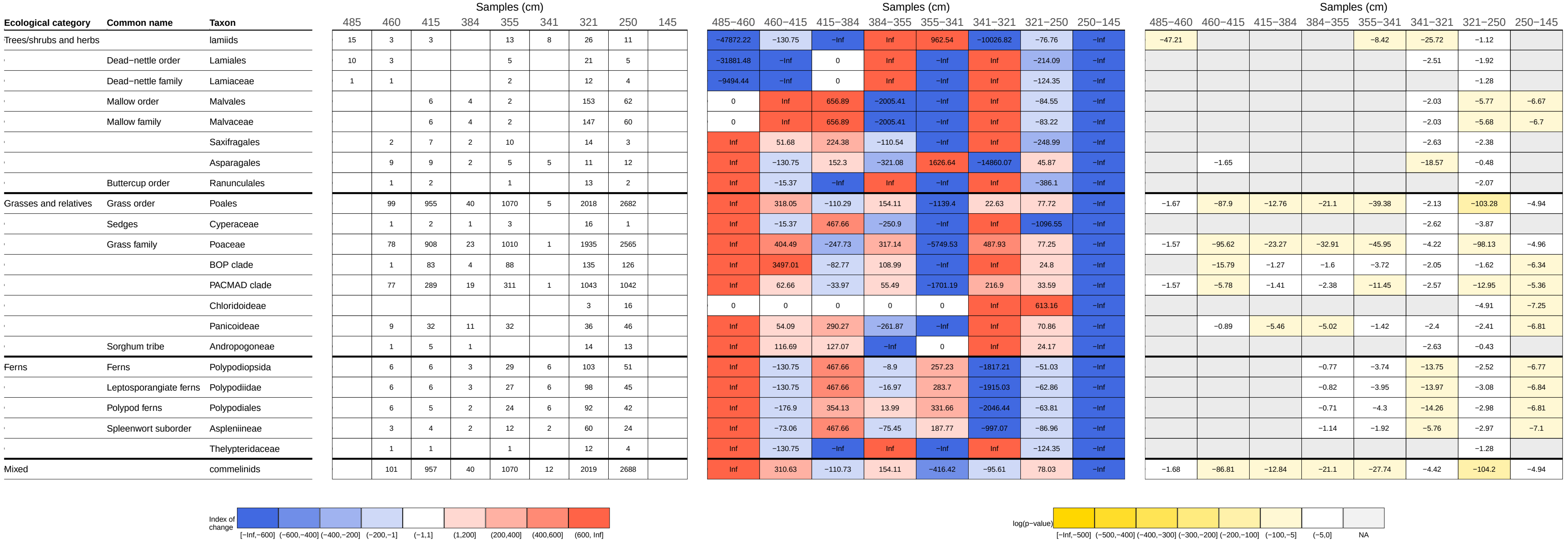


Ecological category	Common name	Taxon	Samples (cm)								
			485	460	415	384	355	341	321	250	145
Freshwater aquatics	Pondweeds	Potamogeton		7	25		10		37	29	
	Bulrush family	Typhaceae		2	3	1	5		10	9	
Salt/brackish aquatics	Manatee−grass family	Cymodoceaceae		19	24	1	18	1	29	25	
	Tasselweeds	Ruppia		13	18	1	10		13	17	
	Eelgrass family	Zosteraceae		209	339	35	315	61	10053	7612	
	Eelgrasses	Zostera		173	302	26	262	51	8867	6658	
Mixed aquatics	Alismatids	Alismatales		1586	2881	287	2481	64	67716	49296	
	Pondweed family	Potamogetonaceae		89	262	18	116		289	212	
	Pondweeds	Stuckenia		7	30	1	11		27	20	
	Reeds	Arundinoideae		11	72	8	58	1	307	246	
	Common reed tribe	Molinieae		9	57	6	47		256	204	
	Common reeds	Phragmites		2	1	1	7		6	9	
Halophytes	Leadwort family	Plumbaginaceae						1	31	5	
Trees/shrubs		Cupressales				1		16		1	
	Cypress family	Cupressaceae						16			
	Birch family	Betulaceae		4	3	1	7		23	12	
	Willow family	Salicaceae	1	8	86	5	52	7	217	385	
	Willow tribe	Saliceae		5	54	2	28	2	116	176	
	Poplars	Populus			15	1	1		10	12	
	Willows	Salix		3	22		10		51	78	
		Amygdaleae		1	4	2	1		19	32	
	Stone fruit trees	Prunus		1	4	2	1		19	30	
Herbs	Catmint subfamily	Nepetoideae		1			1		10	4	
	Arum family	Araceae		2	8	2	1	2	7	7	
Trees/shrubs and herbs	Gymnosperms	Acrogymnospermae	1			1		16	2	2	
		campanulids		1	5	2	8		28	47	
		Asterales		1	1	2	1		21	26	
	Daisy family	Asteraceae		1	1	2	1		19	24	
		Asteroideae		1	1		1		9	17	

Taxon	Samples (cm)							
	485−460	460−415	415−384	384−355	355−341	341−321	321−250	250−145
Potamogeton	Inf	54.78	−Inf	Inf	−Inf	Inf	4.81	−Inf
Typhaceae	Inf	−53.83	278.44	−110.54	−Inf	Inf	20.35	−Inf
Cymodoceaceae	Inf	−82.67	−111.39	70.99	−4.25	−1034.9	15.27	−Inf
Ruppia	Inf	−66.65	−58.54	−5.27	−Inf	Inf	74.86	−Inf
Zosteraceae	Inf	−42.26	17.22	−16.97	234.36	−99.71	1.25	−Inf
Zostera	Inf	−32.18	−2.31	−4.47	236.1	−89.3	0.41	−Inf
Alismatales	Inf	−27.03	13.1	−21.78	−124.52	221.48	−2.73	−Inf
Potamogetonaceae	Inf	27.58	−28.21	−63.35	−Inf	Inf	−1.95	−Inf
Stuckenia	Inf	85.73	−164.24	4.49	−Inf	Inf	−0.96	−Inf
Arundinoideae	Inf	183.66	26.15	−45.2	−235.91	−7.21	7.15	−Inf
Molinieae	Inf	174.47	19.51	−34.39	−Inf	Inf	6.56	−Inf
Phragmites	Inf	−361.49	1035.33	−50.39	−Inf	Inf	100.58	−Inf
Plumbaginaceae	0	0	0	0	Inf	−961.68	−363.66	−Inf
Cupressales	0	0	Inf	−Inf	Inf	−Inf	Inf	−Inf
Cupressaceae	0	0	0	0	Inf	−Inf	0	0
Betulaceae	Inf	−207.66	278.44	−50.39	−Inf	Inf	−43.34	−Inf
Salicaceae	−1099.31	365.88	−51.5	−1.22	132.43	−961.68	137.24	−Inf
Saliceae	Inf	368.05	−137.82	32.99	23.33	−467.45	102.88	−Inf
Populus	0	Inf	−32.12	−952.71	−Inf	Inf	60.46	−Inf
Salix	Inf	217.81	−Inf	Inf	−Inf	Inf	104.51	−Inf
Amygdaleae	Inf	73.35	467.66	−2005.41	−Inf	Inf	125.21	−Inf
Prunus	Inf	73.35	467.66	−2005.41	−Inf	Inf	111.13	−Inf
Nepetoideae	Inf	−Inf	0	Inf	−Inf	Inf	−86.96	−Inf
Araceae	Inf	73.35	183.83	−2005.41	3353.27	−9303.47	33.72	−Inf
Acrogymnospermae	−Inf	0	Inf	−Inf	Inf	−263197.2	33.72	−Inf
campanulids	Inf	116.69	354.13	−163.18	−Inf	Inf	124.46	−Inf
Asterales	Inf	−130.75	2170.66	−2005.41	−Inf	Inf	65.56	−Inf
Asteraceae	Inf	−130.75	2170.66	−2005.41	−Inf	Inf	68.91	−Inf
Asteroideae	Inf	−130.75	−Inf	Inf	−Inf	Inf	152.58	−Inf

Taxon	Samples (cm)							
	485−460	460−415	415−384	384−355	355−341	341−321	321−250	250−145
Potamogeton		−0.75	−1.48			−2.39	−0.14	−7.05
Typhaceae							−0.33	
Cymodoceaceae	−1.62	−1.88	−0.82	−0.72	−0.94	−4.48	−0.38	−7.07
Ruppia		−1.23	−0.73				−1.47	−7.22
Zosteraceae	−2.47	−5.93	−1.16	−1.12	−24.1	−15.51	−0.64	−4.5
Zostera	−2.17	−3.56	−0.74	−0.75	−20.09	−11.55	−0.18	−4.56
Alismatales	−36.11	−85.76	−6.72	−13.13	−42.28	−302.93	−105.95	−4.84
Potamogetonaceae	−1.62	−2.02	−1.17	−2.31	−4.95	−2.12	−0.14	−6.1
Stuckenia		−1.27	−0.98			−2.46	−0.07	−7.23
Arundinoideae		−4.75	−0.99	−1.25	−1.43	−2.2	−0.58	−6.03
Molinieae		−3.75	−0.89	−1.02	−1.99	−2.08	−0.47	−6.11
Phragmites							−1.23	
Plumbaginaceae						−4.5	−5.13	
Cupressales					−26.86	−61.33		
Cupressaceae					−26.86	−61.33		
Betulaceae						−2.48	−0.73	
Salicaceae		−8.62	−1.01	−0.73	−2.69	−12.19	−29.29	−5.82
Saliceae		−5.68	−1.28	−0.71	−1.04	−4.51	−10.5	−6.18
Populus		−3.19	−0.73				−0.67	
Salix		−1.96	−1.32			−2.19	−5.6	−6.57
Amygdaleae						−2.67	−3.21	−6.98
Prunus						−2.67	−2.5	−7
Nepetoideae							−0.52	
Araceae							−0.38	
Acrogymnospermae					−26.86	−59.75		
campanulids						−2.48	−4.23	−6.79
Asterales						−2.51	−1.3	−7.01
Asteraceae						−2.67	−1.63	−7.1
Asteroideae							−2.24	−7.22

Figure 5.22: Stratification analysis for ELF039: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 81.25%. Part 1 of 2.



Ecological category	Common name	Taxon	485	460	415	384	355	341	321	250	145
Trees/shrubs and herbs		lamiids	15	3	3		13	8	26	11	
	Dead−nettle order	Lamiales	10	3			5		21	5	
	Dead−nettle family	Lamiaceae	1	1			2		12	4	
	Mallow order	Malvales			6	4	2		153	62	
	Mallow family	Malvaceae			6	4	2		147	60	
		Saxifragales		2	7	2	10		14	3	
		Asparagales		9	9	2	5	5	11	12	
	Buttercup order	Ranunculales		1	2		1		13	2	
Grasses and relatives	Grass order	Poales		99	955	40	1070	5	2018	2682	
	Sedges	Cyperaceae		1	2	1	3		16	1	
	Grass family	Poaceae		78	908	23	1010	1	1935	2565	
		BOP clade		1	83	4	88		135	126	
		PACMAD clade		77	289	19	311	1	1043	1042	
		Chloridoideae							3	16	
		Panicoideae		9	32	11	32		36	46	
	Sorghum tribe	Andropogoneae		1	5	1			14	13	
Ferns	Ferns	Polypodiopsida		6	6	3	29	6	103	51	
	Leptosporangiate ferns	Polypodiidae		6	6	3	27	6	98	45	
	Polypod ferns	Polypodiales		6	5	2	24	6	92	42	
	Spleenwort suborder	Aspleniineae		3	4	2	12	2	60	24	
		Thelypteridaceae		1	1		1		12	4	
Mixed		commelinids		101	957	40	1070	12	2019	2688	

Samples (cm)

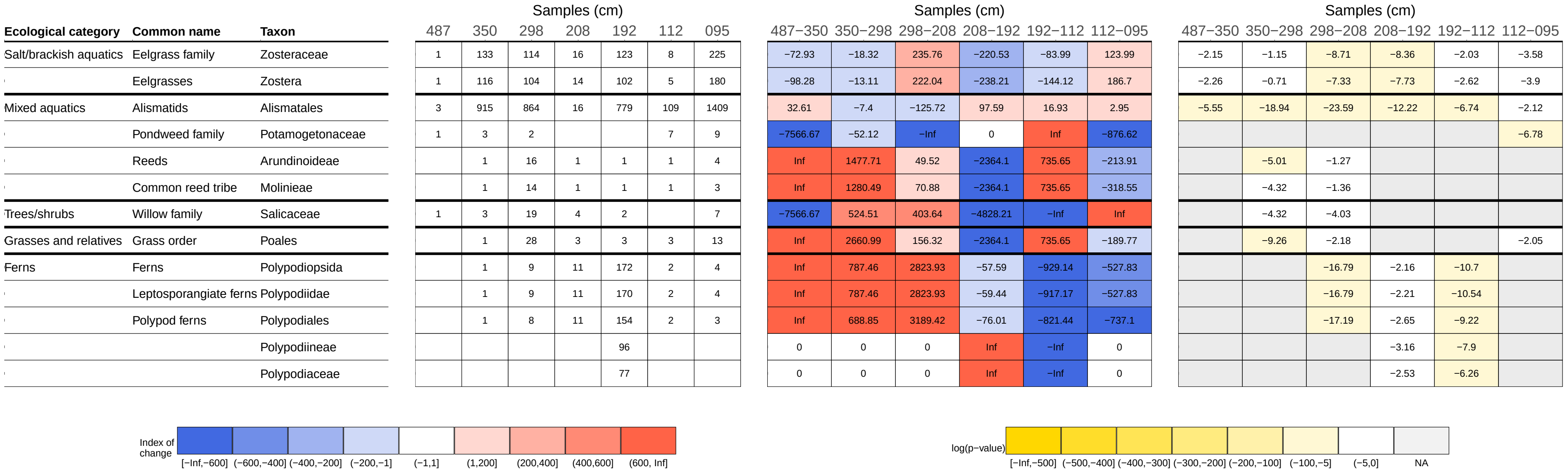
Taxon	485−460	460−415	415−384	384−355	355−341	341−321	321−250	250−145
lamiids	−47872.22	−130.75	−Inf	Inf	962.54	−10026.82	−76.76	−Inf
Lamiales	−31881.48	−Inf	0	Inf	−Inf	Inf	−214.09	−Inf
Lamiaceae	−9494.44	−Inf	0	Inf	−Inf	Inf	−124.35	−Inf
Malvales	0	Inf	656.89	−2005.41	−Inf	Inf	−84.55	−Inf
Malvaceae	0	Inf	656.89	−2005.41	−Inf	Inf	−83.22	−Inf
Saxifragales	Inf	51.68	224.38	−110.54	−Inf	Inf	−248.99	−Inf
Asparagales	Inf	−130.75	152.3	−321.08	1626.64	−14860.07	45.87	−Inf
Ranunculales	Inf	−15.37	−Inf	Inf	−Inf	Inf	−386.1	−Inf
Poales	Inf	318.05	−110.29	154.11	−1139.4	22.63	77.72	−Inf
Cyperaceae	Inf	−15.37	467.66	−250.9	−Inf	Inf	−1096.55	−Inf
Poaceae	Inf	404.49	−247.73	317.14	−5749.53	487.93	77.25	−Inf
BOP clade	Inf	3497.01	−82.77	108.99	−Inf	Inf	24.8	−Inf
PACMAD clade	Inf	62.66	−33.97	55.49	−1701.19	216.9	33.59	−Inf
Chloridoideae	0	0	0	0	0	Inf	613.16	−Inf
Panicoideae	Inf	54.09	290.27	−261.87	−Inf	Inf	70.86	−Inf
Andropogoneae	Inf	116.69	127.07	−Inf	0	Inf	24.17	−Inf
Polypodiopsida	Inf	−130.75	467.66	−8.9	257.23	−1817.21	−51.03	−Inf
Polypodiidae	Inf	−130.75	467.66	−16.97	283.7	−1915.03	−62.86	−Inf
Polypodiales	Inf	−176.9	354.13	13.99	331.66	−2046.44	−63.81	−Inf
Aspleniineae	Inf	−73.06	467.66	−75.45	187.77	−997.07	−86.96	−Inf
Thelypteridaceae	Inf	−130.75	−Inf	Inf	−Inf	Inf	−124.35	−Inf
commelinids	Inf	310.63	−110.73	154.11	−416.42	−95.61	78.03	−Inf

Samples (cm)

Taxon	485−460	460−415	415−384	384−355	355−341	341−321	321−250	250−145
lamiids	−47.21				−8.42	−25.72	−1.12	
Lamiales						−2.51	−1.92	
Lamiaceae							−1.28	
Malvales						−2.03	−5.77	−6.67
Malvaceae						−2.03	−5.68	−6.7
Saxifragales						−2.63	−2.38	
Asparagales		−1.65				−18.57	−0.48	
Ranunculales							−2.07	
Poales	−1.67	−87.9	−12.76	−21.1	−39.38	−2.13	−103.28	−4.94
Cyperaceae						−2.62	−3.87	
Poaceae	−1.57	−95.62	−23.27	−32.91	−45.95	−4.22	−98.13	−4.96
BOP clade		−15.79	−1.27	−1.6	−3.72	−2.05	−1.62	−6.34
PACMAD clade	−1.57	−5.78	−1.41	−2.38	−11.45	−2.57	−12.95	−5.36
Chloridoideae							−4.91	−7.25
Panicoideae		−0.89	−5.46	−5.02	−1.42	−2.4	−2.41	−6.81
Andropogoneae						−2.63	−0.43	
Polypodiopsida				−0.77	−3.74	−13.75	−2.52	−6.77
Polypodiidae				−0.82	−3.95	−13.97	−3.08	−6.84
Polypodiales				−0.71	−4.3	−14.26	−2.98	−6.81
Aspleniineae				−1.14	−1.92	−5.76	−2.97	−7.1
Thelypteridaceae							−1.28	
commelinids	−1.68	−86.81	−12.84	−21.1	−27.74	−4.42	−104.2	−4.94

Figure 5.23: Stratification analysis for ELF039. Part 2 of 2.

5.2.12 ELF040A

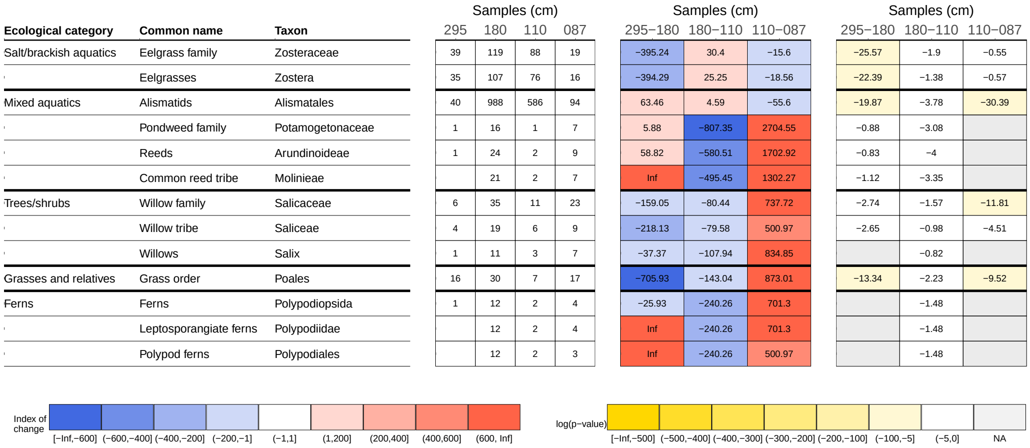


			Samples (cm)							Samples (cm)						Samples (cm)					
Ecological category	**Common name**	**Taxon**	487	350	298	208	192	112	095	487−350	350−298	298−208	208−192	192−112	112−095	487−350	350−298	298−208	208−192	192−112	112−095
Salt/brackish aquatics	Eelgrass family	Zosteraceae	1	133	114	16	123	8	225	−72.93	−18.32	235.76	−220.53	−83.99	123.99	−2.15	−1.15	−8.71	−8.36	−2.03	−3.58
	Eelgrasses	Zostera	1	116	104	14	102	5	180	−98.28	−13.11	222.04	−238.21	−144.12	186.7	−2.26	−0.71	−7.33	−7.73	−2.62	−3.9
Mixed aquatics	Alismatids	Alismatales	3	915	864	16	779	109	1409	32.61	−7.4	−125.72	97.59	16.93	2.95	−5.55	−18.94	−23.59	−12.22	−6.74	−2.12
	Pondweed family	Potamogetonaceae	1	3	2			7	9	−7566.67	−52.12	−Inf	0	Inf	−876.62						−6.78
	Reeds	Arundinoideae		1	16	1	1	1	4	Inf	1477.71	49.52	−2364.1	735.65	−213.91		−5.01	−1.27			
	Common reed tribe	Molinieae		1	14	1	1	1	3	Inf	1280.49	70.88	−2364.1	735.65	−318.55		−4.32	−1.36			
Trees/shrubs	Willow family	Salicaceae	1	3	19	4	2		7	−7566.67	524.51	403.64	−4828.21	−Inf	Inf		−4.32	−4.03			
Grasses and relatives	Grass order	Poales		1	28	3	3	3	13	Inf	2660.99	156.32	−2364.1	735.65	−189.77		−9.26	−2.18			−2.05
Ferns	Ferns	Polypodiopsida		1	9	11	172	2	4	Inf	787.46	2823.93	−57.59	−929.14	−527.83			−16.79	−2.16	−10.7	
	Leptosporangiate ferns	Polypodiidae		1	9	11	170	2	4	Inf	787.46	2823.93	−59.44	−917.17	−527.83			−16.79	−2.21	−10.54	
	Polypod ferns	Polypodiales		1	8	11	154	2	3	Inf	688.85	3189.42	−76.01	−821.44	−737.1			−17.19	−2.65	−9.22	
		Polypodiineae					96			0	0	0	Inf	−Inf	0				−3.16	−7.9	
		Polypodiaceae					77			0	0	0	Inf	−Inf	0				−2.53	−6.26	

Figure 5.24: Stratification analysis for ELF040A: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 50%.

The data yield in ELF040A is erratic, with adjacent samples usually showing quite different read counts. The taxonomic profile in ELF040A is consistent, with all samples appearing as *Zostera*-dominated salt/brackish apart from the mixed aquatic 112. However, the data yield is erratic, resulting in disqualification of the significant changes in *Zostera*, Zosteraceae and Alismatales for most samples. Other taxa generally have too few reads for p-values, but we can see significant appearances of grasses and reeds in sample 298 and ferns from 208 to 192. The overall stratification signal completeness is 50%. There is some evidence for stratification in ELF040A.

5.2.13 ELF041



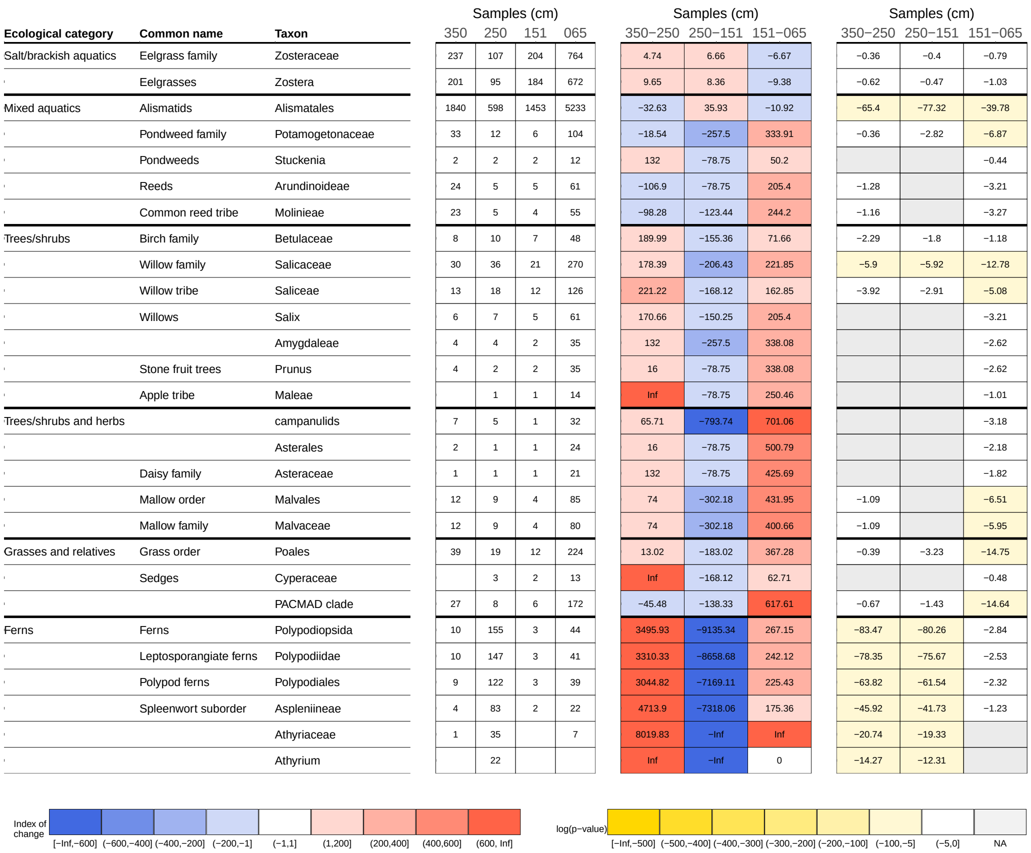
			Samples (cm)				Samples (cm)			Samples (cm)		
Ecological category	**Common name**	**Taxon**	295	180	110	087	295−180	180−110	110−087	295−180	180−110	110−087
Salt/brackish aquatics	Eelgrass family	Zosteraceae	39	119	88	19	−395.24	30.4	−15.6	−25.57	−1.9	−0.55
	Eelgrasses	Zostera	35	107	76	16	−394.29	25.25	−18.56	−22.39	−1.38	−0.57
Mixed aquatics	Alismatids	Alismatales	40	988	586	94	63.46	4.59	−55.6	−19.87	−3.78	−30.39
	Pondweed family	Potamogetonaceae	1	16	1	7	5.88	−807.35	2704.55	−0.88	−3.08	
	Reeds	Arundinoideae	1	24	2	9	58.82	−580.51	1702.92	−0.83	−4	
	Common reed tribe	Molinieae		21	2	7	Inf	−495.45	1302.27	−1.12	−3.35	
Trees/shrubs	Willow family	Salicaceae	6	35	11	23	−159.05	−80.44	737.72	−2.74	−1.57	−11.81
	Willow tribe	Saliceae	4	19	6	9	−218.13	−79.58	500.97	−2.65	−0.98	−4.51
	Willows	Salix	1	11	3	7	−37.37	−107.94	834.85		−0.82	
Grasses and relatives	Grass order	Poales	16	30	7	17	−705.93	−143.04	873.01	−13.34	−2.23	−9.52
Ferns	Ferns	Polypodiopsida	1	12	2	4	−25.93	−240.26	701.3		−1.48	
	Leptosporangiate ferns	Polypodiidae		12	2	4	Inf	−240.26	701.3		−1.48	
	Polypod ferns	Polypodiales		12	2	3	Inf	−240.26	500.97		−1.48	

Figure 5.25: Stratification analysis for ELF041: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 50%.

The ecological profiles in ELF041 show some change, from mostly *Zostera* with underlying grass and trees/shrubs in the deeper three samples to an increase in trees/shrubs at the top (87). This shift is clearly visible in the indices of change, as are simultaneous increases in freshwater aquatics and ferns. Unfortunately, the generally low data yield means that few comparisons have p-values, and three of the four significant changes between the first two samples can be attributed to the effect of data yield on *Zostera* and its parent taxa. Overall, evidence for stratification lower down the core is limited, although the DNA profile of the top sample does appear distinct.

5.2.14 ELF042

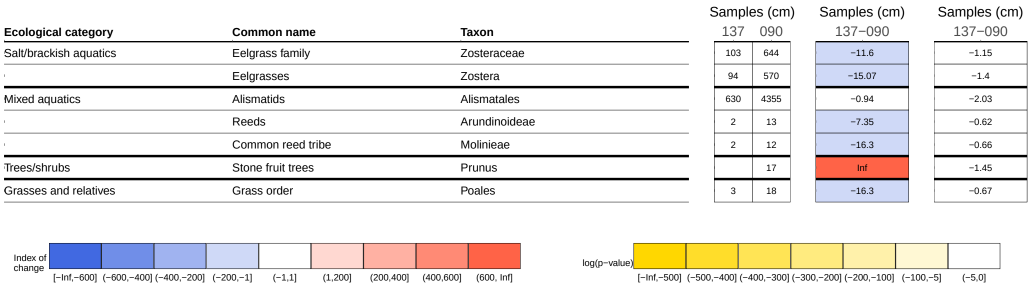
All four samples share high data yields and similar brackish taxonomic profiles with some woody taxa. While the decrease in Alismatales from 530-250 could be attributed to the increase in read count (due to the corresponding increase in *Zostera*/Zosteraceae), there are many other significant changes across all three comparisons. Willows consistently vary, there is a burst of ferns in sample 250, and the top sample is distinguished by appearances of Malvaceae and grasses. Stratification signal completeness for ELF042 is a respectable 83.33%.



			Samples (cm)				Samples (cm)			Samples (cm)		
Ecological category	**Common name**	**Taxon**	350	250	151	065	350−250	250−151	151−065	350−250	250−151	151−065
Salt/brackish aquatics	Eelgrass family	Zosteraceae	237	107	204	764	4.74	6.66	−6.67	−0.36	−0.4	−0.79
	Eelgrasses	Zostera	201	95	184	672	9.65	8.36	−9.38	−0.62	−0.47	−1.03
Mixed aquatics	Alismatids	Alismatales	1840	598	1453	5233	−32.63	35.93	−10.92	−65.4	−77.32	−39.78
	Pondweed family	Potamogetonaceae	33	12	6	104	−18.54	−257.5	333.91	−0.36	−2.82	−6.87
	Pondweeds	Stuckenia	2	2	2	12	132	−78.75	50.2			−0.44
	Reeds	Arundinoideae	24	5	5	61	−106.9	−78.75	205.4	−1.28		−3.21
	Common reed tribe	Molinieae	23	5	4	55	−98.28	−123.44	244.2	−1.16		−3.27
Trees/shrubs	Birch family	Betulaceae	8	10	7	48	189.99	−155.36	71.66	−2.29	−1.8	−1.18
	Willow family	Salicaceae	30	36	21	270	178.39	−206.43	221.85	−5.9	−5.92	−12.78
	Willow tribe	Saliceae	13	18	12	126	221.22	−168.12	162.85	−3.92	−2.91	−5.08
	Willows	Salix	6	7	5	61	170.66	−150.25	205.4			−3.21
		Amygdaleae	4	4	2	35	132	−257.5	338.08			−2.62
	Stone fruit trees	Prunus	4	2	2	35	16	−78.75	338.08			−2.62
	Apple tribe	Maleae		1	1	14	Inf	−78.75	250.46			−1.01
Trees/shrubs and herbs		campanulids	7	5	1	32	65.71	−793.74	701.06			−3.18
		Asterales	2	1	1	24	16	−78.75	500.79			−2.18
	Daisy family	Asteraceae	1	1	1	21	132	−78.75	425.69			−1.82
	Mallow order	Malvales	12	9	4	85	74	−302.18	431.95	−1.09		−6.51
	Mallow family	Malvaceae	12	9	4	80	74	−302.18	400.66	−1.09		−5.95
Grasses and relatives	Grass order	Poales	39	19	12	224	13.02	−183.02	367.28	−0.39	−3.23	−14.75
	Sedges	Cyperaceae		3	2	13	Inf	−168.12	62.71			−0.48
		PACMAD clade	27	8	6	172	−45.48	−138.33	617.61	−0.67	−1.43	−14.64
Ferns	Ferns	Polypodiopsida	10	155	3	44	3495.93	−9135.34	267.15	−83.47	−80.26	−2.84
	Leptosporangiate ferns	Polypodiidae	10	147	3	41	3310.33	−8658.68	242.12	−78.35	−75.67	−2.53
	Polypod ferns	Polypodiales	9	122	3	39	3044.82	−7169.11	225.43	−63.82	−61.54	−2.32
	Spleenwort suborder	Aspleniineae	4	83	2	22	4713.9	−7318.06	175.36	−45.92	−41.73	−1.23
		Athyriaceae	1	35		7	8019.83	−Inf	Inf	−20.74	−19.33	
		Athyrium		22			Inf	−Inf	0	−14.27	−12.31	

Figure 5.26: Stratification analysis for ELF042: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 83.33%.

5.2.15 ELF044



			Samples (cm)		Samples (cm)	Samples (cm)
Ecological category	**Common name**	**Taxon**	137	090	137−090	137−090
Salt/brackish aquatics	Eelgrass family	Zosteraceae	103	644	−11.6	−1.15
	Eelgrasses	Zostera	94	570	−15.07	−1.4
Mixed aquatics	Alismatids	Alismatales	630	4355	−0.94	−2.03
	Reeds	Arundinoideae	2	13	−7.35	−0.62
	Common reed tribe	Molinieae	2	12	−16.3	−0.66
Trees/shrubs	Stone fruit trees	Prunus		17	Inf	−1.45
Grasses and relatives	Grass order	Poales	3	18	−16.3	−0.67

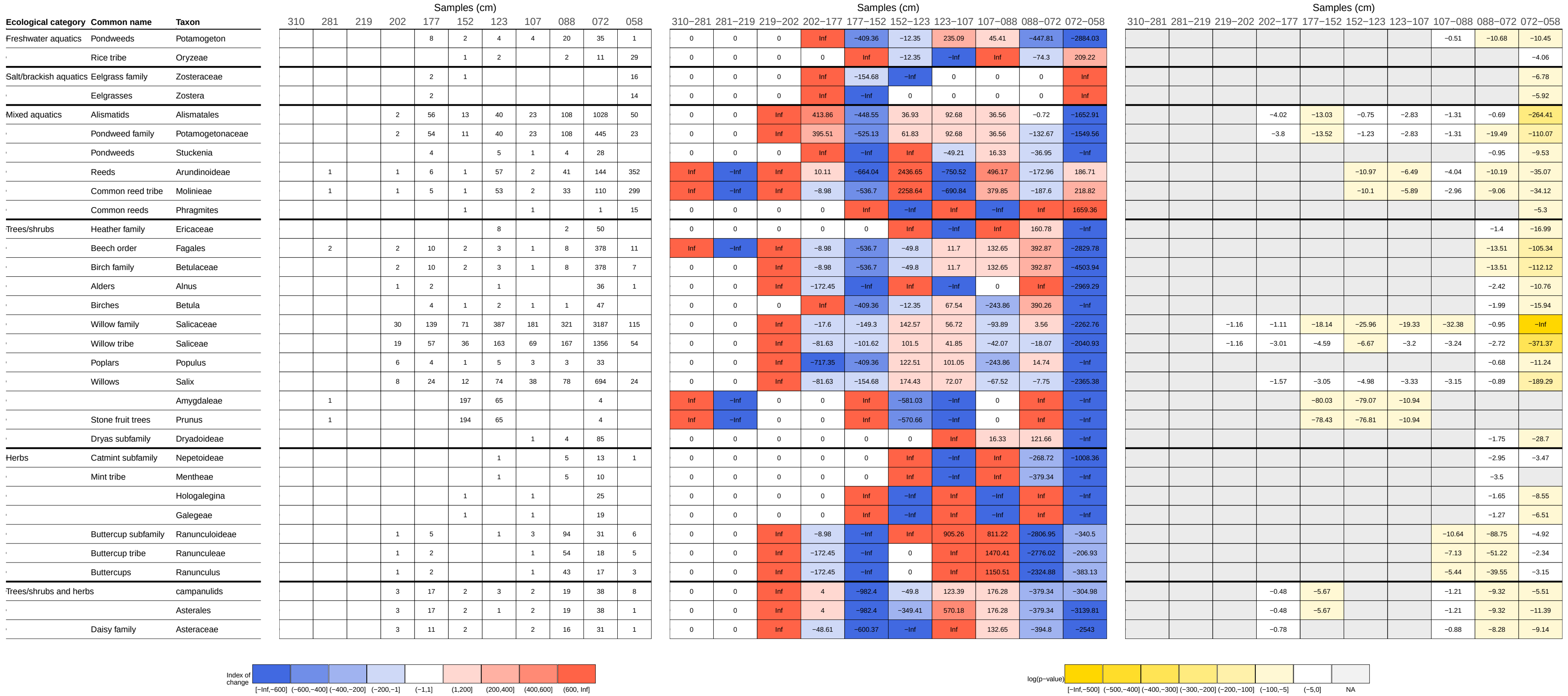
Figure 5.27: Stratification analysis for ELF044: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 0%.

Both samples in ELF044 have a high data yield and strongly *Zostera*-dominated taxonomic profiles. There is little data left to compare other taxa, and no comparisons have statistically significant changes. There is no evidence of stratification in ELF044.

5.2.16 ELF031A

The stratification plot for ELF031A is in two parts: figure 5.28 and figure 5.29. Most samples have predominantly woody taxa. Data yield varies considerably: the first four samples are nearly empty, 177-88 have several hundred Embryophyta reads, and 72 and 58 several thousand. 58 also shows an ecological change to mostly grasses, though reeds are seen through most of the core.

The lack of data in deeper samples is clear in both stratification figures. However, when comparisons are possible, an unusually high proportion are statistically significant (59.5%). Many are concentrated in the last comparison, involving 58, and these include some of the most significant values across all cores. The DNA profile of sample 58 is clearly very distinct. However, many middle samples also show significant changes, resulting in an overall signal completeness of 60%. Where data allows, ELF031A does appear stratified.

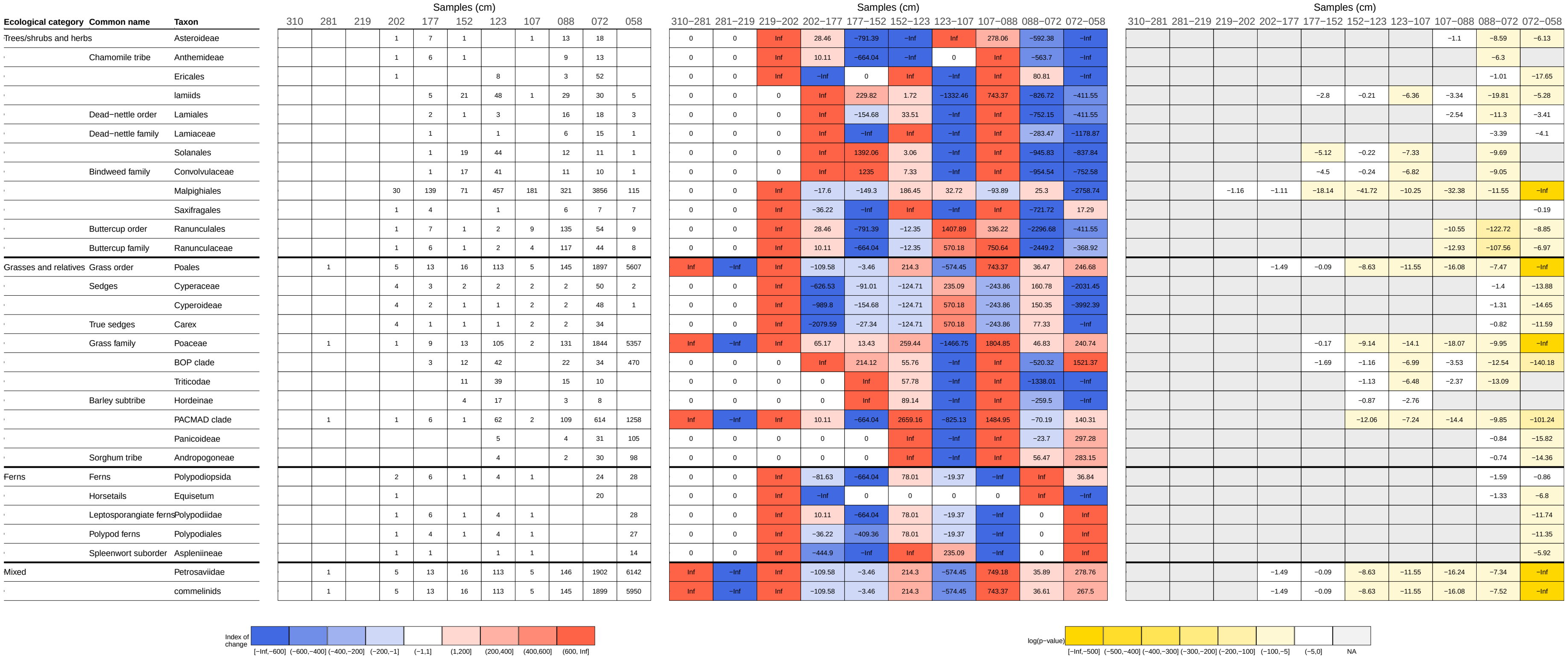


Ecological category	Common name	Taxon	310	281	219	202	177	152	123	107	088	072	058
			Samples (cm)										
Freshwater aquatics	Pondweeds	Potamogeton					8	2	4	4	20	35	1
	Rice tribe	Oryzeae						1	2		2	11	29
Salt/brackish aquatics	Eelgrass family	Zosteraceae					2	1					16
	Eelgrasses	Zostera					2						14
Mixed aquatics	Alismatids	Alismatales				2	56	13	40	23	108	1028	50
	Pondweed family	Potamogetonaceae				2	54	11	40	23	108	445	23
	Pondweeds	Stuckenia					4		5	1	4	28	
	Reeds	Arundinoideae		1		1	6	1	57	2	41	144	352
	Common reed tribe	Molinieae		1		1	5	1	53	2	33	110	299
	Common reeds	Phragmites						1		1		1	15
Trees/shrubs	Heather family	Ericaceae							8		2	50	
	Beech order	Fagales		2		2	10	2	3	1	8	378	11
	Birch family	Betulaceae				2	10	2	3	1	8	378	7
	Alders	Alnus				1	2		1			36	1
	Birches	Betula					4	1	2	1	1	47	
	Willow family	Salicaceae				30	139	71	387	181	321	3187	115
	Willow tribe	Saliceae				19	57	36	163	69	167	1356	54
	Poplars	Populus				6	4	1	5	3	3	33	
	Willows	Salix				8	24	12	74	38	78	694	24
		Amygdaleae		1				197	65			4	
	Stone fruit trees	Prunus		1				194	65			4	
	Dryas subfamily	Dryadoideae								1	4	85	
Herbs	Catmint subfamily	Nepetoideae							1		5	13	1
	Mint tribe	Mentheae							1		5	10	
		Hologalegina						1		1		25	
		Galegeae						1		1		19	
	Buttercup subfamily	Ranunculoideae				1	5		1	3	94	31	6
	Buttercup tribe	Ranunculeae				1	2			1	54	18	5
	Buttercups	Ranunculus				1	2			1	43	17	3
Trees/shrubs and herbs		campanulids				3	17	2	3	2	19	38	8
		Asterales				3	17	2	1	2	19	38	1
	Daisy family	Asteraceae				3	11	2		2	16	31	1

Taxon	310−281	281−219	219−202	202−177	177−152	152−123	123−107	107−088	088−072	072−058
	Samples (cm)									
Potamogeton	0	0	0	Inf	−409.36	−12.35	235.09	45.41	−447.81	−2884.03
Oryzeae	0	0	0	0	Inf	−12.35	−Inf	Inf	−74.3	209.22
Zosteraceae	0	0	0	Inf	−154.68	−Inf	0	0	0	Inf
Zostera	0	0	0	Inf	−Inf	0	0	0	0	Inf
Alismatales	0	0	Inf	413.86	−448.55	36.93	92.68	36.56	−0.72	−1652.91
Potamogetonaceae	0	0	Inf	395.51	−525.13	61.83	92.68	36.56	−132.67	−1549.56
Stuckenia	0	0	0	Inf	−Inf	Inf	−49.21	16.33	−36.95	−Inf
Arundinoideae	Inf	−Inf	Inf	10.11	−664.04	2436.65	−750.52	496.17	−172.96	186.71
Molinieae	Inf	−Inf	Inf	−8.98	−536.7	2258.64	−690.84	379.85	−187.6	218.82
Phragmites	0	0	0	0	Inf	−Inf	Inf	−Inf	Inf	1659.36
Ericaceae	0	0	0	0	0	Inf	−Inf	Inf	160.78	−Inf
Fagales	Inf	−Inf	Inf	−8.98	−536.7	−49.8	11.7	132.65	392.87	−2829.78
Betulaceae	0	0	Inf	−8.98	−536.7	−49.8	11.7	132.65	392.87	−4503.94
Alnus	0	0	Inf	−172.45	−Inf	Inf	−Inf	0	Inf	−2969.29
Betula	0	0	0	Inf	−409.36	−12.35	67.54	−243.86	390.26	−Inf
Salicaceae	0	0	Inf	−17.6	−149.3	142.57	56.72	−93.89	3.56	−2262.76
Saliceae	0	0	Inf	−81.63	−101.62	101.5	41.85	−42.07	−18.07	−2040.93
Populus	0	0	Inf	−717.35	−409.36	122.51	101.05	−243.86	14.74	−Inf
Salix	0	0	Inf	−81.63	−154.68	174.43	72.07	−67.52	−7.75	−2365.38
Amygdaleae	Inf	−Inf	0	0	Inf	−581.03	−Inf	0	Inf	−Inf
Prunus	Inf	−Inf	0	0	Inf	−570.66	−Inf	0	Inf	−Inf
Dryadoideae	0	0	0	0	0	0	Inf	16.33	121.66	−Inf
Nepetoideae	0	0	0	0	0	Inf	−Inf	Inf	−268.72	−1008.36
Mentheae	0	0	0	0	0	Inf	−Inf	Inf	−379.34	−Inf
Hologalegina	0	0	0	0	Inf	−Inf	Inf	−Inf	Inf	−Inf
Galegeae	0	0	0	0	Inf	−Inf	Inf	−Inf	Inf	−Inf
Ranunculoideae	0	0	Inf	−8.98	−Inf	Inf	905.26	811.22	−2806.95	−340.5
Ranunculeae	0	0	Inf	−172.45	−Inf	0	Inf	1470.41	−2776.02	−206.93
Ranunculus	0	0	Inf	−172.45	−Inf	0	Inf	1150.51	−2324.88	−383.13
campanulids	0	0	Inf	4	−982.4	−49.8	123.39	176.28	−379.34	−304.98
Asterales	0	0	Inf	4	−982.4	−349.41	570.18	176.28	−379.34	−3139.81
Asteraceae	0	0	Inf	−48.61	−600.37	−Inf	Inf	132.65	−394.8	−2543

Taxon	310−281	281−219	219−202	202−177	177−152	152−123	123−107	107−088	088−072	072−058
	Samples (cm)									
Potamogeton								−0.51	−10.68	−10.45
Oryzeae										−4.06
Zosteraceae										−6.78
Zostera										−5.92
Alismatales				−4.02	−13.03	−0.75	−2.83	−1.31	−0.69	−264.41
Potamogetonaceae				−3.8	−13.52	−1.23	−2.83	−1.31	−19.49	−110.07
Stuckenia									−0.95	−9.53
Arundinoideae						−10.97	−6.49	−4.04	−10.19	−35.07
Molinieae						−10.1	−5.89	−2.96	−9.06	−34.12
Phragmites										−5.3
Ericaceae									−1.4	−16.99
Fagales									−13.51	−105.34
Betulaceae									−13.51	−112.12
Alnus									−2.42	−10.76
Betula									−1.99	−15.94
Salicaceae			−1.16	−1.11	−18.14	−25.96	−19.33	−32.38	−0.95	−Inf
Saliceae			−1.16	−3.01	−4.59	−6.67	−3.2	−3.24	−2.72	−371.37
Populus									−0.68	−11.24
Salix				−1.57	−3.05	−4.98	−3.33	−3.15	−0.89	−189.29
Amygdaleae					−80.03	−79.07	−10.94			
Prunus					−78.43	−76.81	−10.94			
Dryadoideae									−1.75	−28.7
Nepetoideae									−2.95	−3.47
Mentheae									−3.5	
Hologalegina									−1.65	−8.55
Galegeae									−1.27	−6.51
Ranunculoideae								−10.64	−88.75	−4.92
Ranunculeae								−7.13	−51.22	−2.34
Ranunculus								−5.44	−39.55	−3.15
campanulids				−0.48	−5.67			−1.21	−9.32	−5.51
Asterales				−0.48	−5.67			−1.21	−9.32	−11.39
Asteraceae				−0.78				−0.88	−8.28	−9.14

Figure 5.28: Stratification analysis for ELF031A: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 60%. Part 1 of 2.



Ecological category	**Common name**	**Taxon**	Samples (cm) 310	281	219	202	177	152	123	107	088	072	058
Trees/shrubs and herbs		Asteroideae				1	7	1		1	13	18	
	Chamomile tribe	Anthemideae				1	6	1			9	13	
		Ericales				1			8		3	52	
		lamiids					5	21	48	1	29	30	5
	Dead−nettle order	Lamiales					2	1	3		16	18	3
	Dead−nettle family	Lamiaceae					1		1		6	15	1
		Solanales					1	19	44		12	11	1
	Bindweed family	Convolvulaceae					1	17	41		11	10	1
		Malpighiales				30	139	71	457	181	321	3856	115
		Saxifragales				1	4		1		6	7	7
	Buttercup order	Ranunculales				1	7	1	2	9	135	54	9
	Buttercup family	Ranunculaceae				1	6	1	2	4	117	44	8
Grasses and relatives	Grass order	Poales		1		5	13	16	113	5	145	1897	5607
	Sedges	Cyperaceae				4	3	2	2	2	2	50	2
		Cyperoideae				4	2	1	1	2	2	48	1
	True sedges	Carex				4	1	1	1	2	2	34	
	Grass family	Poaceae		1		1	9	13	105	2	131	1844	5357
		BOP clade					3	12	42		22	34	470
		Triticodae						11	39		15	10	
	Barley subtribe	Hordeinae						4	17		3	8	
		PACMAD clade		1		1	6	1	62	2	109	614	1258
		Panicoideae							5		4	31	105
	Sorghum tribe	Andropogoneae							4		2	30	98
Ferns	Ferns	Polypodiopsida				2	6	1	4	1		24	28
	Horsetails	Equisetum				1						20	
	Leptosporangiate ferns	Polypodiidae				1	6	1	4	1			28
	Polypod ferns	Polypodiales				1	4	1	4	1			27
	Spleenwort suborder	Aspleniineae				1	1		1	1			14
Mixed		Petrosaviidae		1		5	13	16	113	5	146	1902	6142
		commelinids		1		5	13	16	113	5	145	1899	5950

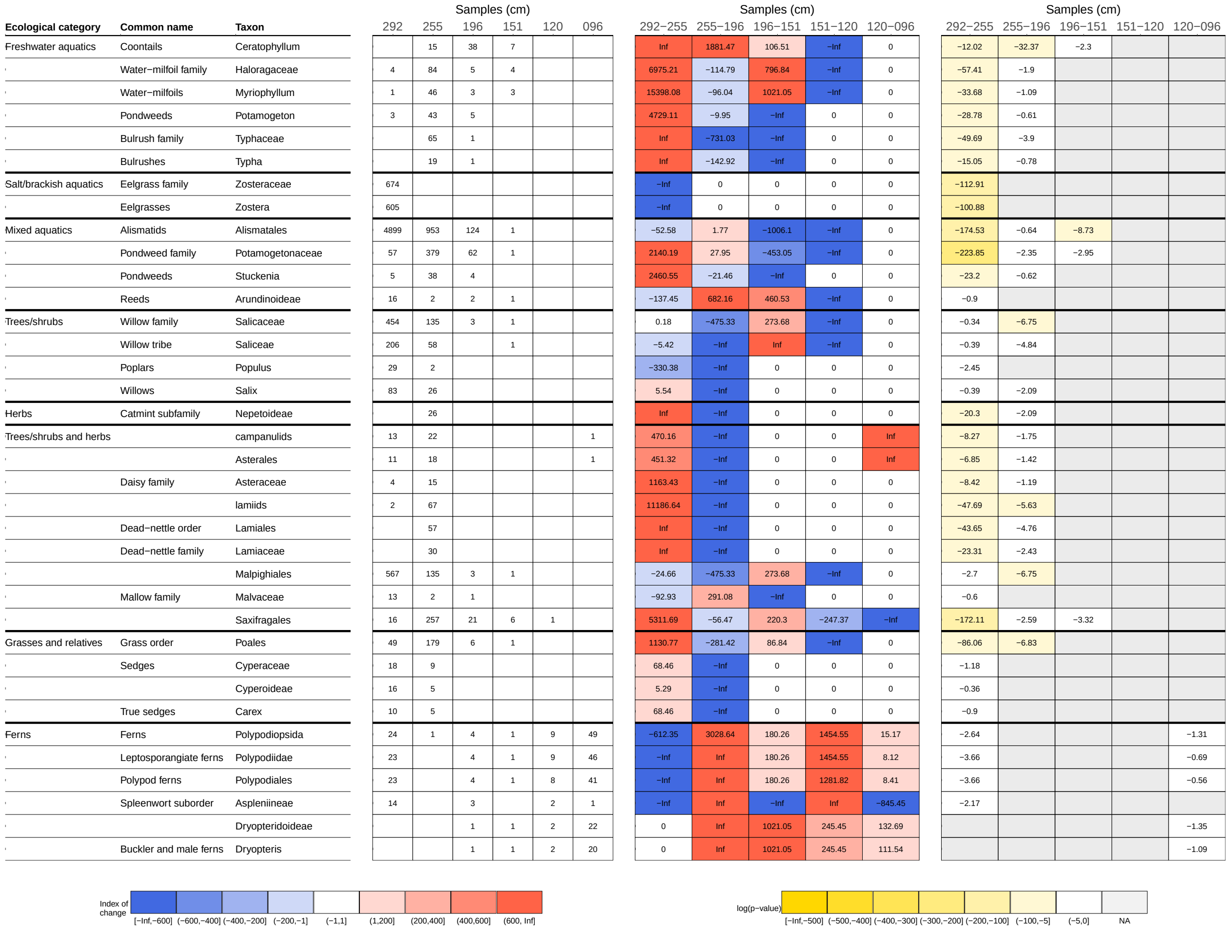
Taxon	Samples (cm) 310−281	281−219	219−202	202−177	177−152	152−123	123−107	107−088	088−072	072−058
Asteroideae	0	0	Inf	28.46	−791.39	−Inf	Inf	278.06	−592.38	−Inf
Anthemideae	0	0	Inf	10.11	−664.04	−Inf	0	Inf	−563.7	−Inf
Ericales	0	0	Inf	−Inf	0	Inf	−Inf	Inf	80.81	−Inf
lamiids	0	0	0	Inf	229.82	1.72	−1332.46	743.37	−826.72	−411.55
Lamiales	0	0	0	Inf	−154.68	33.51	−Inf	Inf	−752.15	−411.55
Lamiaceae	0	0	0	Inf	−Inf	Inf	−Inf	Inf	−283.47	−1178.87
Solanales	0	0	0	Inf	1392.06	3.06	−Inf	Inf	−945.83	−837.84
Convolvulaceae	0	0	0	Inf	1235	7.33	−Inf	Inf	−954.54	−752.58
Malpighiales	0	0	Inf	−17.6	−149.3	186.45	32.72	−93.89	25.3	−2758.74
Saxifragales	0	0	Inf	−36.22	−Inf	Inf	−Inf	Inf	−721.72	17.29
Ranunculales	0	0	Inf	28.46	−791.39	−12.35	1407.89	336.22	−2296.68	−411.55
Ranunculaceae	0	0	Inf	10.11	−664.04	−12.35	570.18	750.64	−2449.2	−368.92
Poales	Inf	−Inf	Inf	−109.58	−3.46	214.3	−574.45	743.37	36.47	246.68
Cyperaceae	0	0	Inf	−626.53	−91.01	−124.71	235.09	−243.86	160.78	−2031.45
Cyperoideae	0	0	Inf	−989.8	−154.68	−124.71	570.18	−243.86	150.35	−3992.39
Carex	0	0	Inf	−2079.59	−27.34	−124.71	570.18	−243.86	77.33	−Inf
Poaceae	Inf	−Inf	Inf	65.17	13.43	259.44	−1466.75	1804.85	46.83	240.74
BOP clade	0	0	0	Inf	214.12	55.76	−Inf	Inf	−520.32	1521.37
Triticodae	0	0	0	0	Inf	57.78	−Inf	Inf	−1338.01	−Inf
Hordeinae	0	0	0	0	Inf	89.14	−Inf	Inf	−259.5	−Inf
PACMAD clade	Inf	−Inf	Inf	10.11	−664.04	2659.16	−825.13	1484.95	−70.19	140.31
Panicoideae	0	0	0	0	0	Inf	−Inf	Inf	−23.7	297.28
Andropogoneae	0	0	0	0	0	Inf	−Inf	Inf	56.47	283.15
Polypodiopsida	0	0	Inf	−81.63	−664.04	78.01	−19.37	−Inf	Inf	36.84
Equisetum	0	0	Inf	−Inf	0	0	0	0	Inf	−Inf
Polypodiidae	0	0	Inf	10.11	−664.04	78.01	−19.37	−Inf	0	Inf
Polypodiales	0	0	Inf	−36.22	−409.36	78.01	−19.37	−Inf	0	Inf
Aspleniineae	0	0	Inf	−444.9	−Inf	Inf	235.09	−Inf	0	Inf
Petrosaviidae	Inf	−Inf	Inf	−109.58	−3.46	214.3	−574.45	749.18	35.89	278.76
commelinids	Inf	−Inf	Inf	−109.58	−3.46	214.3	−574.45	743.37	36.61	267.5

Taxon	Samples (cm) 310−281	281−219	219−202	202−177	177−152	152−123	123−107	107−088	088−072	072−058
Asteroideae								−1.1	−8.59	−6.13
Anthemideae									−6.3	
Ericales									−1.01	−17.65
lamiids					−2.8	−0.21	−6.36	−3.34	−19.81	−5.28
Lamiales								−2.54	−11.3	−3.41
Lamiaceae									−3.39	−4.1
Solanales					−5.12	−0.22	−7.33		−9.69	
Convolvulaceae					−4.5	−0.24	−6.82		−9.05	
Malpighiales			−1.16	−1.11	−18.14	−41.72	−10.25	−32.38	−11.55	−Inf
Saxifragales										−0.19
Ranunculales								−10.55	−122.72	−8.85
Ranunculaceae								−12.93	−107.56	−6.97
Poales				−1.49	−0.09	−8.63	−11.55	−16.08	−7.47	−Inf
Cyperaceae									−1.4	−13.88
Cyperoideae									−1.31	−14.65
Carex									−0.82	−11.59
Poaceae					−0.17	−9.14	−14.1	−18.07	−9.95	−Inf
BOP clade					−1.69	−1.16	−6.99	−3.53	−12.54	−140.18
Triticodae						−1.13	−6.48	−2.37	−13.09	
Hordeinae						−0.87	−2.76			
PACMAD clade						−12.06	−7.24	−14.4	−9.85	−101.24
Panicoideae									−0.84	−15.82
Andropogoneae									−0.74	−14.36
Polypodiopsida									−1.59	−0.86
Equisetum									−1.33	−6.8
Polypodiidae										−11.74
Polypodiales										−11.35
Aspleniineae										−5.92
Petrosaviidae				−1.49	−0.09	−8.63	−11.55	−16.24	−7.34	−Inf
commelinids				−1.49	−0.09	−8.63	−11.55	−16.08	−7.52	−Inf

Figure 5.29: Stratification analysis for ELF031A. Part 2 of 2.

5.2.17 ELF051

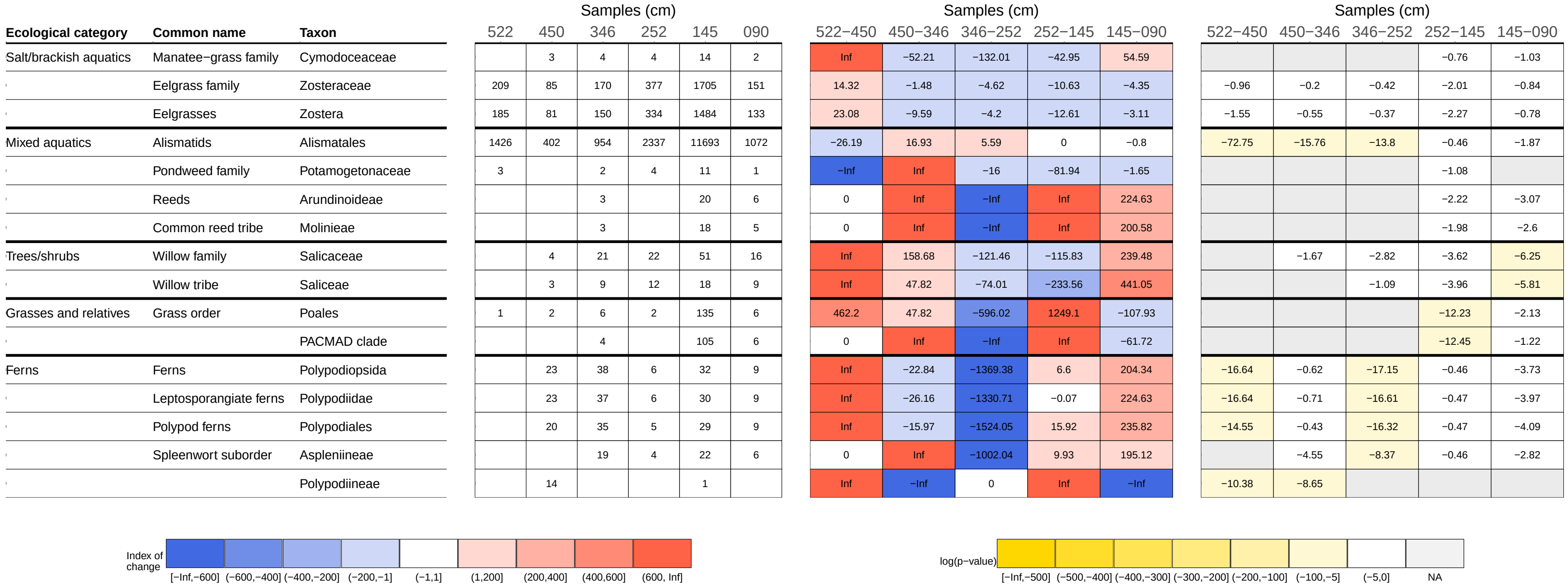
Unusually, ELF051 begins with a *Zostera*-dominated sample followed by three predominantly freshwater and two terrestrial, suggesting inversion or temporary reversal of inundation. The data yield is very high for the deepest sample and rapidly decreases, with the top two samples nearly empty. Accordingly, there are dramatic and statistically significant indices of change between the two deepest samples, but data quickly drops off. The handful of taxon comparisons at the top of the core are not statistically significant. The stratification signal is only 50% complete; it is present towards the bottom of the core only.



Ecological category	Common name	Taxon	Samples (cm) 292	255	196	151	120	096	Samples (cm) 292−255	255−196	196−151	151−120	120−096	Samples (cm) 292−255	255−196	196−151	151−120	120−096
Freshwater aquatics	Coontails	Ceratophyllum		15	38	7			Inf	1881.47	106.51	−Inf	0	−12.02	−32.37	−2.3		
	Water−milfoil family	Haloragaceae	4	84	5	4			6975.21	−114.79	796.84	−Inf	0	−57.41	−1.9			
	Water−milfoils	Myriophyllum	1	46	3	3			15398.08	−96.04	1021.05	−Inf	0	−33.68	−1.09			
	Pondweeds	Potamogeton	3	43	5				4729.11	−9.95	−Inf	0	0	−28.78	−0.61			
	Bulrush family	Typhaceae		65	1				Inf	−731.03	−Inf	0	0	−49.69	−3.9			
	Bulrushes	Typha		19	1				Inf	−142.92	−Inf	0	0	−15.05	−0.78			
Salt/brackish aquatics	Eelgrass family	Zosteraceae	674						−Inf	0	0	0	0	−112.91				
	Eelgrasses	Zostera	605						−Inf	0	0	0	0	−100.88				
Mixed aquatics	Alismatids	Alismatales	4899	953	124	1			−52.58	1.77	−1006.1	−Inf	0	−174.53	−0.64	−8.73		
	Pondweed family	Potamogetonaceae	57	379	62	1			2140.19	27.95	−453.05	−Inf	0	−223.85	−2.35	−2.95		
	Pondweeds	Stuckenia	5	38	4				2460.55	−21.46	−Inf	0	0	−23.2	−0.62			
	Reeds	Arundinoideae	16	2	2	1			−137.45	682.16	460.53	−Inf	0	−0.9				
Trees/shrubs	Willow family	Salicaceae	454	135	3	1			0.18	−475.33	273.68	−Inf	0	−0.34	−6.75			
	Willow tribe	Saliceae	206	58		1			−5.42	−Inf	Inf	−Inf	0	−0.39	−4.84			
	Poplars	Populus	29	2					−330.38	−Inf	0	0	0	−2.45				
	Willows	Salix	83	26					5.54	−Inf	0	0	0	−0.39	−2.09			
Herbs	Catmint subfamily	Nepetoideae		26					Inf	−Inf	0	0	0	−20.3	−2.09			
Trees/shrubs and herbs		campanulids	13	22				1	470.16	−Inf	0	0	Inf	−8.27	−1.75			
		Asterales	11	18				1	451.32	−Inf	0	0	Inf	−6.85	−1.42			
	Daisy family	Asteraceae	4	15					1163.43	−Inf	0	0	0	−8.42	−1.19			
		lamiids	2	67					11186.64	−Inf	0	0	0	−47.69	−5.63			
	Dead−nettle order	Lamiales		57					Inf	−Inf	0	0	0	−43.65	−4.76			
	Dead−nettle family	Lamiaceae		30					Inf	−Inf	0	0	0	−23.31	−2.43			
		Malpighiales	567	135	3	1			−24.66	−475.33	273.68	−Inf	0	−2.7	−6.75			
	Mallow family	Malvaceae	13	2	1				−92.93	291.08	−Inf	0	0	−0.6				
		Saxifragales	16	257	21	6	1		5311.69	−56.47	220.3	−247.37	−Inf	−172.11	−2.59	−3.32		
Grasses and relatives	Grass order	Poales	49	179	6	1			1130.77	−281.42	86.84	−Inf	0	−86.06	−6.83			
	Sedges	Cyperaceae	18	9					68.46	−Inf	0	0	0	−1.18				
		Cyperoideae	16	5					5.29	−Inf	0	0	0	−0.36				
	True sedges	Carex	10	5					68.46	−Inf	0	0	0	−0.9				
Ferns	Ferns	Polypodiopsida	24	1	4	1	9	49	−612.35	3028.64	180.26	1454.55	15.17	−2.64				−1.31
	Leptosporangiate ferns	Polypodiidae	23		4	1	9	46	−Inf	Inf	180.26	1454.55	8.12	−3.66				−0.69
	Polypod ferns	Polypodiales	23		4	1	8	41	−Inf	Inf	180.26	1281.82	8.41	−3.66				−0.56
	Spleenwort suborder	Aspleniineae	14		3		2	1	−Inf	Inf	−Inf	Inf	−845.45	−2.17				
		Dryopteridoideae			1	1	2	22	0	Inf	1021.05	245.45	132.69					−1.35
	Buckler and male ferns	Dryopteris			1	1	2	20	0	Inf	1021.05	245.45	111.54					−1.09

Figure 5.30: Stratification analysis for ELF051: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 50%.

5.2.18 ELF045



Ecological category	Common name	Taxon	522	450	346	252	145	090
Salt/brackish aquatics	Manatee−grass family	Cymodoceaceae		3	4	4	14	2
	Eelgrass family	Zosteraceae	209	85	170	377	1705	151
	Eelgrasses	Zostera	185	81	150	334	1484	133
Mixed aquatics	Alismatids	Alismatales	1426	402	954	2337	11693	1072
	Pondweed family	Potamogetonaceae	3		2	4	11	1
	Reeds	Arundinoideae			3		20	6
	Common reed tribe	Molinieae			3		18	5
Trees/shrubs	Willow family	Salicaceae		4	21	22	51	16
	Willow tribe	Saliceae		3	9	12	18	9
Grasses and relatives	Grass order	Poales	1	2	6	2	135	6
		PACMAD clade			4		105	6
Ferns	Ferns	Polypodiopsida		23	38	6	32	9
	Leptosporangiate ferns	Polypodiidae		23	37	6	30	9
	Polypod ferns	Polypodiales		20	35	5	29	9
	Spleenwort suborder	Aspleniineae			19	4	22	6
		Polypodiineae		14			1	

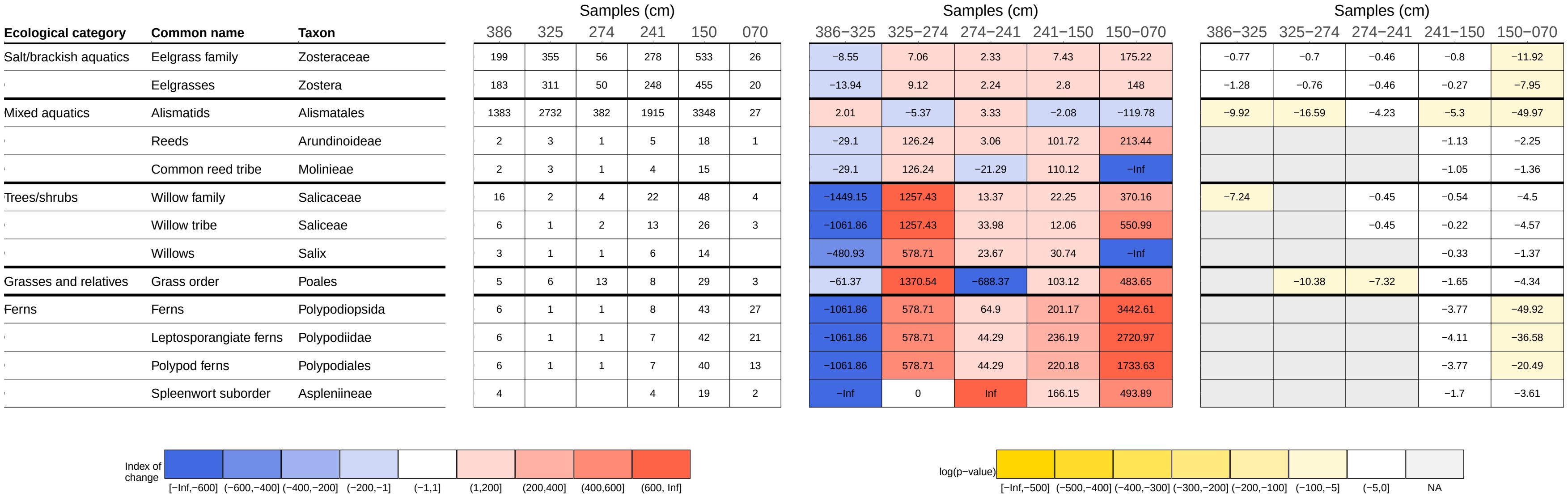
Ecological category	Common name	Taxon	522−450	450−346	346−252	252−145	145−090
Salt/brackish aquatics	Manatee−grass family	Cymodoceaceae	Inf	−52.21	−132.01	−42.95	54.59
	Eelgrass family	Zosteraceae	14.32	−1.48	−4.62	−10.63	−4.35
	Eelgrasses	Zostera	23.08	−9.59	−4.2	−12.61	−3.11
Mixed aquatics	Alismatids	Alismatales	−26.19	16.93	5.59	0	−0.8
	Pondweed family	Potamogetonaceae	−Inf	Inf	−16	−81.94	−1.65
	Reeds	Arundinoideae	0	Inf	−Inf	Inf	224.63
	Common reed tribe	Molinieae	0	Inf	−Inf	Inf	200.58
Trees/shrubs	Willow family	Salicaceae	Inf	158.68	−121.46	−115.83	239.48
	Willow tribe	Saliceae	Inf	47.82	−74.01	−233.56	441.05
Grasses and relatives	Grass order	Poales	462.2	47.82	−596.02	1249.1	−107.93
		PACMAD clade	0	Inf	−Inf	Inf	−61.72
Ferns	Ferns	Polypodiopsida	Inf	−22.84	−1369.38	6.6	204.34
	Leptosporangiate ferns	Polypodiidae	Inf	−26.16	−1330.71	−0.07	224.63
	Polypod ferns	Polypodiales	Inf	−15.97	−1524.05	15.92	235.82
	Spleenwort suborder	Aspleniineae	0	Inf	−1002.04	9.93	195.12
		Polypodiineae	Inf	−Inf	0	Inf	−Inf

Ecological category	Common name	Taxon	522−450	450−346	346−252	252−145	145−090
Salt/brackish aquatics	Manatee−grass family	Cymodoceaceae				−0.76	−1.03
	Eelgrass family	Zosteraceae	−0.96	−0.2	−0.42	−2.01	−0.84
	Eelgrasses	Zostera	−1.55	−0.55	−0.37	−2.27	−0.78
Mixed aquatics	Alismatids	Alismatales	−72.75	−15.76	−13.8	−0.46	−1.87
	Pondweed family	Potamogetonaceae				−1.08	
	Reeds	Arundinoideae				−2.22	−3.07
	Common reed tribe	Molinieae				−1.98	−2.6
Trees/shrubs	Willow family	Salicaceae		−1.67	−2.82	−3.62	−6.25
	Willow tribe	Saliceae			−1.09	−3.96	−5.81
Grasses and relatives	Grass order	Poales				−12.23	−2.13
		PACMAD clade				−12.45	−1.22
Ferns	Ferns	Polypodiopsida	−16.64	−0.62	−17.15	−0.46	−3.73
	Leptosporangiate ferns	Polypodiidae	−16.64	−0.71	−16.61	−0.47	−3.97
	Polypod ferns	Polypodiales	−14.55	−0.43	−16.32	−0.47	−4.09
	Spleenwort suborder	Aspleniineae		−4.55	−8.37	−0.46	−2.82
		Polypodiineae	−10.38	−8.65			

Figure 5.31: Stratification analysis for ELF045: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 50%

All samples in ELF045 contain mostly *Zostera* with a variable terrestrial signal. Data yield is high but variable, casting doubt on the three significant changes in Alismatales, which are all associated with corresponding changes in read count and proportion of *Zostera*/Zosteraceae. Nevertheless, significant changes in ferns among deeper samples and grasses and woody taxa towards the top result in an 50% complete stratification signal. Despite the homogeneous taxonomic profiles, there is some evidence for stratification across ELF045.

5.2.19 ELF047

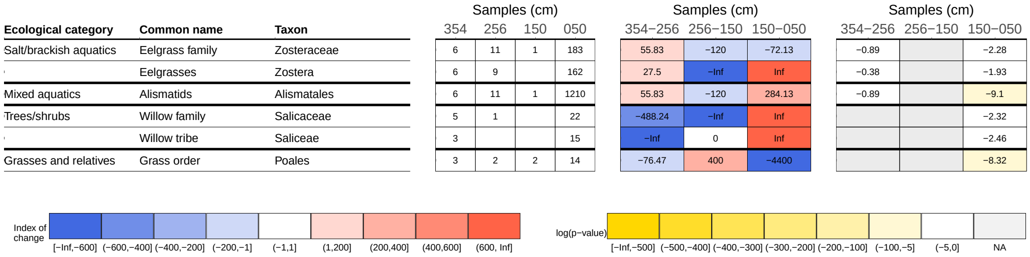


			Samples (cm)						Samples (cm)					Samples (cm)				
Ecological category	**Common name**	**Taxon**	386	325	274	241	150	070	386−325	325−274	274−241	241−150	150−070	386−325	325−274	274−241	241−150	150−070
Salt/brackish aquatics	Eelgrass family	Zosteraceae	199	355	56	278	533	26	−8.55	7.06	2.33	7.43	175.22	−0.77	−0.7	−0.46	−0.8	−11.92
	Eelgrasses	Zostera	183	311	50	248	455	20	−13.94	9.12	2.24	2.8	148	−1.28	−0.76	−0.46	−0.27	−7.95
Mixed aquatics	Alismatids	Alismatales	1383	2732	382	1915	3348	27	2.01	−5.37	3.33	−2.08	−119.78	−9.92	−16.59	−4.23	−5.3	−49.97
	Reeds	Arundinoideae	2	3	1	5	18	1	−29.1	126.24	3.06	101.72	213.44				−1.13	−2.25
	Common reed tribe	Molinieae	2	3	1	4	15		−29.1	126.24	−21.29	110.12	−Inf				−1.05	−1.36
Trees/shrubs	Willow family	Salicaceae	16	2	4	22	48	4	−1449.15	1257.43	13.37	22.25	370.16	−7.24		−0.45	−0.54	−4.5
	Willow tribe	Saliceae	6	1	2	13	26	3	−1061.86	1257.43	33.98	12.06	550.99			−0.45	−0.22	−4.57
	Willows	Salix	3	1	1	6	14		−480.93	578.71	23.67	30.74	−Inf				−0.33	−1.37
Grasses and relatives	Grass order	Poales	5	6	13	8	29	3	−61.37	1370.54	−688.37	103.12	483.65		−10.38	−7.32	−1.65	−4.34
Ferns	Ferns	Polypodiopsida	6	1	1	8	43	27	−1061.86	578.71	64.9	201.17	3442.61				−3.77	−49.92
	Leptosporangiate ferns	Polypodiidae	6	1	1	7	42	21	−1061.86	578.71	44.29	236.19	2720.97				−4.11	−36.58
	Polypod ferns	Polypodiales	6	1	1	7	40	13	−1061.86	578.71	44.29	220.18	1733.63				−3.77	−20.49
	Spleenwort suborder	Aspleniineae	4			4	19	2	−Inf	0	Inf	166.15	493.89				−1.7	−3.61

Figure 5.32: Stratification analysis for ELF047: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 50%.

Taxonomic profiles in ELF047 are all *Zostera*-dominated with a small woody signal, although the top sample (70) appears slightly more terrestrial. This is reflected in a string of significant comparisons between sample 70 and its neighbour, particularly for the increase in ferns. Unfortunately, other significant changes are few, and the very variable read count could explain the significant changes in Alismatales, *Zostera*, and Zosteraceae, reducing stratification signal completeness to 50%. Every sample pair shows some stratification, but each in a single direction only. Fortunately, because the direction differs between comparisons, we can at least infer that ELF047 overall was not subject to a single large vertical transport of DNA, such as the downwards percolation of sheep DNA in (Haile *et al.* 2007).

5.2.20 ELF047A



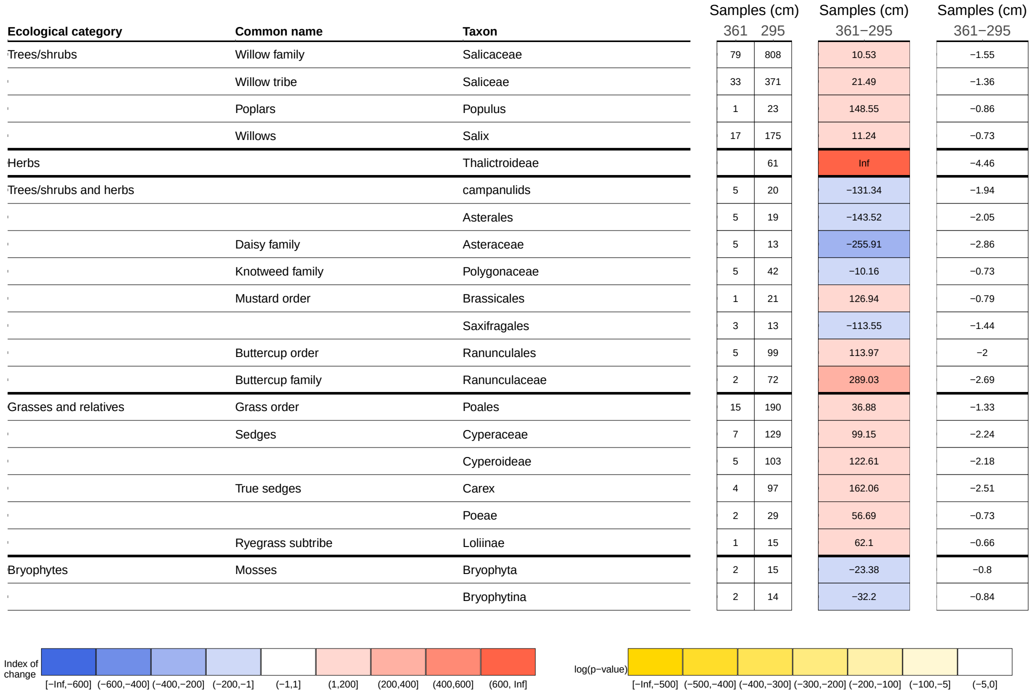
			Samples (cm)				Samples (cm)			Samples (cm)		
Ecological category	**Common name**	**Taxon**	354	256	150	050	354−256	256−150	150−050	354−256	256−150	150−050
Salt/brackish aquatics	Eelgrass family	Zosteraceae	6	11	1	183	55.83	−120	−72.13	−0.89		−2.28
	Eelgrasses	Zostera	6	9		162	27.5	−Inf	Inf	−0.38		−1.93
Mixed aquatics	Alismatids	Alismatales	6	11	1	1210	55.83	−120	284.13	−0.89		−9.1
Trees/shrubs	Willow family	Salicaceae	5	1		22	−488.24	−Inf	Inf			−2.32
	Willow tribe	Saliceae	3			15	−Inf	0	Inf			−2.46
Grasses and relatives	Grass order	Poales	3	2	2	14	−76.47	400	−4400			−8.32

Figure 5.33: Stratification analysis for ELF047A: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 33.33%.

ELF047A shows a strong *Zostera* signal alongside a small terrestrial, although data yield is very low for all but the top sample (50). Only a handful of comparisons between the top two samples have enough data to calculate p-values and only two are significant. One of those is an increase in Alismatales that at first glance appears to follow the proportional-increase-with-data-yield trend, but as *Zostera*, Zosteraceae, and Alismatales are nearly absent from the previous sample, I interpret this as basically an appearance instead of a proportional change. Along with a decrease in Poales, it suggests stratification in both directions for sample 50. Signal completeness is only 33.33% because of the absence of data in deeper samples.

5.2.21 ELF049

The two samples in ELF049 share a reasonable data yield and similar taxonomic profiles consisting of mostly floodplain woodland. Accordingly, most indices of change are moderate and none are associated with significant p-values. There is no clear evidence for stratification between these two samples.



Ecological category	Common name	Taxon	Samples (cm)		Samples (cm)	Samples (cm)
			361	295	361−295	361−295
Trees/shrubs	Willow family	Salicaceae	79	808	10.53	−1.55
	Willow tribe	Saliceae	33	371	21.49	−1.36
	Poplars	Populus	1	23	148.55	−0.86
	Willows	Salix	17	175	11.24	−0.73
Herbs		Thalictroideae		61	Inf	−4.46
Trees/shrubs and herbs		campanulids	5	20	−131.34	−1.94
		Asterales	5	19	−143.52	−2.05
	Daisy family	Asteraceae	5	13	−255.91	−2.86
	Knotweed family	Polygonaceae	5	42	−10.16	−0.73
	Mustard order	Brassicales	1	21	126.94	−0.79
		Saxifragales	3	13	−113.55	−1.44
	Buttercup order	Ranunculales	5	99	113.97	−2
	Buttercup family	Ranunculaceae	2	72	289.03	−2.69
Grasses and relatives	Grass order	Poales	15	190	36.88	−1.33
	Sedges	Cyperaceae	7	129	99.15	−2.24
		Cyperoideae	5	103	122.61	−2.18
	True sedges	Carex	4	97	162.06	−2.51
		Poeae	2	29	56.69	−0.73
	Ryegrass subtribe	Loliinae	1	15	62.1	−0.66
Bryophytes	Mosses	Bryophyta	2	15	−23.38	−0.8
		Bryophytina	2	14	−32.2	−0.84

Figure 5.34: Stratification analysis for ELF049: cumulative read counts, indices of change, and corresponding ln(p-values) between adjacent samples. "N/A" indicates insufficient data. Stratification signal completeness: 0%.

5.2.22 Attempting to exclude vertical DNA movement in non-stratified samples using alternative environmental proxies

Stratification analysis searches for positive evidence of DNA stratification. Similar DNA profiles between adjacent samples could be explained by DNA movement, but also by a similar environment. This section therefore compares the stratification analysis results with those from three other environmental proxies: geology, foraminifera, and ostracods. Martin Bates, Richard Bates, and John Whittaker conducted the analyses and determined the environmental reconstructions used here as part of the Europe's Lost Frontiers project. Because foraminifera and ostracods are aquatic organisms, there is a bias towards alluvial sediments. Nevertheless, the reconstructions overlap considerably with the sedaDNA samples: information was available for 108 of the 131 sample comparisons in the stratification analysis. The 12% of comparisons without information from other proxies are excluded in this section, but should be revisited when other results, such as pollen and plant macrofossils, become available.

Table 5.3 summarises the environmental interpretations derived from the geology, foraminifera, and ostracods (collected in "core reports") and the stratification analysis results. These agree for 32% of comparisons: either the core report shows no environmental change and the stratification analysis shows similar DNA profiles, or the core report suggests environmental change and the stratification analysis suggests at least partial stratification (lack of DNA movement either upwards or downwards). However, 64% of comparisons have no evidence of environmental change in the core report, but do suggest stratification, and usually in both directions. For many comparisons, the sedaDNA appears to show changes despite a constant environment. This is simply due to the higher sensitivity of the stratification analysis: it is designed to find differences in DNA profiles even if they do not necessarily represent changes in the biomass of taxa, in plant community, or in the wider environment. The environmental reconstructions from the core reports, on the other hand, are summaries and do not consider individual taxa unless they alter the overall reconstruction. It is unsurprising to see changes in DNA profiles in apparently constant environments.

Table 5.3: Comparing stratification results with environmental reconstructions based on geology, foraminifera, and ostracods analysed by colleagues in Europe's Lost Frontiers. Sample comparisons without a reconstruction are excluded. The "Agreement" column notes whether changes were observed in the environmental reconstruction, the DNA profiles, or both. Blue indicates more change in DNA profiles; red indicates more in the environmental reconstructions, which suggests DNA movement in the sediment.

		Environment according to core report		Stratification signal			
Core	Comparison	Below	Above	Upwards	Downwards	Both	Agreement
ELF027	177-174	Estuary strand-line	Estuary strand-line	1	1	1	Change in DNA only
ELF027	174-160	Estuary strand-line	Estuary strand-line	1	1	1	Change in DNA only
ELF059	378-359	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF059	359-337	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF059	337-280	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF059	280-270	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF059	270-230	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF059	230-210	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF059A	355-320	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF059A	320-250	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF059A	250-190	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF059A	190-135	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF060	465-420	Tidal river, mudflats	Tidal river, mudflats	0	0	0	No change in either
ELF060	420-350	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF060	350-250	Tidal river, mudflats	Tidal river, mudflats	0	1	0	Partial change in DNA only
ELF032A	177-153	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF032A	153-117	Tidal river, mudflats	Tidal river, mudflats	1	0	0	Partial change in DNA only
ELF032A	117-95	Tidal river, mudflats	Tidal river, mudflats	0	1	0	Partial change in DNA only
ELF032A	95-74	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF032A	74-47	Tidal river, mudflats	Tidal river, mudflats	1	1	1	Change in DNA only
ELF033	203-195	Vegetated river fringe	Vegetated river fringe	1	0	0	Partial change in DNA only
ELF033	183-176	Tidal river, saltmarsh	Tidal river, saltmarsh	1	1	1	Change in DNA only
ELF033	176-155	Tidal river, saltmarsh	Tidal river, saltmarsh	0	1	0	Partial change in DNA only
ELF033	155-118	Tidal river, saltmarsh	Brackish mudflats	1	0	0	Change in report; partial change in DNA
ELF033	118-75	Brackish mudflats	Brackish mudflats	1	1	1	Change in DNA only
ELF033	75-46	Brackish mudflats	Brackish mudflats	0	0	0	No change in either
ELF033A	98-70	Brackish mudflats	Brackish mudflats	0	0	0	No change in either
ELF033A	70-50	Brackish mudflats	Brackish mudflats	0	0	0	No change in either
ELF034A	166-146	Brackish mudflats, saltmarsh	Brackish mudflats, saltmarsh	0	0	0	No change in either
ELF034A	146-126	Brackish mudflats, saltmarsh	Brackish mudflats, saltmarsh	0	1	0	Partial change in DNA only
ELF034A	126-81	Brackish mudflats, saltmarsh	Brackish mudflats, saltmarsh	1	1	1	Change in DNA only
ELF034	219-202	Vegetated river	Vegetated river	1	0	0	Partial change in DNA only
ELF034	202-185	Vegetated river	Vegetated river	0	0	0	No change in either
ELF034	185-177	Vegetated river	Vegetated river	0	0	0	No change in either
ELF034	177-157	Vegetated river	Vegetated river	0	0	0	No change in either
ELF034	157-132	Vegetated river	Vegetated river	0	0	0	No change in either
ELF034	132-94	Vegetated river	Vegetated river	0	0	0	No change in either
ELF034	94-79	Vegetated river	Brackish mudflats, saltmarsh	0	0	0	Change in core report only
ELF034	79-61	Brackish mudflats, saltmarsh	Brackish mudflats, saltmarsh	0	0	0	No change in either
ELF054	182-140	Brackish algae/seagrass	Brackish algae/seagrass	0	0	0	No change in either
ELF054	140-58	Brackish algae/seagrass	Brackish algae/seagrass	0	0	0	No change in either
ELF039	415-384	Freshwater	Freshwater	1	1	1	Change in DNA only
ELF039	384-355	Freshwater	Freshwater	1	1	1	Change in DNA only
ELF039	355-341	Freshwater	Estuary	1	1	1	Change in both
ELF039	341-321	Estuary	Estuary	1	1	1	Change in DNA only
ELF039	321-250	Estuary	Brackish tidal flats	1	1	1	Change in both

Continued on next page

Table 5.3 continued

Core	Comparison	Environment according to core report Below	 Above	Stratification signal Upwards	 Downwards	 Both	Agreement
ELF039	250-145	Brackish tidal flats	Brackish tidal flats	0	1	0	No change in either
ELF040A	350-298	Brackish mudflats, saltmarsh	Brackish mudflats, saltmarsh	1	1	1	Change in DNA only
ELF040A	298-208	Brackish mudflats, saltmarsh	Brackish mudflats, saltmarsh	1	1	1	Change in DNA only
ELF040A	208-192	Brackish mudflats, saltmarsh	Brackish mudflats, saltmarsh	1	1	1	Change in DNA only
ELF040A	192-112	Brackish mudflats, saltmarsh	Brackish mudflats, saltmarsh	1	1	1	Change in DNA only
ELF040A	112-95	Brackish mudflats, saltmarsh	Brackish mudflats, saltmarsh	0	1	0	Partial change in DNA only
ELF041	295-180	Brackish mudflats	Brackish mudflats	1	1	1	Change in DNA only
ELF041	180-110	Brackish mudflats	Brackish mudflats	0	0	0	No change in either
ELF041	110-87	Brackish mudflats	Brackish mudflats	1	1	1	Change in DNA only
ELF042	350-250	Brackish mudflats; outer estuarine?	Brackish mudflats; outer estuarine?	1	1	1	Change in DNA only
ELF042	250-151	Brackish mudflats; outer estuarine?	Brackish mudflats; outer estuarine?	1	1	1	Change in DNA only
ELF042	151-65	Brackish mudflats; outer estuarine?	Brackish mudflats; outer estuarine?	1	1	1	Change in DNA only
ELF044	137-90	Outer estuarine	Outer estuarine	0	0	0	No change in either
ELF031A	310-281	River	River	0	0	0	No change in either
ELF031A	281-219	River	River	0	0	0	No change in either
ELF031A	219-202	River	River	0	0	0	No change in either
ELF031A	202-177	River	River	0	0	0	No change in either
ELF031A	123-107	Tidal, vegetated river	Tidal, vegetated river	1	1	1	Change in DNA only
ELF031A	107-88	Tidal, vegetated river	Tidal, vegetated river	1	1	1	Change in DNA only
ELF031A	88-72	Tidal, vegetated river	Tidal, vegetated river	1	1	1	Change in DNA only
ELF031A	72-58	Tidal, vegetated river	Tidal, vegetated river	1	1	1	Change in DNA only
ELF051	292-255	Freshwater	Vegetated margin of tidal river or lake	1	1	1	Change in both
ELF051	255-196	Vegetated margin of tidal river or lake	Vegetated margin of tidal river or lake	1	1	1	Change in DNA only
ELF051	196-151	Vegetated margin of tidal river or lake	Vegetated margin of tidal river or lake	0	1	0	Partial change in DNA only
ELF051	151-120	Vegetated margin of tidal river or lake	River	0	0	0	Change in core report only
ELF051	120-96	River	Brackish mudflats	0	0	0	Change in core report only
ELF045	522-450	Estuarine	Estuarine	1	1	1	Change in DNA only
ELF045	450-346	Estuarine	Estuarine	1	1	1	Change in DNA only
ELF045	346-252	Estuarine	Brackish mudflats	1	1	1	Change in both
ELF045	252-145	Brackish mudflats	Brackish mudflats	1	0	0	Partial change in DNA only
ELF047	386-325	Outer estuary	Outer estuary	1	1	1	Change in DNA only
ELF047	325-274	Outer estuary	Outer estuary	1	1	1	Change in DNA only
ELF047	274-241	Outer estuary	Outer estuary	0	1	0	Partial change in DNA only
ELF047	241-150	Outer estuary	Outer estuary	0	1	0	Partial change in DNA only
ELF047	150-70	Outer estuary	Outer estuary	1	1	1	Change in DNA only
ELF047A	354-256	Brackish mudflats	Brackish mudflats	0	0	0	No change in either
ELF047A	256-150	Brackish mudflats	Brackish mudflats	0	0	0	No change in either
ELF047A	150-50	Brackish mudflats	Brackish mudflats	1	1	1	Change in DNA only
ELF049	361-295	Outer estuarine	Outer estuarine	0	0	0	No change in either

This could potentially be explained by greater resolution in the sedaDNA data, particularly for the terrestrial environment. The core reports are limited to broader environmental changes reflected in the geology, foraminifera, and ostracods, such as the arrival of tides or salt water with marine inundation. However, similar habitats can support different plant communities, and only one taxon would need to be significantly different for the stratification analysis to find a signal. For example, in ELF031A, the core reports show a tidal, vegetated river for both samples 88 and 72. However, between these samples the stratification analysis reports small but significant decreases in *Potamogeton*, reeds, and *Ranunculus* (buttercups), but a significant increase in Betulaceae. These changes in plant community are not necessarily associated with a broader environmental change that would create visible differences to geology. The terrestrial community in particular may undergo significant changes (such as ecological succession) independently of aquatic factors affecting foraminifera or ostracods, such as temperature or salinity. Therefore, it is unsurprising that the stratification analysis finds differences in DNA profiles across similar broad environments.

Instead, the main aim of this validation analysis is to determine whether the environment was constant for sample comparisons where DNA profiles also appear similar: where there is no evidence of DNA stratification. A change in the environment in that case would suggest movement of DNA between the strata. If we assume a change in environment results in a different plant community and therefore the DNA entering the sediment, yet the DNA profiles remain statistically indistinguishable, the simplest explanation would be the mixing of DNA between strata. Table 5.3 highlights three such comparisons where the DNA profiles are similar yet the core reports show environmental change. While I would argue that the "Vegetated margin of tidal river or lake" to "River" transition in ELF051_151-120 is not sufficiently different, ELF034_94-79 and ELF051_120-96 require investigation.

ELF034_94-79 moves from "Vegetated river" to "Brackish mudflats, saltmarsh". The reason for no change in the DNA profile is immediately apparent: both samples have zero Embryophyta reads. There is therefore no reason to suspect DNA movement. For ELF051_120 and 96, the core report states "River" and "Brackish mudflats" respectively. These samples did have Embryophyta reads; enough for five taxa to be compared in the stratification analysis, none of which appeared significantly different. However, with only 11 and 53 reads, neither sample can be considered particularly representative of the environment. I would argue that the lack of change in DNA profile here could also be attributed to a lack of data, particularly as there were many significant changes observed between samples further down ELF051 that did have more reads (notably not 151-120 mentioned above, which have only 30 reads between them).

In conclusion, evidence of DNA stratification was absent for many samples, but in the vast majority of cases, the environment also appeared consistent according to geology, foraminifera, and ostracods. Those similar DNA profiles can therefore be interpreted as deriving from similar plant communities. Of the two samples where alternative proxies do suggest an environmental change despite no difference in DNA profiles, the similar profiles can reasonably be attributed to a lack of data. There is no convincing evidence of vertical DNA movement in these cores.

However, validating the stratification analysis results against the low-resolution environmental reconstructions in the core reports may have failed to detect other instances where similar DNA profiles were derived from dissimilar environments. While vertical DNA movement can be excluded for many samples in these cores, a more thorough investigation of the remainder should be performed when results from other environmental proxies are available, particularly those with information on terrestrial plants, such as pollen. Further information about the sediment itself could also suggest whether water movement, thought to be one mechanism of DNA leaching (Giguet-Covex *et al.* 2014), was likely in these cores. For example, ongoing sediment analysis by Martin Bates has measured the percentage moisture content for samples in five cores as part of loss-on-ignition tests. The results are summarised in table 5.4. On average, samples were relatively dry despite coming from marine cores. This does not exclude movement of water through the sediments before inundation, but is evidence against the percolation of seawater after burial. Furthermore, a lack of any vertical water movement is supported by the relatively consistent uranium:thorium ratio measured during OSL dating of cores by Tim Kinnaird (2020; personal communication). The taphonomy of sedaDNA after deposition is not yet well understood (Birks and Birks 2016), but the detailed contextual information that will soon be available from Europe's Lost Frontiers may allow some hypotheses to be tested.

Table 5.4: Mean moisture (%) in loss-on-ignition samples from five cores. Produced by Martin Bates as part of Europe's Lost Frontiers.

Core	n	Mean moisture (%)	St. dev.
ELF003	22	18.204	11.066
ELF031	37	22.552	10.264
ELF034	83	49.057	18.136
ELF044A	42	19.164	5.422
ELF031A	65	17.282	7.349

5.2.23 Summary of stratification analysis results

Table 5.5: Summary of stratification analysis by core. Signal completeness is the percentage of instances demonstrating stratification, where there each comparison has two possible instances: upward and downward movement. 100% means that every sample shows at least one statistically significant increase and decrease.

Core	Embryophyta data yield	Number of comparisons	Signal completeness (%)
ELF022	Low	0	N/A (single sample)
ELF027	Variable	2	100.00
ELF059	High	6	83.33
ELF059A	High	4	100.00
ELF060	High	3	50.00
ELF032A	High	5	80.00
ELF033	Variable	8	68.75
ELF033A	Low	5	10.00
ELF034A	Variable	10	25.00
ELF034	Low	8	6.25
ELF054	Variable	7	7.14
ELF039	High	8	81.25
ELF040A	Variable	6	50.00
ELF041	Variable	3	50.00
ELF042	High	3	83.33
ELF044	High	1	0.00
ELF044A	High	0	N/A (single sample)
ELF031	Low	2	N/A (insufficient data)
ELF031A	Variable	10	50.00
ELF051	Variable	5	50.00
ELF045	High	5	80.00
ELF047	Variable	5	50.00
ELF047A	Variable	3	33.33
ELF049	Variable	1	0.00
ELF046A	NA	0	N/A (single sample)
ELF050	Low	3	N/A (insufficient data)
ELF053	Low	4	N/A (insufficient data)

Table 5.5 summarises the stratification results from each core using signal completeness. Recall that every sample comparison could potentially demonstrate a lack of vertical DNA movement in two instances: from sample A to B (downwards) and from B to A (upwards). Signal completeness is the percentage of instances showing stratification. 19 of the 21 cores suitable for stratification analysis show some signal, as did 72 of the 108 sample comparisons. "Some signal" includes comparisons with evidence against DNA movement either upwards or downwards only, such as those in ELF047. This leaves open the possibility of DNA movement in the other direction between those samples. Fortunately, the direction frequently differed from one sample comparison to the next. No core showed evidence of stratification in only a single direction. While smaller episodes of vertical DNA movement in cores with <100% signal completeness cannot be excluded, the widespread evidence against movement suggests this is unlikely.

Furthermore, for a subset of sample comparisons that did not appear stratified, three alternative environmental proxies overwhelmingly confirmed a constant environment: the similar DNA profiles could be explained by a similar plant community contributing DNA to the sediment. The three exceptions had very little data. While this validation was of limited resolution and could not prove a lack of vertical DNA movement, it did not produce evidence to the contrary. Current investigations have produced no reason to suspect that leaching occurred in any of these cores. The sedaDNA interpretations in this project can broadly be assumed to be in context, and therefore comparable with other environmental proxies from the same strata. Stratification analysis has helped to demonstrate the validity of sedaDNA for palaeoenvironmental reconstruction.

However, as a very new tool, limitations have become apparent and it would benefit from further development. For example, as is becoming the case for many analyses in this project, the strength of the stratification signal appears correlated with data yield, despite the use of proportional counts. There is a minimum data requirement, and the more taxa meet this requirement, the more opportunities there are for statistically significant increases and decreases between samples. Most higher-yield cores had relatively high signal completeness, although there were exceptions (such as ELF044 where almost all reads were assigned to *Zostera*, leaving few other taxa with enough to compare).

Furthermore, the behaviour of different taxa needs further investigation. Most higher taxa were overly influenced by cumulative counting, and *Zostera* and its parent taxa by data yield (Chapter 4: Taxonomic results). However, there was no systematic study of taxon-specific patterns or, more fundamentally, the difficulties of comparing taxa of different ranks. PIA will not assign reads to leaf taxa, so datasets comprising many different ranks will likely be a problem in similar studies. At what level does cumulative counting stop taxa from being informative? For example, commelinids in ELF039 are a higher taxon and capture a large proportion of reads across the core, yet show several large and significant changes. Because ranks are artificial constructs, the level may well vary considerably for different taxa. Future development could consider a more sophisticated way to deal with parent and child taxa than cumulative counting.

Other future work could apply the stratification analysis to different questions. Firstly, it could search for evidence of stratification in data from other quantitative environmental proxies, such as pollen. Taxonomic results from other proxies typically have a smaller range of ranks than DNA data, so may be less affected by the difficulties of comparing different ranks. Secondly, the analysis could be applied horizontally instead of vertically, to compare DNA profiles from different cores but potentially the same strata. Future work for Europe's Lost Frontiers will involve assigning dates to sediment samples where possible (see Chapter 7: Discussion for a preliminary attempt). Samples may also be linked by geology or other environmental proxies. If other evidence suggests that samples originate from the same stratum and are approximately the same age, do we see significant changes in the DNA profile? The validation results above suggest that this may be possible. This could inform the taphonomy of DNA in these environments; perhaps adjacent DNA profiles differ because the taxa involved leave a very local sedaDNA signal, for instance. Statistical analysis of differences between taxonomic results is an important addition to the sedaDNA toolkit. It has demonstrated the ability to produce useful evidence for stratification of DNA in sediment and may also prove useful for answering other quantitative questions in environmental reconstruction.

Chapter 6

Mesophilic taxa and human disturbance indicators

Chapter 5 considered authentication on a sample-by-sample basis. This chapter instead evaluates specific Embryophyta taxa that may have implications for climate or human activity. Some are not native to Great Britain, so would be a surprising find in a recent palaeoenvironment in a similar location if it experienced a similar climate. For example, grasses in Panicoideae are currently native to southern Europe only, so their presence in the Southern River valley may imply warmer temperatures than today. *Tilia* is another mesophilic taxon, native to Britain but near the edge of its thermal limit (Eaton *et al.* 2016). Some *Tilia* reads have been dated to before its expected arrival date in Britain (Chapter 7: Discussion). Pollen records suggest that *Tilia* arrived from the southeast (Brewer *et al.* 2017), so new early evidence in the Southern River region would fit with existing data, but would benefit from further support via damage authentication. Finally, some taxa may suggest human-aided dispersal, such as *Hordeum* (barley). There are native British barleys, but *H. vulgare* was an important domesticate of early northern European farmers (Betti *et al.* 2020).

For each taxon, the MetaDamage analysis was performed on pooled reads from across all samples. Many mesophilic or potentially human-associated taxa were noted in Chapter 4 (Taxonomic results), including from Metazoa, but this chapter considers only those with the most data, so the most likely to provide informative results. However, there were still rarely sufficient reads for a reliable MetaDamage signal, so this chapter also makes use of additional Embryophyta data from cores ELF001-20, which were otherwise analysed by Roselyn Ware, and preliminary results from the deep sequencing of selected cores. MetaDamage results from the original data are presented first, followed by results from the addition of all other data.

However, even pooling data did not always provide sufficient reads for a confident MetaDamage result. Therefore, the issue of finding a damage signal was also approached from an alternative angle. First, an average damage signal was calculated from all Embryophyta data. Then, the binomial distribution was used to calculate the probability of finding a non-zero signal assuming that the quantity of data for a given taxon showed that average damage rate. This binomial probability analysis allowed some quantification of whether a lack of any signal at all could be explained by a low sample size, and whether a non-zero damage signal was above or below average.

This chapter starts with an explanation of the binomial probability analysis followed by two examples of uncontroversial taxa showing good damage signals to aid later comparison. These also demonstrate the improvement in signal that can be achieved when the amount of data is increased. Finally, MetaDamage results from mesophilic and potentially human-associated taxa are discussed in this context.

6.1 Binomial probability analysis

The binomial probability analysis calculates the probability of any MetaDamage outcome if we assume prior knowledge of the damage level. It uses the binomial distribution to calculate the probability of every possible mismatch rate that MetaDamage could return for a given taxon. The three parameters required to calculate a binomial probability are the number of trials, the probability of a trial being successful, and an observed success rate (from which the MetaDamage mismatch rate is derived):

1. The number of trials is the number of position-zero C bases in the reference sequences (roughly 25% of the total read count, but calculated precisely by MetaDamage) for the taxon in question. Each is an opportunity for a C-to-T mismatch in the corresponding sample sequence.
2. The probability of a trial being successful is the assumed mismatch rate: the proportion of those C bases that should appear as Ts if the DNA shows a given amount of damage. Here, this is the position-zero C-to-T mismatch rate from a combined MetaDamage analysis of all Embryophyta reads, including the additional data. This came to 0.125 (3 d.p.) and was associated with a very clear damage signal (figure 6.1). On average, every read should have a 0.125 probability of a C-to-T mismatch. We may expect variation due to age and preservation (Kistler *et al.* 2017), but not with taxonomic assignment. Does this taxon show average damage?
3. The observed success rate is simply the number of position-zero C bases (trials) that *did* show a C-to-T mismatch for that taxon. We calculate a probability for every possible success rate; if there were 40 position-zero C bases for a taxon - 40 trials - the success rate could be any integer from 0-40. The potential success rates are equivalent to the potential MetaDamage mismatch rates, the only difference being that a success rate is a count of C-to-T mismatches whereas MetaDamage presents them as a proportion. So, a success rate of 20/40 is equivalent to a MetaDamage mismatch rate of 0.5. Both would have the same probability.

The result is a probability distribution for all possible MetaDamage outcomes for a given taxon (e.g. figure 6.4). Assuming a mismatch rate of 0.125, we can first ask the

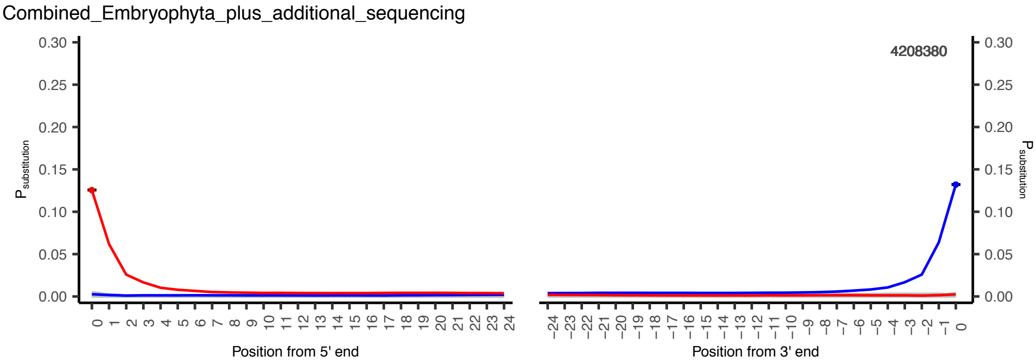


Figure 6.1: MetaDamage profile for all Embryophyta reads, including additional sequencing. $P_{substitution}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right.

probability of MetaDamage returning a zero mismatch rate (no damage signal). Recall that Chapter 5 (Authentication results) found many samples to lack MetaDamage signals, but all could be explained by a lack of data. Mismatches may be absent because there are none to be found, but also by chance if the sample is small. A taxon with fewer reads will be increasingly likely to return zero mismatches, even if the "right" answer is above zero, resulting in a shift of the probability distribution towards zero as trial number decreases. Furthermore, fewer trials allow fewer possible MetaDamage outcomes, reducing the precision of the result, yet an outcome of zero mismatches is always possible. A high probability of no signal and/or a low resolution distribution strongly suggest that the number of reads is too small.

However, if the most probable outcome is close to the assumed mismatch rate of 0.125, we can confirm that this amount of data is sufficient to detect an average signal, so our confidence in the observed MetaDamage outcome is improved. We can then compare the expected and observed outcomes. If they are significantly different, we can conclude that the actual mismatch value for this taxon is above or below the measured average.

6.2 Comparative results from more confident taxa

6.2.1 Fagaceae

Figures 6.2 and 6.3 show the MetaDamage results for Fagaceae, the oak family, which also includes beeches and chestnuts. These three genera are native to Great Britain and are very plausible in Doggerland sediments, with *Quercus* (oak) pollen reported from cores in Wolters *et al.* (2010), Krüger *et al.* (2017), and Brown *et al.* (2018). We could expect Fagaceae reads to show damage. However, Fagaceae was chosen as a comparison for less secure taxa because it shares with several of them a relatively low read count: the original data only returned 130 reads. This was rarely enough for a convincing signal in Chapter 5 (Authentication results) and is close to the limit of sensitivity (100-300 reads). Nevertheless, figure 6.2 shows a clear uptick in age-related mismatches on both ends, although the confidence interval for G-to-A overlaps the rate further along the molecule. There is a reasonable argument for the Fagaceae reads being ancient, so they can be interpreted in samples with some confidence.

The signal is improved by adding more data, as expected. Figure 6.3 shows the results from the original data, cores ELF001-20, and the additional deep sequencing of some cores: 1,067 reads in total. The confidence intervals are narrowed and all mismatch types show a smoother curve. Note that the position-zero mismatch values are similar, but figure 6.2 has an extended *y*-axis to accommodate its wide confidence intervals.

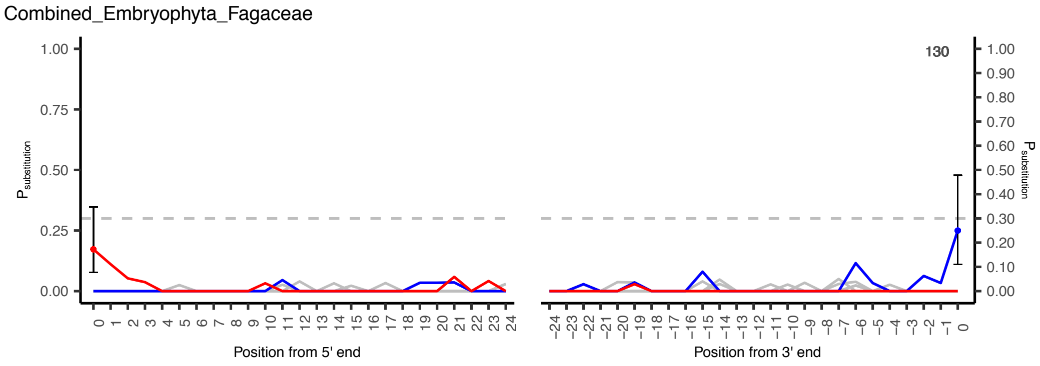


Figure 6.2: MetaDamage profile for Fagaceae (original data only). The position-zero C-to-T mismatch rate is 0.172 95% CI [0.077-0.347].

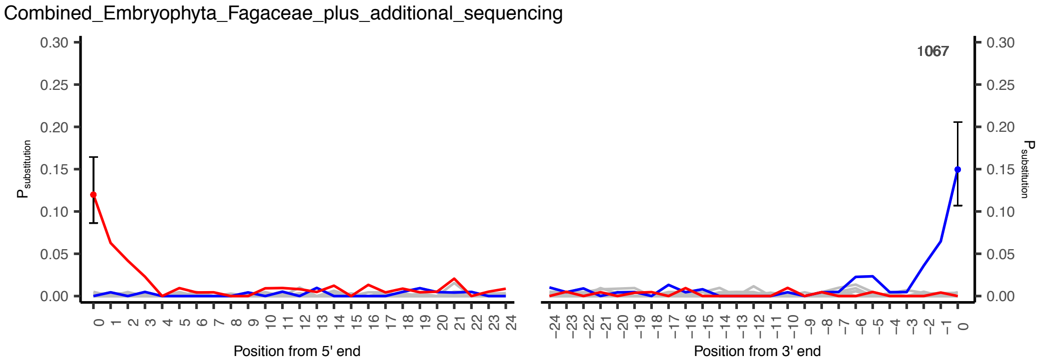


Figure 6.3: MetaDamage profile for Fagaceae (all data). The position-zero C-to-T mismatch rate is 0.120 95% CI [0.086-0.164].

The next step is to use the bionomial probability analysis to assess whether, if the 1,067 Fagaceae reads showed the average mismatch rate of 0.125, their 267 trials (number of position-0 C bases in the 1,067 reference sequences) would be sufficient for MetaDamage to detect it. Is there enough data for the average mismatch rate to be found, or is the sample small enough that a rate of zero is likely to be returned by chance? Figure 6.4 shows the probability distribution for possible MetaDamage outcomes based on the average mismatch rate and 267 trials. The probability of zero mismatches is <0.001 and the most probable outcome is instead 0.124, very close to the expected value. This confirms that MetaDamage would most likely return approximately the "right" answer from 267 trials. This increases our confidence in the observed MetaDamage output.

Now we can compare the observed MetaDamage output to the expected output derived from the average mismatch rate. The shaded area in figure 6.4 shows the observed mismatch rate and its 95% confidence interval: 0.120 95% CI [0.086-0.164] (red). The peak of the distribution, the expected MetaDamage outcome from the average mismatch rate, lies well within the confidence interval. The damage level for Fagaceae is therefore not significantly different from the average. Both the MetaDamage output and the bionomial probability analysis present good evidence that the Fagaceae reads are ancient.

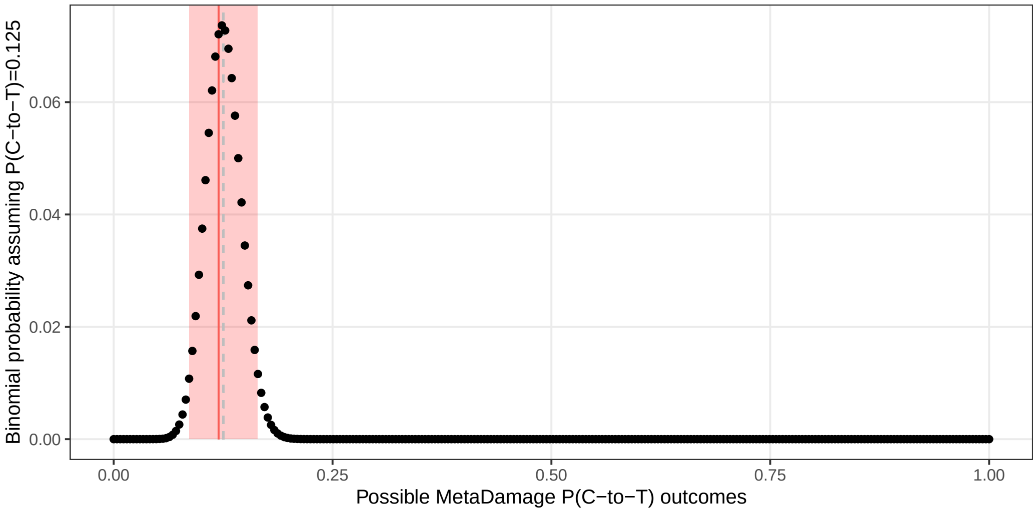


Figure 6.4: Probability distribution of all possible position-zero C-to-T mismatch rates produced by MetaDamage for Fagaceae (all data; 267 trials), assuming the average mismatch rate of 0.125 (dashed line). The probability of zero mismatches is <0.001 and the most probable outcome is 0.124. The observed mismatch rate and its 95% confidence interval are in red.

6.2.2 Corylus

Corylus (hazel) is the second example of a plausible taxon from Holocene Doggerland. Although its parent family Betulaceae was very common in the data, far fewer reads were assigned specifically to *Corylus* than Fagaceae, so this provides a more realistic comparison for low-frequency exotic taxa. Figure 6.5 shows the MetaDamage profile for the 28 reads in the original data. Understandably, few mismatches were found, but these include age-associated C-to-T and G-to-A mismatches at the very ends. The additional data only increases the read count to 96, but presents a surprisingly convincing damage signal (figure 6.6). The upticks remain, and despite wide confidence intervals, both exclude variation from further along the molecules or from other mismatch types. A good damage signal with <100 reads suggests that *Corylus* shows above-average damage.

The binomial probability analysis can test this. The smaller number of trials compared to Fagaceae is apparent in the lower resolution of figure 6.7 and the non-negligible probability of MetaDamage failing to find a signal: 0.023. However, the most probable outcome is above zero, and relatively close to the assumed value at 0.107. There should

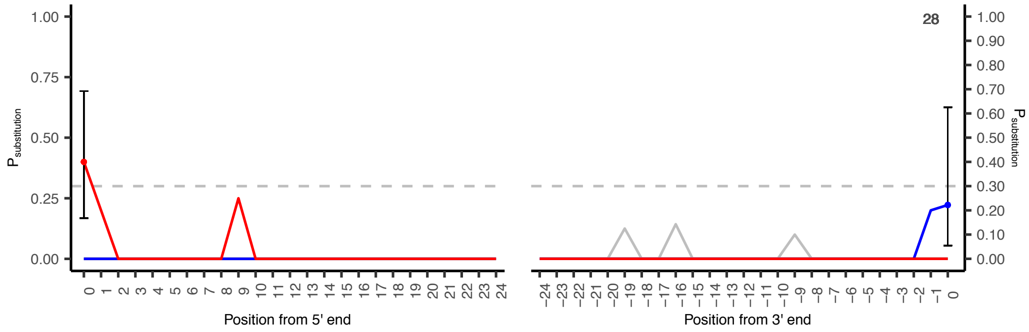

Combined_Embryophyta_Corylus

Figure 6.5: MetaDamage profile for *Corylus* (original data only). The position-zero C-to-T mismatch rate is 0.400 95% CI [0.167, 0.692].

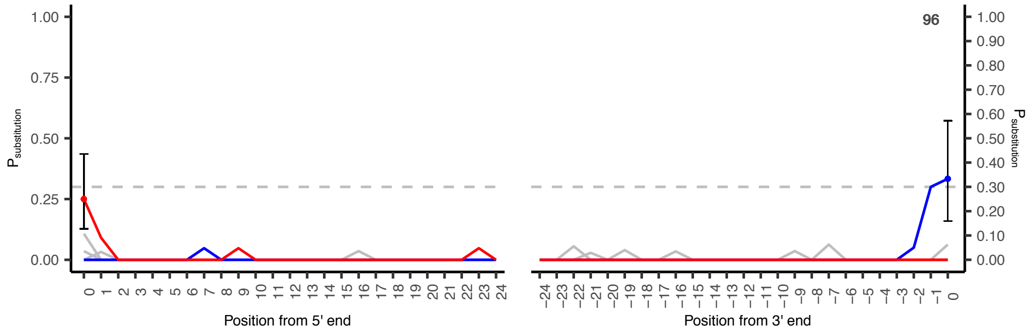

Combined_Embryophyta_Corylus_plus_additional_sequencing

Figure 6.6: MetaDamage profile for *Corylus* (all data). The position-zero C-to-T mismatch rate is 0.250 95% CI [0.127, 0.435].

be sufficient *Corylus* data to detect an average damage signal, albeit less accurately than Fagaceae.

The observed mismatch rate is actually rather high at 0.250 95% CI [0.127, 0.435] (figure 6.6). This is significantly higher than 0.107, so the observed mismatch rate is in fact greater than average. *Corylus* shows a convincing damage signal despite limited data.

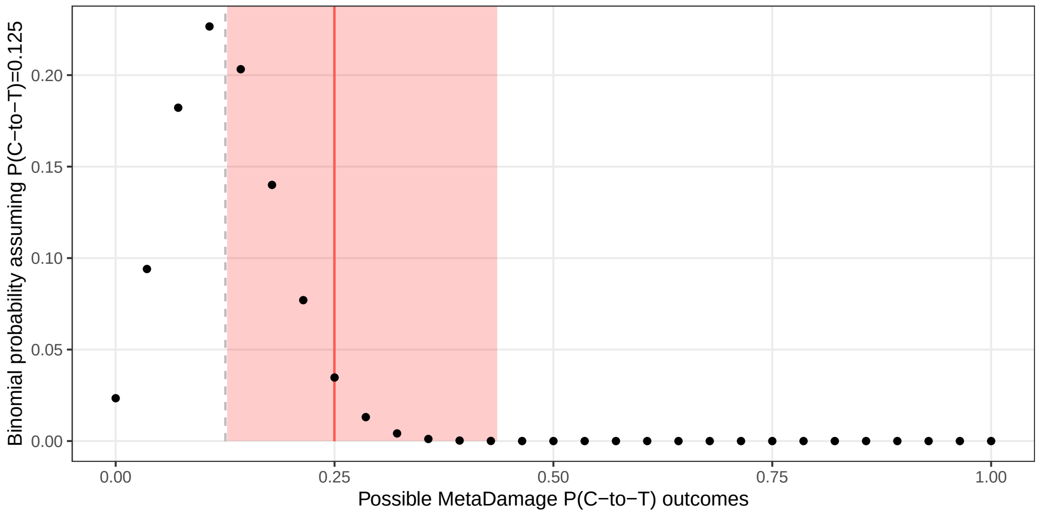


Figure 6.7: Probability distribution of all possible position-zero C-to-T mismatch rates produced by MetaDamage for *Corylus* (all data; 28 trials), assuming the average mismatch rate of 0.125 (dashed line). The probability of zero mismatches is 0.023 and the most probable outcome is 0.107. The observed mismatch rate and its 95% confidence interval are in red.

6.3 Results from mesophilic taxa and human disturbance indicators

6.3.1 Panicoideae

The first of the unexpected taxa to consider is Panicoideae. This subfamily contains grasses native to southern Europe, but not to Great Britain, yet occurred in many samples and often with relatively high biogenomic mass. However, Panicoideae also appeared in some negative controls, suggesting possible modern contamination, and it also includes the heavily-studied maize and sorghum, which suggests potential assignment error.

Figures 6.8 and 6.9 show the MetaDamage profiles for the original data and all data respectively. Unlike for previous taxa, the read count of the original data was already very high, and the differences between the two plots are minor. There is a very clear damage signal. The sequences do therefore appear ancient, so there is no evidence of significant modern contamination.

However, the high overall mismatch rate suggests limited similarity to the first BLAST hits, which in turn suggests that these reads were matched to related taxa rather than the true taxa during the taxonomic assignment BLAST. A genuine Panicoideae signal would imply warmer temperatures in the Southern River than present in 53% of samples, so accurate assignment is important. Do these reads derive from Panicoideae?

The high mismatch rate can potentially be explained by over-represented taxa in Panicoideae, which are not addressed by MetaDamage but are by PIA. Because the GenBank database contains so many maize and sorghum reads, it is likely that the first BLAST hit for any genuine Panicoideae read will be to maize or sorghum. PIA considers many BLAST hits to reduce the impact of over-represented taxa, but MetaDamage matches a read to a reference sequence using only the first BLAST hit. This is simply

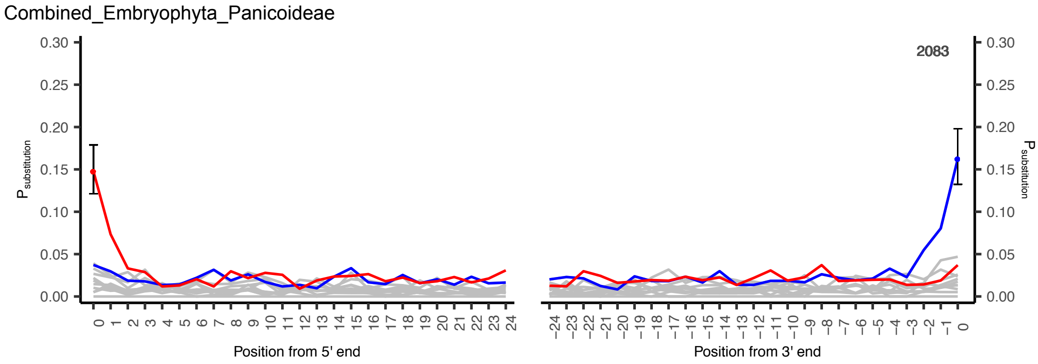


Figure 6.8: MetaDamage profile for Panicoideae (original data only). The position-zero C-to-T mismatch rate is 0.148 95% CI [0.121, 0.179]. Data from 76 samples.

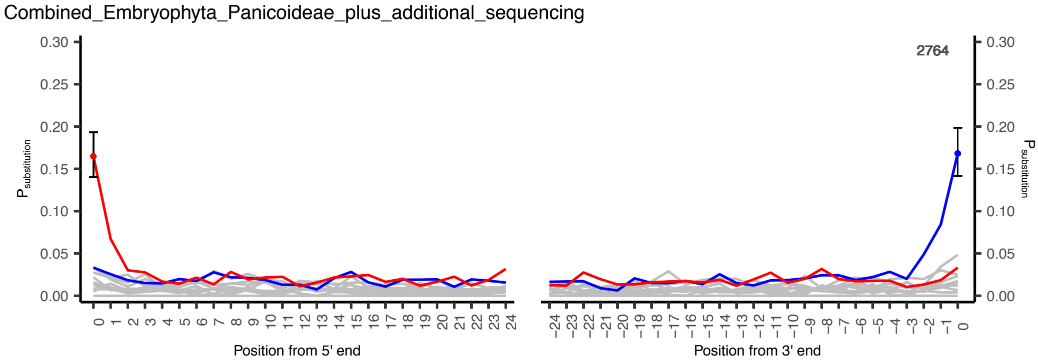


Figure 6.9: MetaDamage profile for Panicoideae (all data). The position-zero C-to-T mismatch rate is 0.165 95% CI [0.140, 0.193].

because searching for a closer match would reduce the number of mismatches, which can be informative for MetaDamage. Assuming that (a) the reads do not actually derive from maize or sorghum, which are not European, and (b) Panicoideae has the general plant substitution rate of approximately 6 x 10^{-9} substitutions per site per year (Wolfe *et al.* 1989), we can use the mean mismatch rate found by MetaDamage to estimate the typical divergence time between reads and their reference sequences. The mean mismatch rate, omitting C-to-T and G-to-A rates from the end five bases to ignore the damage signal, was 0.010 (3 d.p.; n = 580, standard deviation = 0.007). This produces an estimated divergence time of ~1.6 million years; well within the age of Panicoideae (Burke *et al.* 2016). The high background mismatch rate can therefore be explained by the over-represented taxa in Panicoideae, which PIA is designed to address. The MetaDamage output gives us no reason to question the assignment of the reads to Panicoideae.

However, MetaDamage cannot confirm the assignment either. The question of misassignment was re-examined by correlating the read counts of Panicoideae with every other European taxon across the initial sequencing data from cores ELF001-60. The ten closest-correlating taxa besides Panicoideae itself are reported in table 6.1. Unlike in the correlation analysis with the non-European taxa in Chapter 5 (Authentication results), it is possible that the reads do derive from Panicoideae. Many of the taxa in table 6.1 are child or parent taxa of Panicoideae, so their co-occurrence is consistent with Panicoideae reads being assigned at various ranks. The four taxa highlighted in red are not in the Panicoideae taxonomy; these could alternatively be explained by similar habitat preferences. However, the second-highest coefficient belongs to *Phragmites*. Not only is this often dominant to the point of monoculture (Holdredge and Bertness 2011, Packer *et al.* 2017), it was noted in Chapter 5 as a strong candidate for mis-assignment to maize and sorghum. *Phragmites* and its parent tribe Molinieae, also positively correlated, are close relatives of Panicoideae and are significantly less well-represented in the reference database. Although this is not decisive evidence, *Phragmites* offers a more parsimonious explanation for the Panicoideae reads.

Table 6.1: The ten taxa most correlated with Panicoideae apart from Panicoideae itself, in descending order of Pearson's correlation coefficient.

Taxon	Correlation with Panicoideae
Paniceae	0.952
Phragmites	0.793
Panicum	0.658
BOP clade	0.634
Andropogoneae	0.629
Cenchrinae	0.519
Asplenium	0.510
Poaceae	0.501
Melinidinae	0.482
Molinieae	0.417

The damage signal is clear, but let us compare it to the average. The binomial probability distribution in figure 6.10 has high resolution, <0.001 probability of zero mismatches, and the most probable outcome is equal to the assumed mismatch rate, 0.125. There is enough data for a high probability of returning the true value. The actual mismatch rate found by MetaDamage is significantly higher at 0.165 95% CI [0.140, 0.193], although the difference may be explained by the high background mismatch rate. The Panicoideae reads reveal a limitation of the binomial probability analysis with its exclusive focus on position zero: the analysis expects position zero to be the tip of an uptick, but high variation or a high baseline can contradict this assumption. I would not be confident interpreting above-average damage in Panicoideae reads because of this, but it is clear that they do show a strong damage signal, making modern contamination unlikely. The reads do appear ancient. However, their assignment remains questionable.

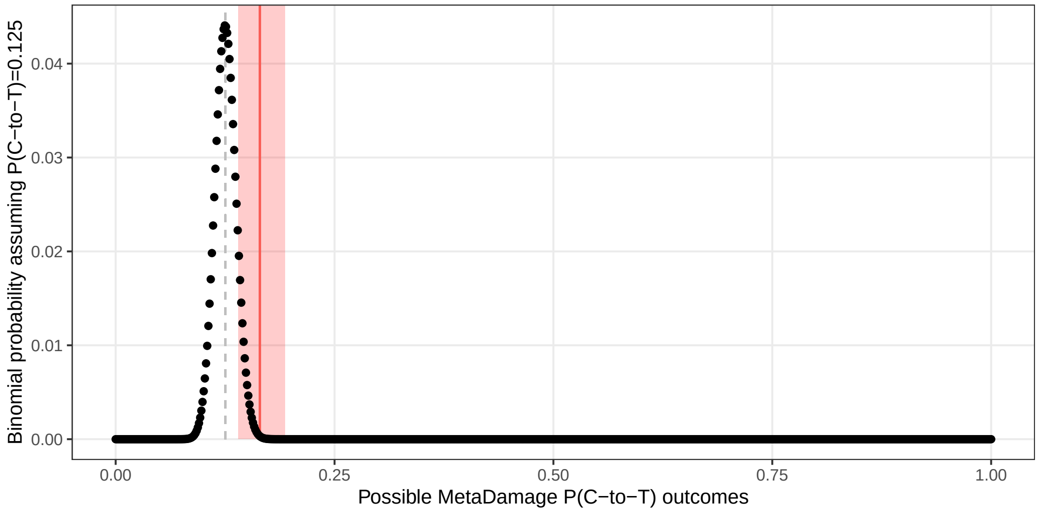


Figure 6.10: Probability distribution of all possible position-zero C-to-T mismatch rates produced by MetaDamage for Panicoideae (all data; 746 trials), assuming the average mismatch rate of 0.125 (dashed line). The probability of zero mismatches is <0.001 and the most probable mismatch rate is 0.125. The observed mismatch rate and its 95% confidence interval are in red.

6.3.2 Tilioideae

The subfamily Tilioideae has only one European species: *Tilia* (limes). *Tilia* is native to Great Britain, but near the edge of its thermal range, as it requires warm summers to produce fertile seed (Eaton *et al.* 2016). It can therefore act as another indicator of a relatively warm climate. However, Tilioideae is in the same family as several widely used and over-represented organisms, especially cottons (*Gossypium*, *Gossypioides*), raising the possibility of contamination from modern sources or mis-assignment. It therefore requires careful authentication.

Unfortunately, Tilioideae reads only occur at low frequency. The MetaDamage profiles (figures 6.11 and 6.12) have suggestions of upticks, but the confidence intervals overlap non-age-associated mismatch types. It is not a convincing signal, even with the additional data.

Does the binomial probability analysis suggest a lack of data impeding the result? The distribution in figure 6.13 has relatively low resolution and a 0.046 probability of finding zero mismatches. However, the most probable outcome, 0.130, is also the closest to the assumed mismatch value. There does appear to be enough data for a reliable MetaDamage result.

The observed MetaDamage outcome is 0.217 95% CI [0.098, 0.422]. The confidence interval is wide, reflecting the variability in the data, but does include 0.130. Despite the weak signal in the MetaDamage plots, the behaviour of the Tilioideae data is consistent with an average level of damage.

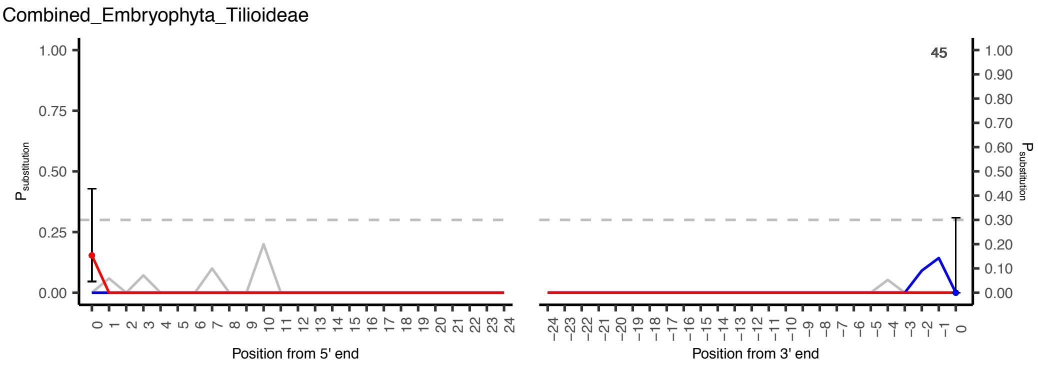


Figure 6.11: MetaDamage profile for Tilioideae (original data only). Tilioideae appeared in samples ELF027_174 and 160; ELF032A_177; ELF033_155; ELF039_415, 321, and 250; ELF042_350 and 065; and ELF059_270. The position-zero C-to-T mismatch rate is 0.154 95% CI [0.047, 0.428].

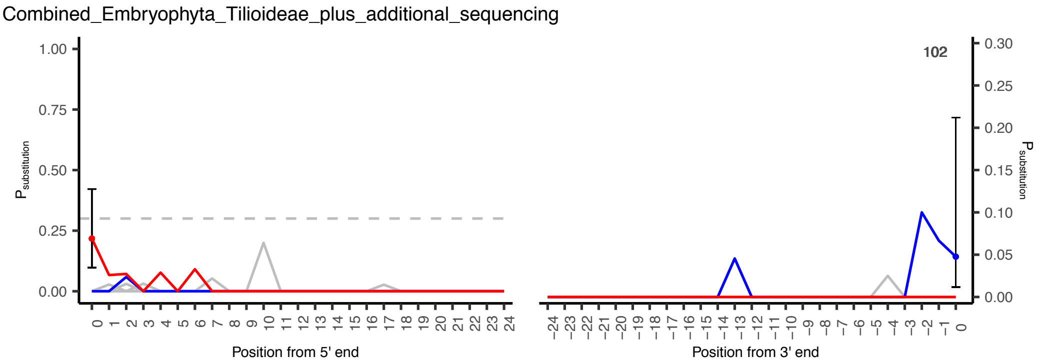


Figure 6.12: MetaDamage profile for Tilioideae (all data). The position-zero C-to-T mismatch rate is 0.217 95% CI [0.098, 0.422].

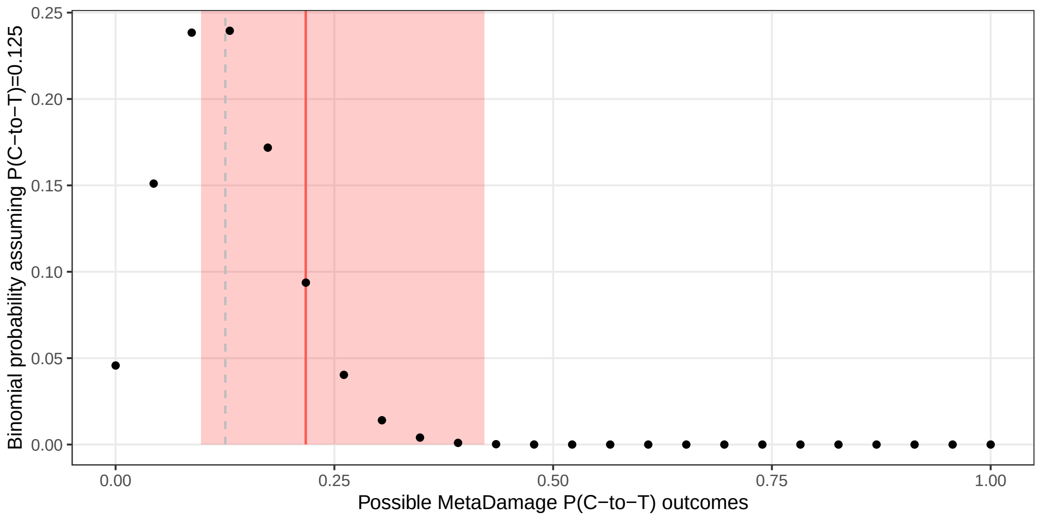


Figure 6.13: Probability distribution of all possible position-zero C-to-T mismatch rates produced by MetaDamage for Tilioideae (all data; 23 trials), assuming the average mismatch rate of 0.125 (dashed line). The probability of zero mismatches is 0.046 and the most probable mismatch rate is 0.130. The observed mismatch rate and its 95% confidence interval are in red.

6.3.3 Juglandaceae

Another taxon associated with warmer climates is Juglandaceae, the walnut family. The only European representative, *Juglans regia*, is today native to western Asia and is only thought to have colonised Great Britain naturally during previous interglacials (Beer *et al.* 2008), although it has been cultivated in Britain for thousands of years and tolerates the present climate (de Rigo *et al.* 2016). Also, recent work on pollen and population genetics by Pollegioni *et al.* (2017) suggests that *J. regia* may have persisted in southern Europe during the last glacial period. The sedaDNA in these cores may represent an early northern colonisation, possibly including transport by humans. Juglandaceae is absent from negative controls and does not contain any over-represented taxa, so appears plausible from a data point of view. Authentication with mapDamage2.0 could clarify the situation.

Unfortunately, very few reads are available for damage analysis. Only a single mismatch was observed in the original data (figure 6.14) and the additional data does not show any signal.

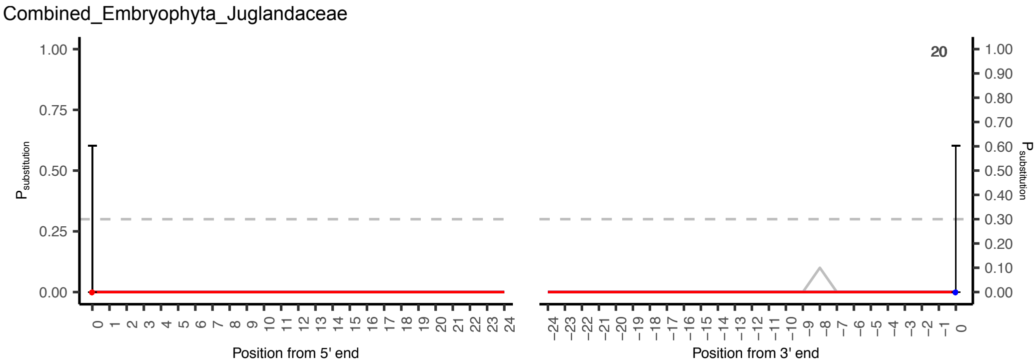


Figure 6.14: MetaDamage profile for Juglandaceae (original data only). Juglandaceae occurred in samples ELF031A_281, ELF045_346 and 252, ELF050_450, and ELF059A_135. The position-zero C-to-T mismatch rate is 0.000 95% CI [0.000, 0.602].

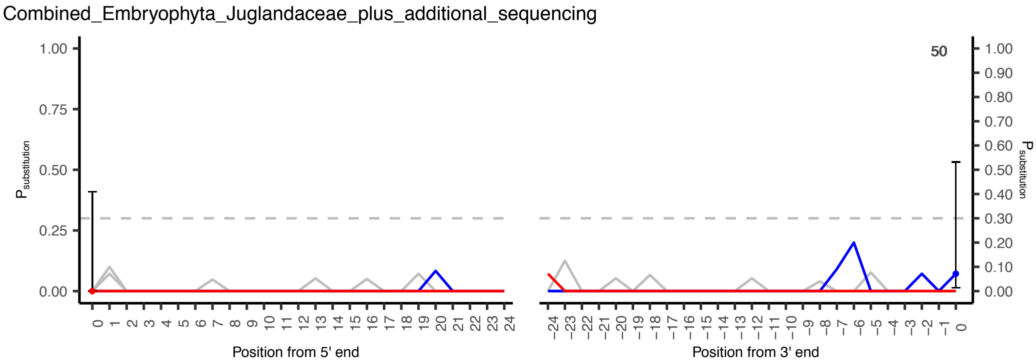


Figure 6.15: MetaDamage profile for Juglandaceae (all data). The position-zero C-to-T mismatch rate is 0.000 95% CI [0.000, 0.410].

The binomial probability analysis shows that the lack of signal can be explained by the lack of data. Assuming a true C-to-T mismatch rate of 0.125, the most probable MetaDamage outcome for Juglandaceae is zero, at 0.447 (figure 6.16). There is insufficient data to confidently assess damage. Therefore, even though the observed result is also zero [0.000, 0.410], there is no reason to suspect a genuine absence of damage signal, and therefore that the Juglandaceae reads are modern. The result is simply inconclusive.

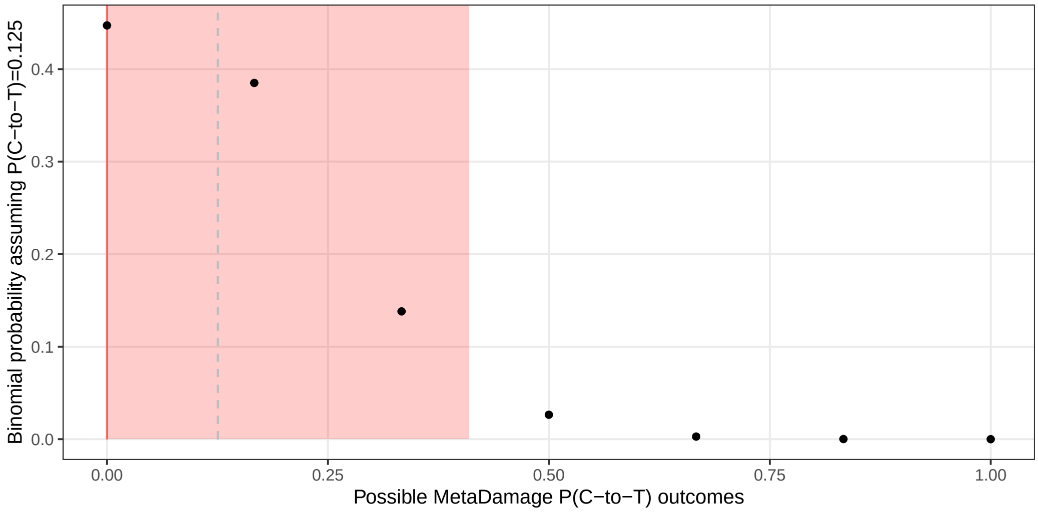


Figure 6.16: Probability distribution of all possible position-zero C-to-T mismatch rates produced by MetaDamage for Juglandaceae (all data; 6 mismatches), assuming the average mismatch rate of 0.125 (dashed line). Zero mismatches has the highest probability: 0.447. The observed mismatch rate and its 95% confidence interval are in red.

6.3.4 Posidonia

Posidonia (Neptune-grasses) is a Mediterranean seagrass that occurs at low frequency in samples with other saltwater taxa. Genuine *Posidonia* in the Southern River samples would suggest significantly warmer sea temperatures and dramatically change the biogeographic history of the genus (Chefaoui *et al.* 2017). It is absent from negative controls and not particularly closely related to any over-represented taxa, being in its own family, so its presence is difficult to explain.

Unfortunately, the MetaDamage result is similarly inconclusive to Juglandaceae. Neither the original nor full dataset shows a damage signal, and even with the additional data there are only 27 reads (figures 6.17 and 6.18). However, what is notable is the high background mismatch rate. Recall with Juglandaceae that the initial 20 reads showed only one mismatch. To see so many mismatches with so few reads here strongly suggests that the *Posidonia* reads do not match their MetaDamage references well. Unlike with Panicoideae, where this could be explained by closely-related over-represented taxa, there are no clear candidates for mis-assignment around *Posidonia*.

Instead, it could potentially be due to variation within *Posidonia*. Average divergence time between the reads and their reference sequences was calculated assuming the general plant substitution rate of approximately 6 x 10^{-9} substitutions per site per year (Wolfe *et al.* 1989). The mean mismatch rate in this case, omitting C-to-T and G-to-A rates from the end five bases to ignore the damage signal, was 0.011 (3 d.p.; n = 580, standard deviation = 0.050). The resulting estimated divergence time was ∼1.9 million years. The most recent study of the evolutionary history of *Posidonia*, Aires *et al.* (2011), report a divergence time between *P. oceanica* and the Southern Hemisphere species of ∼ 68 million years. This would suggest there is more than enough variation among *Posidonia* species to explain the high background mismatch rate, particularly as Aires *et al.* (2011) also found that the substitution rate for *Posidonia* was actually particularly low (perhaps due to the clonal nature and extended generation times of *Posidonia* and other seagrasses), making the divergence time between reads and references an underestimate.

However, of the 5905 *Posidonia* nucleotide records in the GenBank database used in this thesis (downloaded 2019-09-15), only 1.25% are from Southern Hemisphere species. The probability of *P. oceanica* reads receiving first BLAST hits from their distant southern relatives, with resulting mismatch, would actually be very small. The variation within *Posidonia* is unlikely to explain the divergence between reads and references seen here. Instead, perhaps the reads do not derive from *Posidonia*, but were mis-assigned from a less well-represented relative. The co-occurrence of *Posidonia* with other taxa was examined as for Panicoideae, but the highest results do not contain any close relatives; 18 reads are probably too few for a reliable result. However, one good candidate may be *Ruppia*: this is the closest relative to *Posidonia* represented in the data (Janssens *et al.* 2020) and is found in the North Sea today, but is less well-represented in the GenBank database than *Posidonia* with only around 600 records. The unlikeliness of *Posidonia* and the divergent BLAST hits revealed by MetaDamage suggests that PIA may have assigned these reads incorrectly.

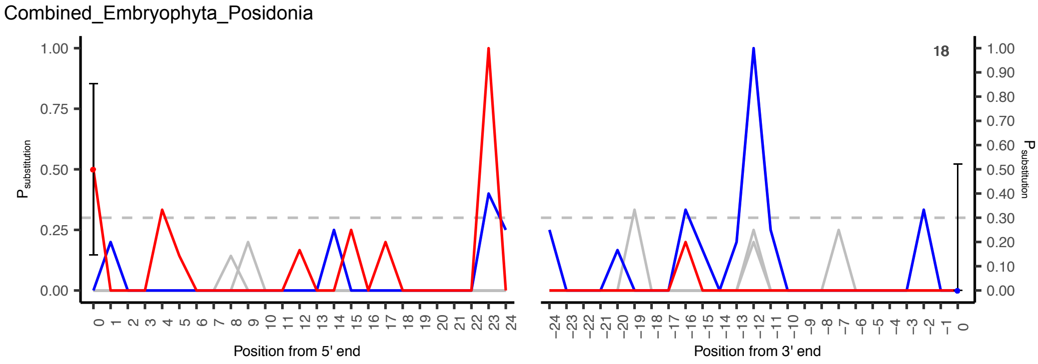


Figure 6.17: MetaDamage profile for *Posidonia* (original data only). *Posidonia* occurred in samples ELF032A_177; ELF034A_126; ELF039_460, 355, 321, and 250; ELF045_145; ELF051_292; ELF059_280 and 270; and ELF059A_250, 190, and 135. The position-zero C-to-T mismatch rate is 0.500 95% CI [0.147, 0.853].

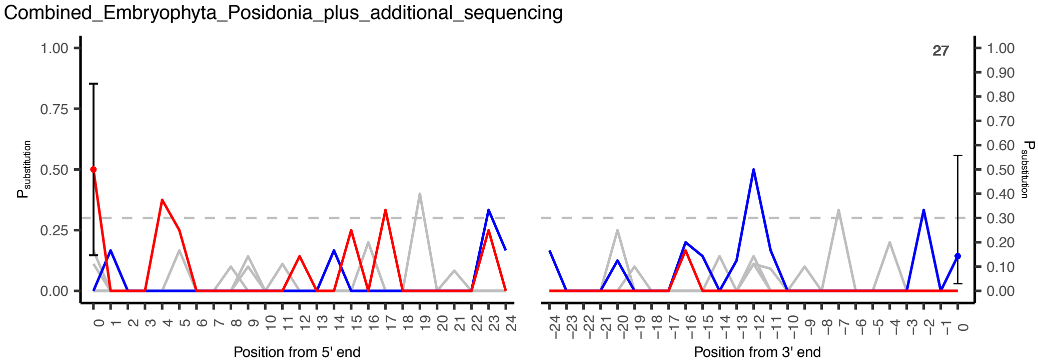


Figure 6.18: MetaDamage profile for *Posidonia* (all data). The position-zero C-to-T mismatch rate is also 0.500 95% CI [0.147, 0.853].

Despite the MetaDamage results suggesting taxonomic information, the binomial probability analysis confirms that there is insufficient data for a reliable estimate of damage. The most probable outcome is zero, at 0.585 (figure 6.19), and only five outcomes are possible. There is not enough data for a reliable MetaDamage result. Note that the confidence interval of the observed result (0.500 95% CI [0.147, 0.853]) is actually greater than zero, but this can be attributed to the high background mismatch rate. Neither the damage signal nor the assignment of the *Posidonia* reads are sufficiently secure to inform climatic reconstructions.

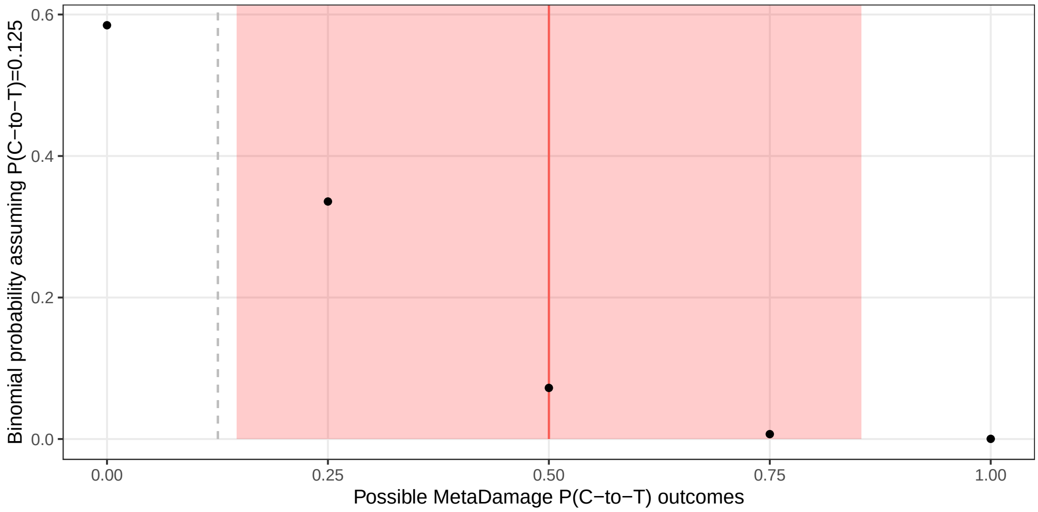


Figure 6.19: Probability distribution of all possible position-zero C-to-T mismatch rates produced by MetaDamage for *Posidonia* (all data; 4 trials), assuming the average mismatch rate of 0.125 (dashed line). Zero mismatches has the highest probability: 0.585. The observed mismatch rate and its 95% confidence interval are in red.

6.3.5 Triticinae

Triticinae, the wheat subtribe, includes data from at least one British native (*Thinopyrum*), but most taxa originate from southern Europe or Asia. Triticinae is highlighted here not as an indicator of climate, but of human activity, wheat being a key European crop (Tresset and Vigne 2011). Two potentially domesticated species are represented in the data: *Triticum monoccum* (einkorn wheat) and *T. turgidum* (durum wheat). However, various Triticinae taxa also regularly appeared in negative controls. Wheat and its relatives are also particularly difficult to assign because of the over-representation of wheat sequences in databases and their history of hybridising into polyploids, which results in new species containing genomes from multiple donor species. The wheat genomes are also particularly rich in transposable elements, comprising >85% of the *T. aestivum* (bread wheat) genome (Wicker *et al.* 2018), which can be shared across diverse taxa. Therefore, authentication by damage signal is particularly important for Triticinae.

The MetaDamage profiles show a high background mismatch rate and no clear damage signal. The original data contained only 11 reads (figure 6.20), yet like *Posidonia*, shows many mismatches. The additional data increases the read count to a reasonable 228, yet there is still no signal. The high background mismatch rate is maintained. As

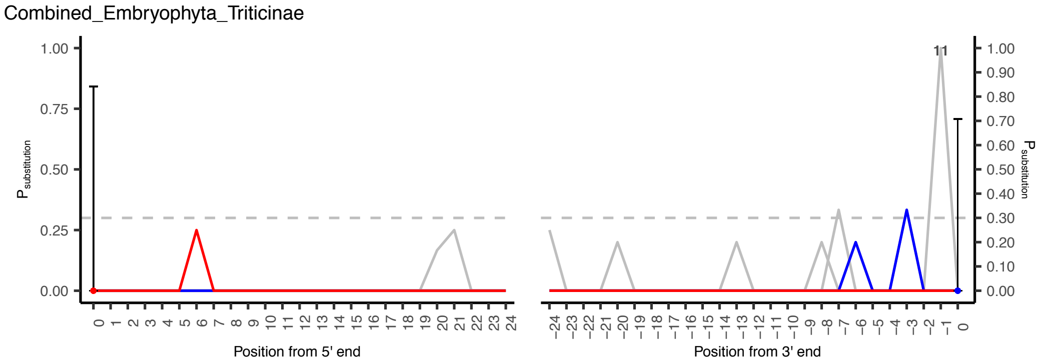


Figure 6.20: MetaDamage profile for Triticinae (original data only). The position-zero C-to-T mismatch rate is 0.000 95% CI [0.000, 0.842].

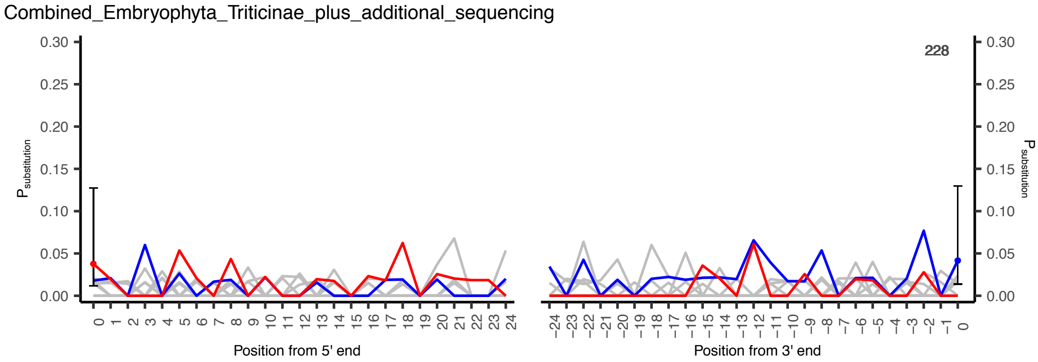


Figure 6.21: MetaDamage profile for Triticinae (all data). The position-zero C-to-T mismatch rate is 0.038 95% CI [0.012, 0.127].

with Panicoideae, the background rate may reflect the loose approach that MetaDamage takes to pairing reads with reference sequences; perhaps many non-wheat Triticinae reads were paired with over-represented wheat references. We can again estimate the typical divergence time between reads and references using the general plant substitution rate of approximately 6 x 10^{-9} substitutions per site per year (Wolfe *et al.* 1989). The mean mismatch rate for Triticinae, excluding the end five bases for C-to-T and G-to-A despite the lack of visible damage signal, came to 0.007 (3 d.p; n = 580, standard deviation = 0.012). This places the divergence time at ~1 million years. The estimated age of Triticinae is imprecise but greater, at approximately 23.8-5.3 million years (Feldman and Levy 2015). Therefore, the evolutionary distance between members of Triticinae can more than explain the background mismatch rate seen in the Triticinae data. While this suggests that the high background does not indicate a problem with assignment, as PIA is designed to handle over-representation, it will still complicate the probability analysis.

The data quantity for Triticinae across all sequences was reasonable. Accordingly, the probability of zero mismatches for Triticinae is only 0.001 (figure 6.22). If the Triticinae reads had a mismatch rate of 0.125, MetaDamage is highly likely to have found something, with the most probable outcome being 0.113. The observed outcome, 0.038 95% CI [0.012, 0.127], is consistent with an average damage level. However, this result may be misleading, as it does not consider the high background mismatch rate.

If we extend our interpretation beyond position zero, the damage signal from Triticinae is unconvincing. The presence of Triticinae taxa in negative controls suggests a possible modern contamination issue. There is not enough evidence to demonstrate the presence of wheat in the Southern River.

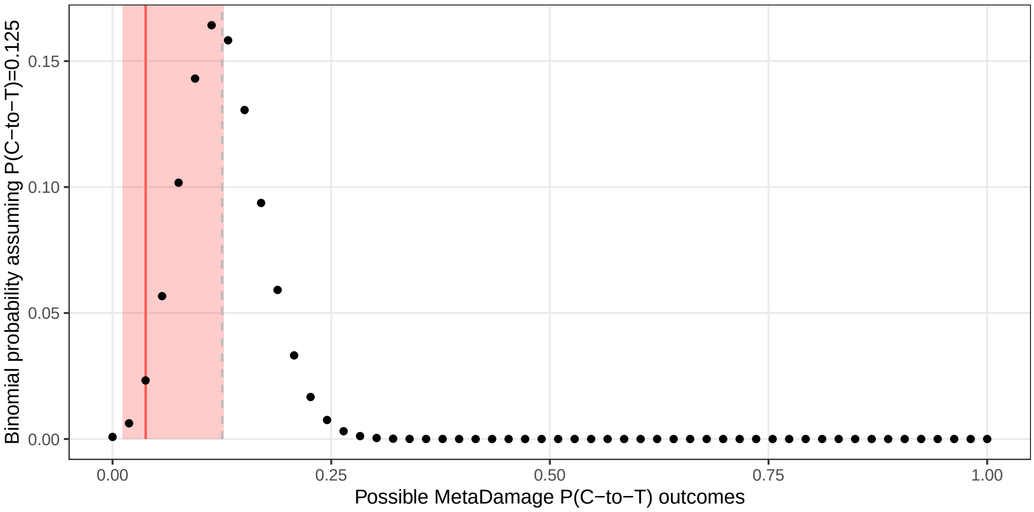


Figure 6.22: Probability distribution of all possible position-zero C-to-T mismatch rates produced by MetaDamage for Triticinae (all data; 53 trials), assuming the average mismatch rate of 0.125 (dashed line). The probability of zero mismatches is 0.001 and the most probable mismatch rate is 0.113. The observed mismatch rate and its 95% confidence interval are in red.

6.3.6 Other cereals

Wheat and its close relatives are difficult to interpret. Fortunately, several other cereal taxa appeared in these samples, including other early European domesticates:

- Aveninae (oat subtribe)
- *Avena* (oats)
- Triticodae
- Triticeae (wheat tribe)
- Hordeinae (barley subtribe)
- *Hordeum* (barleys)
- *Elymus* (couch grasses)
- *Secale* (ryes)
- *Hordeum vulgare* (cultivated barley)
- *Secale cereale* (cultivated rye)

Avena, *Hordeum*, and *Elymus* are native to Great Britain, and all others to Europe apart from the cultivated species. Note that reads derived from Triticinae may potentially be included if they were assigned to the higher taxa Triticeae or Triticodae.

Figures 6.23 and 6.24 show the MetaDamage profiles for these cereals. Similar to Triticinae, the original data has very few reads but already shows a high background mismatch rate, with no particular damage signal. Adding the additional data produces potential upticks, but their wide confidence intervals overlap the persistently high background rate. Some of this background mismatch may be explained by first BLAST hits of less- to more-represented cereals, such as reads from *Secale* being matched with the over-represented *Hordeum*. These taxa also have large genomes (estimated C-values of 6.720-11.364 compared to 0.325 for *Arabidopsis thaliana*; appendix D) and a high transposable element content (Liu *et al.* 2019, Bauer *et al.* 2017), particularly *Hordeum* (Wicker *et al.* 2017). These transposable elements can be shared with a wide range of other plants. Excluding the end 5 bases to avoid age-related mismatches, the mean mismatch rate was 0.006 (3 d.p.; n = 580, standard deviation = 0.014). With the general plant substitution rate (~6 x 10^{-9} substitutions per site per year; Wolfe *et al.* 1989), this estimates mean divergence time at ~930,000 years. Considering this could potentially include reads from organisms across Triticeae being matched to something like *Hordeum*, 960,000 years of divergence is very possible. The high background mismatch rate is inconvenient, but does not cast doubt on the assignment of these reads.

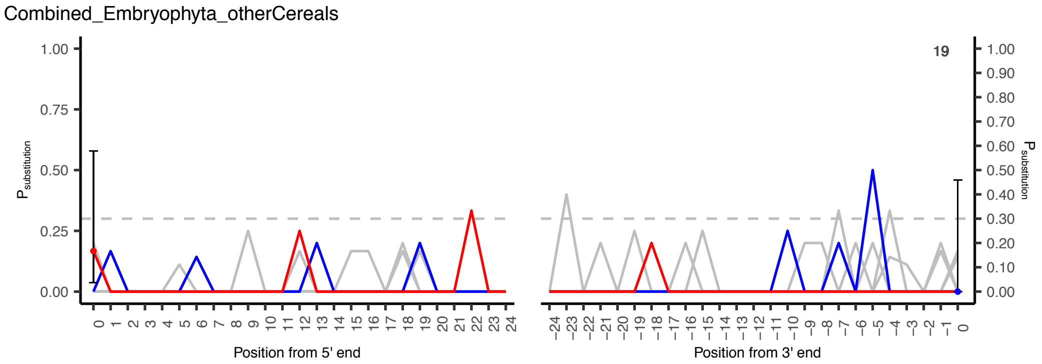


Figure 6.23: MetaDamage profile for other cereals (original data only). The position-zero C-to-T mismatch rate is 0.167 95% CI [0.037, 0.579].

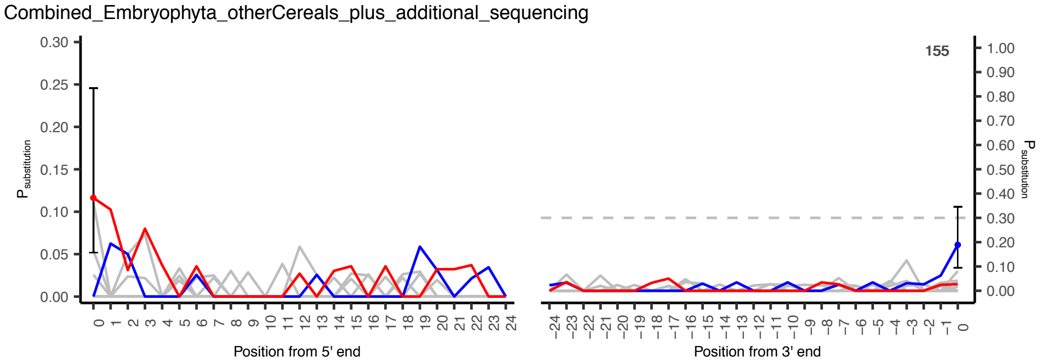


Figure 6.24: MetaDamage profile for other cereals (all data). The position-zero C-to-T mismatch rate is also 0.167 95% CI [0.037, 0.579].

The binomial probability test suggests that the 155 reads in the combined analysis are sufficient to find a damage signal. Figure 6.25 shows only a 0.003 probability of observing no mismatches, and the most probable outcome is close to the assumed mismatch rate at 0.116. The observed MetaDamage outcome of 0.167 95% CI [0.037, 0.579] is consistent with the average damage signal, but the width of the confidence intervals and degree of background variation in figure 6.24 mean that the signal is not strong. While we can interpret some damage signal in the other cereals, it is not enough to exclude the possibility of some contamination. Nevertheless, the result is more promising than for Triticinae.

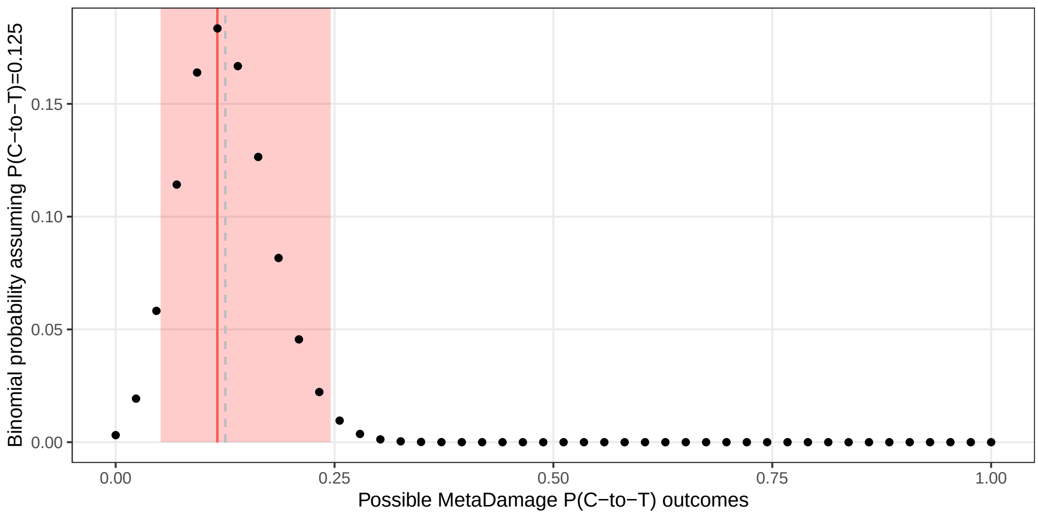


Figure 6.25: Probability distribution of all possible position-zero C-to-T mismatch rates produced by MetaDamage for other cereals (all data; 43 trials), assuming the average mismatch rate of 0.125 (dashed line). The probability of zero mismatches is 0.003 and the most probable mismatch rate is 0.116. The observed mismatch rate and its 95% confidence interval are in red.

6.4 Conclusion

Pooling additional data from across sediment samples, including from other cores and deeper sequencing, has allowed damage-based authentication of several individual taxa. Furthermore, the use of a new binomial probability analysis has suggested other taxa where the apparent lack of damage signal could be explained by low data quantity. This chapter began by demonstrating the methods on two plausible tree taxa, Fagaceae and *Corylus*, both of which performed well. However, the focus of the chapter was on unexpected taxa with implications for changing our interpretation of climate or human activity in this region of Doggerland.

Panicoideae, Tilioideae, Juglandaceae, and *Posidonia* are mesophilic plants, adapted to moderate temperatures. Tilioideae would indicate a similar climate to today, which may prove useful when samples are dated (Chapter 7: Discussion). Panicoideae, Juglandaceae, and *Posidonia* are not native to Great Britain and may indicate warmer temperatures in the Southern River region than present, which could also link dated samples to warmer periods or even suggest a greater degree of warming than expected. Panicoideae showed a very strong damage signal, suggesting that the reads are genuinely ancient despite an uneasy presence in negative controls, but also high co-occurrence with *Phragmites*, from which reads may have been mis-assigned. The uncertain assignment of these reads casts doubt on the use of Panicoideae as an indicator of warm climate. More promisingly, Tilioideae also passed age authentication, although can inform only a few samples due to its limited appearances. Juglandaceae returned too few reads to assess damage, but its uncomplicated assignment and lack of presence in negative controls suggest that the Juglandaceae reads may be genuine. They should be taken with caution, but have not been excluded by the damage analysis. Finally, MetaDamage was also inconclusive for *Posidonia* due to limited data, but the questionable assignment and radical biogeographic implications for this taxon make it less acceptable than Juglandaceae. The *Posidonia* reads may be better explained as mis-assigned *Ruppia*, a temperate seagrass, and are therefore not sufficiently persuasive to suggest a Mediterranean climate in this region of Doggerland.

The remaining taxa, Triticinae and other cereals, contain species likely to have been important for early Neolithic activity: wheat, barley, rye, and oats. These and their relatives did occur in negative controls and have features that would complicate taxonomic assignment. Triticinae, the wheat subtribe, did not produce a clear damage signal despite what was probably sufficient data. However, the other cereals did show a possible signal. Samples containing these taxa should be prioritised for comparison with other environmental proxies when data becomes available. It would also be particularly interesting to find supporting evidence of Juglandaceae

Calculating the binomial probability of possible MetaDamage outcomes proved useful for quantifying when a taxon had too few reads to detect the average Embryophyta signal, and for comparing signals from individual taxa to that average. It demonstrated that the negative Juglandaceae result was not positive evidence of modern contamination, for example. However, the probability analysis was less useful for taxa whose MetaDamage profiles showed a high background mismatch rate. This was due to its exclusive focus on position zero, and only at the 5' end. Chapter 5 (Authentication results) demonstrated that signal strength as manually interpreted from the complete plots cannot be predicted from position-zero values alone: strong signals showed a well-defined range of values, but this also included occasional values from plots with weak or absent signals. On average, position zero will show the maximum C-to-T mismatch rate, but in samples with a less-than-perfect signal it is advisable to also consider the full curve, the curves of other mismatch types, and the confidence intervals. The unexpected taxa considered in this chapter rarely have enough data for smooth curves, yet it is for these high-variation curves that the probability analysis is most valuable. One solution may be to include the background rate as a variable, altering the baseline from zero to the background rate where necessary. Quantitatively evaluating outcomes is a new direction for MetaDamage and would benefit from further development.

However, a high background mismatch rate can actually be informative even if the damage signal is obscured. By suggesting unusually low similarity between reads and their first BLAST hits, it can identify taxa which may be attracted to over-represented taxa in the database, which may therefore be at greater risk of mis-assignment. PIA can still be expected to assign a small minority of reads incorrectly, but the patterns behind these errors are not yet well understood. MetaDamage may actually be a useful tool to investigate this further.

Chapter 7

Discussion

7.1 Addressing the project aims

This section discusses progress towards the following aims:

1. Using sedaDNA to reconstruct the environments sampled by these sediment cores.
2. Identifying and evaluating key taxa indicating an unexpected climate or human activity.
3. Demonstrating the potential of a new suite of sedaDNA analysis methods.

7.1.1 Using sedaDNA to reconstruct the environments sampled by these sediment cores

Environmental reconstruction from sedaDNA

The primary aim of this project was to produce an environmental reconstruction of the Southern River. Figure 7.1 is a reminder of the broad taxonomic results from the sediment samples in this project. While there are no obvious spatial trends, it illustrates the range of different environments inferred from these cores. ELF022 was uninformative for the Dogger Bank, but ELF027 revealed a coastal environment with some freshwater influence. This turned out to be shared with most cores in the Southern River transect.

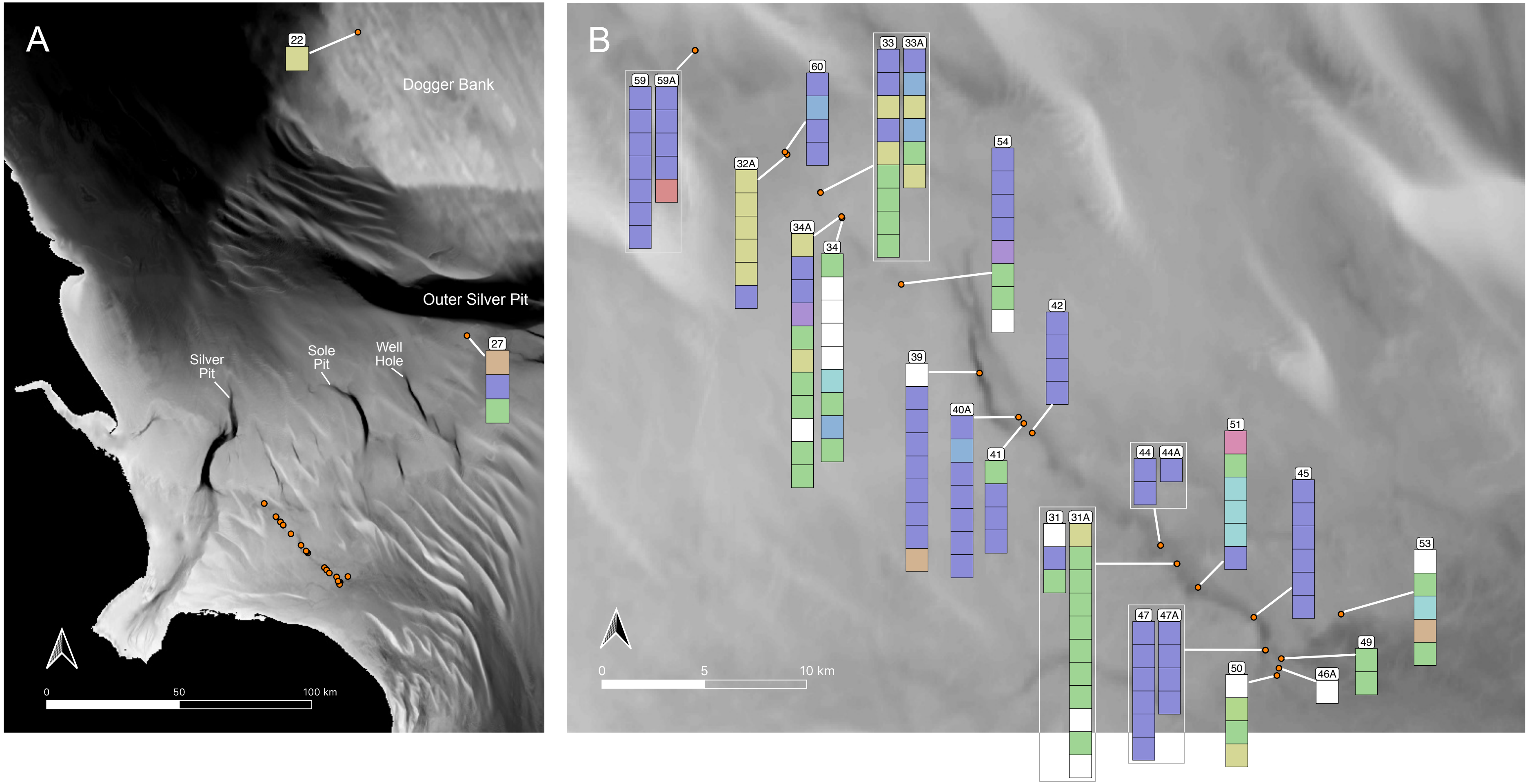


Figure 7.1: Map of cores with summarised sedaDNA results reproduced from Chapter 4. Cores from the same locations are boxed. Each block represents one sample. Depth between samples is omitted. Blocks are coloured according to most frequent Embryophyta ecological category. See figure 7.2 for colour key.

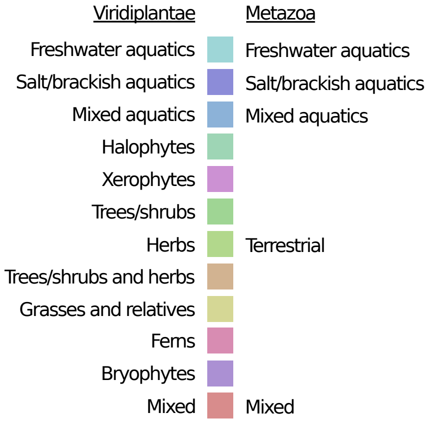


Figure 7.2: Key to ecological categories.

The transect runs from about 20 km northwest of the visible palaeochannel to its mouth, with the addition of ELF053 a short distance along the headland. Approximately half of samples presented a mostly saltwater aquatic plant profile dominated by the seagrass *Zostera*, with occasional *Ruppia* (tasselweeds) and *Phragmites* specifying brackish rather than fully marine conditions. In particular, the *Phragmites* suggest that this may represent the post-inundation brackish reed vegetation, including many *P. australis* microfossils, reported in the Doggerland core of Wolters *et al.* (2010). However, all of these also contained a small signal from terrestrial plants, strongly suggesting that the sedaDNA was not all immediately local. We could expect river and estuarine sediments to contain material washed in from the surrounding catchment, and coastal sediments from runoff along the shore, as noted by Geary *et al.* (2017) in their study of conventional proxies from a core in a nearby palaeochannel. However, the survival time of environmental DNA in freshwater environments is not well characterised and may be limited to days or weeks (Barnes and Turner 2016), so the terrestrial signal in these otherwise brackish samples may still be relatively local. The terrestrial signal usually comprised floodplain woodland taxa, such as Salicaceae and *Alnus*, and wetland Poaceae taxa such as Arundinoideae (reeds). These would be very plausible in a riparian environment. Also typical but at lower frequency were trees of drier ground suggesting woodland in the wider environment: *Quercus*, *Prunus*, *Ulmus*, Betulaceae including *Betula* and *Corylus*, and *Tilia*. These taxa are supported by previous studies based on conventional proxies (Wolters *et al.* 2010, Geary *et al.* 2017, Brown *et al.* 2018). Finally, some *Zostera*-dominated samples with more data also contained a small signal from halophytes or xerophytes indicative of coastal environments, such as Plumbaginaceae (the leadwort family) and Chenopodiaceae (the goosefoot family). This is again congruent with previous studies: halophytes were also found in marine-associated strata by Wolters *et al.* (2010) in their eastern Doggerland core.

Other samples were instead dominated by grasses. Most grass reads were assigned to higher taxa, such as Poaceae, but the general difficulty of assigning grasses (due to the presence of many over-represented crop species) and the presence of wetland taxa suggested that these could be interpreted as salt marsh or reed beds. These two environments were also recorded as intermediate phases between terrestrial and marine in other studies (Wolters *et al.* 2010, Geary *et al.* 2017, Brown *et al.* 2018). The range of taxa was usually very similar to the smaller wetland grass signal found in most brackish samples, suggesting different samples of the same broad environment.

The third main group of samples were primarily composed of the floodplain woodland taxa also noted in most brackish samples. These more terrestrial samples typically returned less data than the brackish or wetland samples, perhaps because terrestrial sediment did not benefit from the effect of salt on DNA preservation (Lindahl and Nyberg 1972, Kistler *et al.* 2017). Lack of data was a frequent limitation in this project. However, there were exceptions. ELF049 returned a surprisingly large number of reads for a Salicaceae-type profile. Uniquely, it also contained several lower-frequency taxa that together suggest a cool climate. These include *Picea* (spruces), Ericaceae (the heather family), and Cyperaceae (sedges), which were also found in the Wolters *et al.* (2010) core before Holocene warming. A few samples from other cores also presented a strong freshwater signal; this too was typically composed of taxa found in most samples, but with proportionally more biogenomic mass. These included *Myriophyllum* (water-milfoils), *Ceratophyllum* (coontails), and *Typha* (bulrushes). There is overlap between the freshwater and floodplain-woodland-type samples. Wolters *et al.* (2010) interpreted both together as a willow carr system which may also be appropriate here.

The validity of these interpretations is supported by the use of minimally biased shotgun sequencing, which brings no prior assumptions of which taxa should be present, and taxonomic assignment by PIA, which has an estimated accuracy of 96% when the source organism is not represented in the database (Chapter 3). The plant sequences for most samples also showed an age-associated damage signal; the remainder were inconclusive due to insufficient data, but the overall positive result from all Embryophyta sequences demonstrates a lack of significant modern contamination among plants (Chapter 5, section 1). Finally, 90% of applicable cores showed at least some evidence of stratification, demonstrating that the movement of DNA out of its sedimentary context was also not a significant issue (Chapter 5, section 2).

The plant sedaDNA therefore provided a rich source of information for environmental reconstruction. The broad results were consistent with previous research and could be validated using these new methods. Unfortunately, results from animal DNA were less useful. The damage signal from pooled Metazoa reads was small and obscured by a high background mismatch rate; this may be explained by modern human contamination and mis-assignment of reads to particularly over-represented animal taxa. Other analyses were limited by the small amount of data: the typical Metazoa read count per sample was <20, even for samples with thousands of Embryophyta reads. This is despite most samples actually having more reads assigned to animals than plants towards the beginning of the analysis pipeline, but the vast majority of this animal data was assigned to Primates and is therefore highly likely to be modern human contamination (Chapter 6: Authentication results). Consequently, the animal data was rarely informative for ecological reconstruction and did not contribute to stratification analysis.

Environmental changes over time

The presence of terrestrial and freshwater samples also provided evidence for environmental change in several cores. A key aim of this project was to search for the effects of marine inundation: it appears to have been captured in ELF033, ELF033A, ELF034A, ELF054, and ELF031A. We see change up the cores from typical terrestrial or freshwater samples to brackish, often with intervening salt marsh or reed bed, again mirroring previous work (Wolters *et al.* 2010, Geary *et al.* 2017, Brown *et al.* 2018).

To expand on this, preliminary radiocarbon (Martin Bates and Derek Hamilton) and OSL (Tim Kinnaird) dates can now put a selection of samples in the context of time. We can now infer environmental changes over broad time periods and across approximately contemporary samples from different cores. The dated analysis involves seventeen of the ELF020-60 cores detailed in this thesis, eight ELF001-20 cores analysed by Roselyn Ware, and also preliminary results from deep sequencing. Figure 7.6 (page 379) summarises taxonomic results from the 26 dated cores from the Late Pleistocene to Late Holocene. Figures 2.1 (page 26) and 7.5 (page 378) provide spatial context for the cores, while figure 7.4 (page 377), reproduced from Platt *et al.* 2017, again provides climatic context via reconstructed Greenland air temperatures through most of this time period.

The cores are presented by transect to allow some interpretation of environmental changes across the landscape. Figure 2.1 maps the two main transects: cores ELF001-20 running from west to east across the head of the Southern River system, and cores ELF021-60 from the mouth to the head of the Southern River system proper, with some cores in adjacent terrestrial areas (note the reversed order compared to elsewhere in this thesis). The two exceptions are ELF027, ∼100 km northeast on the edge of the Outer Silver Pit (omitted from the figure); and ELF019, which was taken in the west-east transect but is located towards the mouth of the river channel, so is listed with the second transect.

Relevant cores are also plotted by depth in figure 7.5. As with figure 4.148 in Chapter 4 (Taxonomic results), the top sample for each core is plotted at the seabed surface depth. The ELF001-20 cores show the same range of elevations as ELF0021-60, including ELF019 inside the palaeochannel between ELF031A and ELF039. The new cores support the trend for salt/brackish-dominated samples further up the cores, as would be expected in an inundation sequence, but there are no other clear patterns.

Finally, figure 7.6 collapses samples from each core into a dated sequence corresponding to eleven broad time periods. The sedaDNA samples do not fully overlap with dated samples for most cores, so the ages should be taken as a guide only. The aim is to present a broad picture of the landscape over time rather than to anchor specific taxa to specific dates. Furthermore, radiocarbon and OSL dates frequently gave conflicting results (cores marked *); carbon isotope ratios and bleaching of minerals can be influenced by different aspects of the depositional environment (Lee *et al.* 2011). However, the similar sets of taxa in samples of similar dates suggest that the dates broadly align.

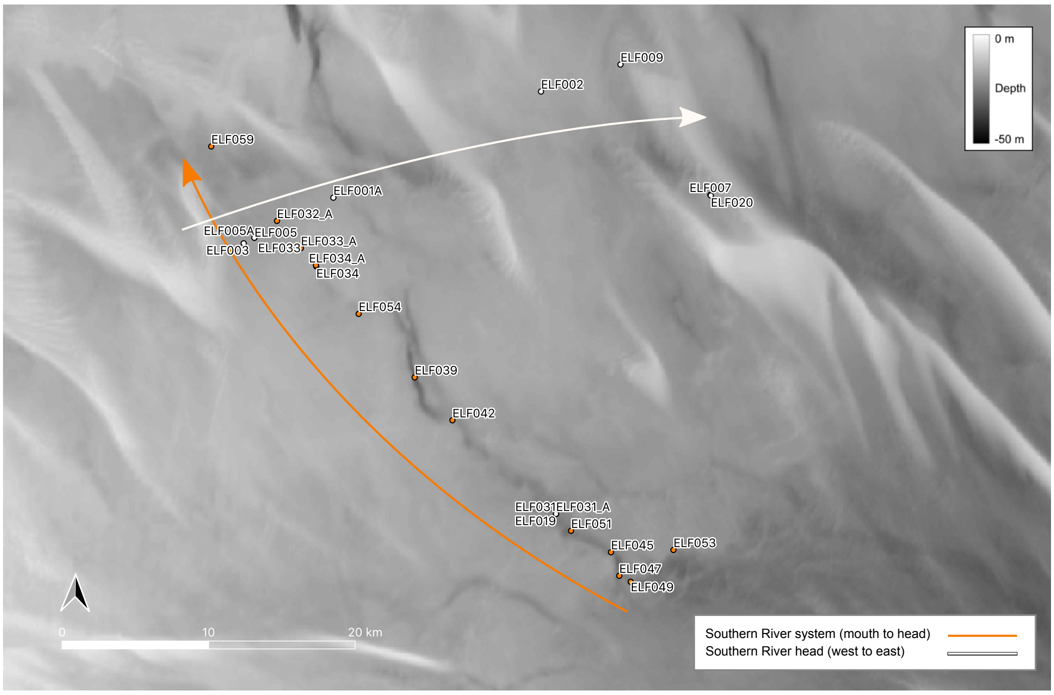


Figure 7.3: Bathymetry map showing core locations for 25 of the 26 dated cores. ELF027 is further north on the edge of the Outer Silver Pit (Chapter 2, figure 2.2). Darker grey indicates greater depth. Arrows indicate the order of samples in figure 7.6 along the two transects.

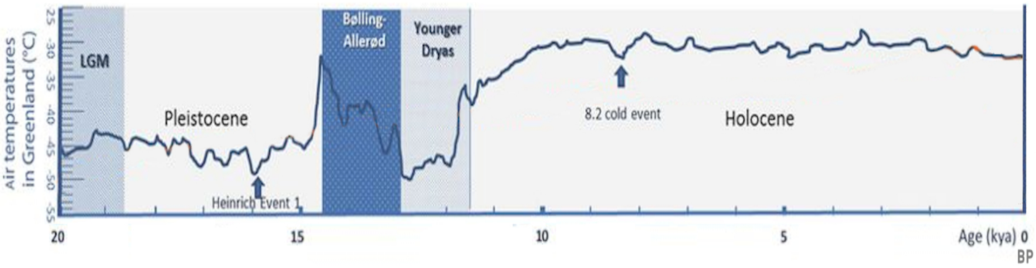


Figure 7.4: "Evolution of temperature in the Post-Glacial period according to Groeanland ice cores" by Platt *et al.* (2017), reproduced without alteration under a Creative Commons 4.0 license (https://creativecommons.org/licenses/by/4.0/). Note that the Holocene began around 11.7 ka, at the end of the Younger Dryas stadial, but this boundary is marked around 11.5 ka in this figure.

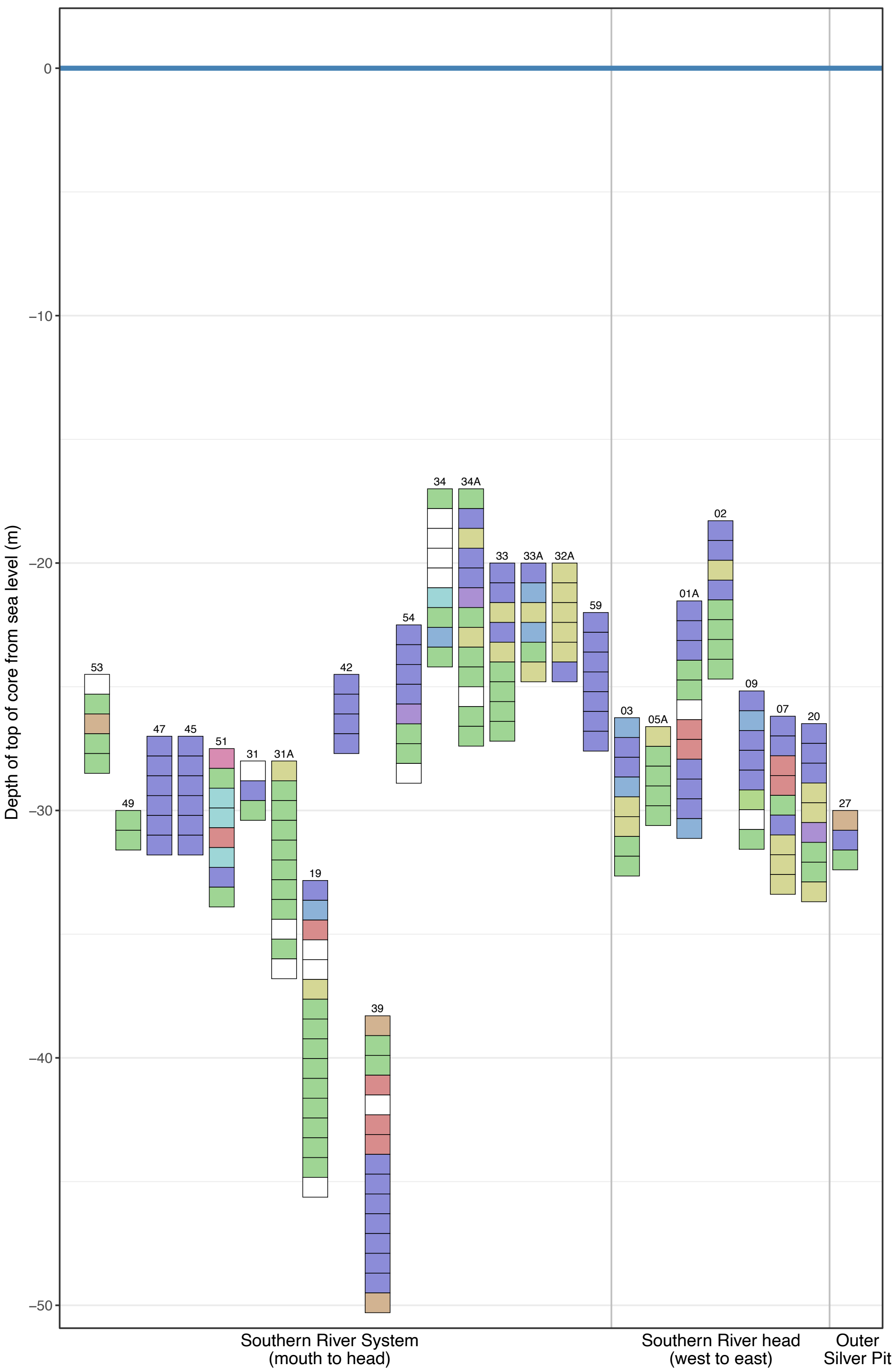


Figure 7.5: Dated cores by starting depth (m) below sea level (blue line at depth=0) with summarised sedaDNA results. Samples within a core are not to scale, but are represented by blocks, with the top block representing the top of the core. Depth between samples is omitted. Blocks are coloured according to most frequent Embryophyta ecological category for that sample; see figure 7.2 for colour key.

	Southern River system (mouth to head)																	Southern River head (west to east)							Outer Silver Pit
Time period (ka)	ELF053	ELF049	ELF047	ELF045	ELF051*	ELF031	ELF031A	ELF019*	ELF042	ELF039	ELF054	ELF034	ELF034A	ELF033	ELF033A	ELF032A	ELF059	ELF003	ELF005A	ELF001A	ELF002	ELF009	ELF007*	ELF020*	ELF027
Late Holocene 6,000-3,000								Willow, *Zostera*	*Zostera* , willow		*Zostera*										*Zostera*				
Mid Holocene II 7,000-6,000								Willow			*Zostera*						*Zostera*			*Zostera*					
Mid Holocene I 8,000-7,000				*Zostera*				Willow		*Zostera*							Grasses, *Zostera*			*Zostera* , willow	Willow	*Zostera*	*Zostera*		
Tsunami & inundation 8,300-8,000													*Zostera* , willow, grasses						Betulaceae, grasses, *Quercus*	*Zostera* , grasses, willow, Betulaceae			Grasses, willow	*Zostera* , grasses	
Early Holocene IV 8,500-8,300													*Zostera*, Betulaceae, *Quercus*					*Zostera*							
Early Holocene III 9,000-8,500			*Zostera* , willow		Ferns	*Zostera* , willow, *Prunus*	Grasses, willow	Willow		*Zostera*	Willlow				Grasses, willow, *Prunus*	Grasses, *Zostera*		Grasses		*Zostera*				Betulaceae, *Quercus*	
Early Holocene II 9,500-9,000			*Zostera* , willow	*Zostera* , willow	Betulaceae	Willow	Willow, grasses	Willow				*Quercus*		Grasses, *Zostera* , Betulaceae, *Quercus*											
Early Holocene I 10,500-9,500					Willow, freshwater		Willow, *Prunus*				Willlow														*Zostera* , grasses, Betulaceae
Pleistocene/Holocene boundary 11,700-10,500				*Zostera* , willow	Ferns			Willow					Willow, *Betula*								Willow				*Zostera* , grasses, Betulaceae
Late Pleistocene II 13,500-11,700					Willow, freshwater					*Zostera* , grasses, willow		Freshwater, willow												Willow	
Late Pleistocene I 40,000-13,500	Willow	Willow, *Carex*										Willow, freshwater						Grasses, Betulaceae, Fagaceae			Willow, *Carex*			Grasses, willow	

Figure 7.6: Taxonomic summaries of dated cores by time period. Cores are ordered by position and cores around the head of the Southern River system are considered separately. Cells can be influenced by multiple samples, so significant taxa are listed instead of dominant ecological category. Includes preliminary resequencing results and initial data from cores 01-20 (analysed by Roselyn Ware). Cores marked * have conflicting radiocarbon and OSL dates, so placement in time periods is less certain.

The oldest samples date to >13,500 BP, in the Late Pleistocene I phase. Samples from cores in the Southern River system all contain the willow group (Salicaceae and its child taxa). While the willow group was mostly interpreted in a riparian context, *Salix* and *Populus* are also associated with boreal forest and tundra (Pividori *et al.* 2016, Anderson *et al.* 1994). In ELF049, these are accompanied by lower-frequency taxa like Ericaceae and Dryadoideae that together suggest a cold climate (Chapter 4: Taxonomic results). The ELF049 samples could therefore date from one of the stadials in this period, and are potentially among the oldest samples studied (figure 7.4). The other Late Pleistocene I samples lack specific cold indicators, and the clear freshwater signal in ELF034 supports a riparian environment, suggesting that these samples are more likely to date from the warmer Bølling or Allerod interstadials. The one core where the willow group is not particularly frequent is ELF003 from the river head transect. This instead notes the first significant appearance of other trees: Betulaceae (most likely *Betula* or *Corylus*) and Fagaceae (most likely *Quercus*). All three genera appear relatively early in pollen records around the southern North Sea (Wolters *et al.* 2010, Preece and Day 1994). *Betula* in particular can indicate cool boreal forest (Pividori *et al.* 2016) and has been recorded relatively consistently since at least 15,000 BP (Brewer *et al.* 2017). *Corylus* and *Quercus* also have limited pollen records from 14,000 BP, becoming more frequent from around 10,000 BP (Brewer *et al.* 2017). The >13,500 BP Betulaceae and *Quercus* reads in ELF003 and ELF005 support these early appearances. Note that arrival dates for mainland Great Britain are typically much later (Brewer *et al.* 2017), but we could expect an earlier signal in these cores because of the direction of colonisation. Many are expected to have colonised from the east (Birks 1989, Brewer *et al.* 2002), through the remnants of Doggerland, and should have arrived at the Southern River valley before the mainland.

The Late Pleistocene II phase is more precisely dated to 13,500-11,700 BP, beginning in the Allerod interstadial and then covering the Younger Dryas cool period (figure 7.4). The four relevant cores are again rich in the willow group, but three also contain aquatics, which suggest the willow may represent floodplain woodland instead of particularly cold environments. ELF039 shows the first saltwater signal (*Zostera*), accompanied by the typical grass signal which can be interpreted as reeds, but this does not necessarily indicate marine inundation. ELF039 was taken from a lake, which may have become brackish before inundation due to tidal movements or a rising water table. Note that the youngest sample from ELF039 in this phase also contains a small well-timed signal from Dryadoideae, interpreted as *Dryas*, the indicator of the Younger Dryas cool period.

Moving up to the Pleistocene/Holocene boundary (11,700-10,050), we still see predominantly the willow group, perhaps equivalent to the willow-carr phase observed by Wolters *et al.* 2010 in their southern North Sea core 10,700-9,550 BP. Similarly, *Betula* and/or *Corylus* in ELF034A and also ELF027, on the edge of the Outer Silver Pit (figure 7.1), may correspond to their birch woodland phase. *Betula* pollen in Britain suggests a similar arrival date on the east coast of 10,400-10,000 BP (Birks

1989). These cores appear to capture early woodlands. There is also further *Zostera* in ELF045 and also ELF027. These two cores are near the mouth of the Southern River valley and on the edge of the Outer Silver Pit, suggesting an approaching coastline in these locations. We would expect some evidence of Holocene warming around this time, so a potential increase in tree cover and sea level is promising.

This pattern continues during the Early Holocene I to IV phases (10,500-8,300 BP), a period of more gradual warming (figure 7.4). Cores in the Southern River system transect show a blend of *Zostera*, the willow group, and the less water-associated trees in Betulaceae, *Quercus*, and *Prunus*. The *Quercus* signals in ELF034 and ELF033 from 9,500-9,000 BP, and in ELF020 from the river head transect from 9,000-8,500 BP, coincide with a steep increase in *Quercus* pollen in Britain from around 9,500 BP (Brewer *et al.* 2017). Evidence of both terrestrial and brackish environments is in broad agreement with the sea level model of Walker *et al.* (2020), which places the Southern River region near the edge of the retreating Dogger Archipelago coastline around 8,200 BP (Chapter 1, figure 1.2).

Samples from 8,300-8,000 BP were given special attention by Europe's Lost Frontiers because of the discovery of deposits from the Storegga tsunami in several cores. This narrow time period sees possible salt marsh in ELF034A, but most sedaDNA data is from the river head transect: all five cores are rich in grasses, which may also represent marsh, but woodland is more frequent than *Zostera*. This terrestrial signal may originate from inland material washed into the channel (ELF001A, ELF005, ELF005A) or into the sea (ELF007, ELF0020) by the tsunami (Gaffney *et al.* 2020). It at least confirms continued Betulaceae and *Quercus* woodland in the region.

The Mid- to Late Holocene periods (8,000-3,000) no longer contain significant tree signals other than the willow group, although a variety of taxa can still be found at low frequency (Chapter 4: Taxonomic results). Instead, most samples show predominantly *Zostera*, with some floodplain woodland and grasses suggesting coastal environments. The persistence of the willow group into even the most recent samples, younger than 6,000 BP in ELF019 and ELF042, supports the survival of a Dogger Littoral after most of the landscape was inundated (Chapter 1, figure 1.2).

This preliminary analysis of dated sedaDNA results by time period has revealed a steady transition from woodland to coastal environments across the study area. The transition has some geographical pattern, starting at the mouth of the Southern River and proceeding up the valley transect, and then into cores from the head transect. The two exceptions are ELF039, which was taken from a lake, and the samples from ELF019 that appear terrestrial into the Mid- to Late Holocene. ELF019 had dating conflicts; the older dates would in fact push those samples further back in time to around 9,000 BP, which would fit with the inundation pattern. This pattern is also generally consistent with the expected climate and sea level. Dates linked samples from different cores over many kilometres into a richer, three-dimensional environmental reconstruction. Further work by Europe's Lost Frontiers will use additional data, such as geology and other environmental proxies, to map common strata across cores and continue building the many individual samples into a coherent narrative.

7.1.2 Identifying and evaluating key taxa, especially "exotics"

The filtering process highlighted possible but unexpected taxa. In Chapter 6 (Mesophilic taxa and human disturbance indicators), those with sufficient data were subject to individual damage-based authentication with MetaDamage and a further binomial probability analysis to compare their estimated damage level to the Embryophyta average. This was supported by additional data from cores ELF001-20 and deep sequencing of selected samples.

Evidence of damage was found for some mesophilic taxa - those suggesting a warmer climate. *Posidonia* was rejected and, although reads assigned to Panicoideae showed a strong damage signal, they may be better interpreted as *Phragmites*. However, Juglandaceae and early Tilioideae appear genuine. Preliminary radiocarbon (Martin Bates and Derek Hamilton) and OSL (Tim Kinnaird) dating support the presence of Juglandaceae and Tilioideae at least as far back as the Early Holocene. Table 7.1 places three samples containing Juglandaceae reads between 8,950-8,900 and 12,000-10,000 BP, although note the conflict between the radiocarbon and OSL estimates. These reads may represent an early northern colonisation, possibly facilitated by humans, and perhaps suggest that temperatures similar to today were reached more quickly in the North Sea region than expected. Perhaps these early *Juglans* were then pushed back into their present-day southern regions during the 8,200 BP cold event.

Table 7.1: Date estimates for a subset of samples containing Juglandaceae reads, from youngest to oldest. Note the two conflicting estimates for ELF031A_281.

Core	Depth	Date type	Date estimate (BP)
ELF031A	281	cal-C14	8,950-8,900
ELF045	252	OSL	11,000-9,000
ELF045	346	OSL	11,000-9,000
ELF031A	281	OSL	12,000-10,000

Several Tilioideae reads were given similar age estimates. Trace amounts of *Tilia* pollen have been found around the southern North Sea from 13,000 BP, but it does not make a significant appearance in Britain until around 8,000 BP (Brewer *et al.* 2017), and even today England is near the edge of its thermal limit (Eaton *et al.* 2016). However, table 7.2 shows Tilioideae reads from these cores dating at least as far as the Pleistocene/Holocene boundary. This supports those very early pollen records, and is also consistent with the colonisation of Great Britain from the east (Birks 1989). Finally, the potential presence of Juglandaceae and Tilioideae reads before 10,000 BP, when temperatures are expected to have still been rising after the Younger Dryas, may suggest faster warming in the North Sea region. Temperatures similar to today may have been reached relatively early compared to what air temperature data suggest (figure 7.4), and this could have implications for human activity.

The authentication of potentially human-associated grasses was less promising. There was no evidence of age-related damage for taxa in Triticinae, the wheat sub-

Table 7.2: Date estimates for a subset of samples containing Tilioideae reads, from youngest to oldest.

Core	Depth	Date type	Date estimate (BP)
ELF042	65	OSL	<5,000
ELF042	350	OSL	6,000-5,000
ELF059	270	OSL	7,500-6,000
ELF039	250	OSL	8,000-7,000
ELF039	321	OSL	9,000-8,000
ELF033	155	cal-C14	<9300
ELF027	160	OSL	12,000-10,000
ELF027	174	OSL	12,000-10,000
ELF039	145	OSL	>15,000-11,000

tribe, although the size and polyploid nature of wheat genomes present a challenge for both PIA and MetaDamage. Reads from other cereal taxa were analysed in a second group that still contained difficult taxa, but did return some signal, suggesting that at least some reads may be genuine. This offers some support for the presence of cereals that were available for human exploitation. However, there is too little data to distinguish between cultivated varieties that may have been imported and their wild relatives native to Great Britain. Overall, the limited data quantity from the potentially mesophilic or human-associated taxa makes drawing conclusions more difficult than from whole samples.

7.1.3 Demonstrating the potential of a new suite of sedaDNA analysis methods

This project has involved the development of several new tools for sedaDNA analysis. The most important for basic ecological reconstruction is PIA. Chapter 3 details the favourable benchmarking and accuracy test results already published. Chapter 4 (Taxonomic results) supported this by examining how PIA performed with a much larger dataset: PIA continued to assign more selectively than the closest alternative tool, MEGAN, and assigned the vast majority of reads to realistic (European) taxa. A very small number of false positives, non-European taxa, were nevertheless produced, and Chapter 5 (Authentication results) began to investigate their origin. Co-occurrences of these taxa with others suggested that mis-assignment to over-represented taxa may be the main source of error, despite a key aim of PIA being to recognise over-assignment. Nevertheless, these mis-assignments were very rare, and the co-occurrence analysis may prove useful for further development of PIA to improve accuracy even further. Chapter 6 (Mesophilic taxa and human disturbance indicators) also revealed that the simple BLAST used in MetaDamage may identify taxa at increased risk of mis-assignment, suggesting a further avenue for investigation.

Another important method for reconstruction, first published in Gaffney *et al.* (2020), is the use of genome size adjustment to replace raw read count with biogenomic mass. By controlling for genome size, this brings the data one step closer to biomass, the ideal proxy. The mean C-values used to represent genome size in this project were not derived from complete or perfectly curated databases, but size adjustment will become more accurate as the databases continue to develop.

The two remaining tools concern authentication of the DNA profiles built from assigned reads and their biogenomic masses. First, age-authentication via damage analysis is a crucial step for excluding modern contamination (Renaud *et al.* 2019) but its use has been limited in sedaDNA. This can be attributed to the prevalence of metabarcoding studies, where the damaged ends of the molecules are not sequenced (Chapter 3: PIA), and the fact that existing methods of damage analysis, such as mapDamage2.0 (Jónsson *et al.* 2013) and PMDtools (Skoglund *et al.* 2014), are designed for single taxa only. This project instead used shotgun sequencing to record full molecules and the new MetaDamage analysis to estimate damage for reads of varied and sometimes uncertain assignment. MetaDamage is in preparation for publication along with thorough testing performed by Rosie Everett, but this thesis demonstrated its successful use on a large and diverse dataset. MetaDamage found no damage signals for negative controls but did find signals for the Embryophyta data in samples with the minimum recommended data quantity (>100-300 reads); negative results for other samples could be explained by the lack of data. This thesis also suggested methods of quantifying the MetaDamage output, particularly via position-zero C-to-T mismatch values. While these did allow insights only possible through quantitative analysis, such as comparing the damage signals from individual taxa to the Embryophyta average in Chapter 6, the methods were limited. Further development of MetaDamage should include more holistic quantification options, such as incorporating the 95% confidence intervals and background mismatch rate that disproportionately affect interpretation when data quantity is low. The small number of reads for most individual taxa in this project emphasises the value of a metagenomic approach to sedaDNA analysis.

Finally, this project also represented the first large-scale application of the stratification analysis first published in Gaffney *et al.* (2020). Evidence for stratification is evidence against the vertical movement of DNA, leaching, that dissociates sequences from their context. Previous sedaDNA studies have demonstrated stratification through the presence or absence of key taxa with known time ranges (e.g. Graham *et al.* 2016, Ficetola *et al.* 2018), but this analysis is the first attempt to quantify differences between adjacent samples regardless of which taxa are involved. This requires no prior assumptions on the taxonomic composition, lending itself to little-known ancient environments. The analysis suggested stratification in at least one direction in 90% of suitable cores. Further comparison with geology and microfossil results from Europe's Lost Frontiers colleagues demonstrated that apparently unstratified samples shared similar environments, so there was no evidence for leaching. The analysis successfully provided sensitive, quantified evidence for stratification in these cores. However, the stratification analysis was complicated by the differing ranks of taxa involved. PIA output typically involves taxa higher than species, and the cumulative counting method that accounts for nested ranks reduced the sensitivity of the stratification analysis for higher taxa. Future development should consider alternative ways to incorporate ranks, as this aspect of PIA cannot be compromised. Fortunately, the stratification analysis could potentially be applied to other quantifiable environmental proxies such as pollen, where nested ranks are less of an issue. The statistical framework may also prove useful for comparing DNA profiles in other contexts, notably of samples from different cores but apparently the same strata. Evidence of different DNA profiles in the same strata could have implications for the taphonomy of sedaDNA, which is not yet well understood.

In conclusion, this thesis has presented the development and interpretation of important new methods that have great potential to improve the accuracy and richness of information obtained from future sedaDNA studies. By maximising the range and complexity of sequences obtained through shotgun sequencing, this project could build ecological reconstructions with minimal taxonomic bias or subjective interpretation of the validity of taxa. Sequences could be age-authenticated via damage signals in the context of samples rather than a handful of isolated taxa for which reference genomes are available, and a similar taxon-independent approach allowed the rejection of leaching across all sample pairs with sufficient data without reliance on well-studied taxa. All have demonstrated success on this dataset.

7.2 Future directions

This sedaDNA reconstruction of the Southern River provides valuable insight into the composition and behaviour of an ancient environment undergoing marine inundation. Future work in the Europe's Lost Frontiers project will combine these results with those from other environmental proxies and other sediment cores to build a more detailed picture, including further investigation of unexpected taxa such as cereals. There will be close comparison and evaluation of the different proxies. More complete and precise time slices will be constructed using OSL and radiocarbon dating. Finally, this data will be incorporated into models of sea level, ecology, and human activity. The wider project will substantially improve our knowledge of Doggerland and reiterate its importance for the archaeology of Europe.

A parallel investigation will continue with the sedaDNA data from these cores. The enormous quantity of Metazoa reads assigned to Primates were assumed to be human contamination and were excluded for this thesis. However, it is possible that some derive from ancient humans. Future work will use methods optimised for human DNA, such as PMDtools (Skoglund *et al.* 2014) and PALEOMIX (Schubert *et al.* 2014), to assess damage and search for ancient genotypes in any potentially ancient reads. This project may not have returned conclusive results of human-associated taxa, but we have not yet searched for humans themselves.

We can also expect exciting future developments for the sedaDNA analysis tools presented here. All are relatively new and will continue to improve as they are applied to different datasets by the wider sedaDNA community. Furthermore, taxonomic assignment of shotgun sequences will improve as GenBank continues to grow (Sayers *et al.* 2021); nearly 490,000,000 sequences have been added between the release used in this thesis (August 2019) and the latest at the time of writing (February 2021; National Center for Biotechnology Information (US) 2021). There is scope to re-analyse this data in the future, when more organisms have been sequenced with greater coverage, increasing representation and also potentially making regions of the database less uneven, both of which will increase the number of reads that PIA can confidently assign. This thesis has demonstrated the potential of these tools and suggested directions for their improvement. Future work will bring more objective, quantifiable analysis to sedaDNA data from beyond Doggerland.

Bibliography

Aires, T., N. Marbà, R. Cunha, G. Kendrick, D. Walker, E. Serrão, C. Duarte and S. Arnaud-Haond (2011). 'Evolutionary history of the seagrass genus *Posidonia*'. *Marine Ecology Progress Series* **421**, 117–130.

Allaby, R. G., R. L. Ware, B. Cribdon and R. Everett (2021). 'Revealing the Hidden Landscapes of Doggerland with SedaDNA'. In: *Lost Frontiers and Drowned Landscapes in Britain and Beyond*. Ed. by G. Bailey and V. Gaffney. Vol. **1**. Society of Antiquaries of London, p. 5.

Amkreutz, L. and M. Spithoven (2019). 'Hunting beneath the waves. Bone and antler points from North Sea Doggerland off the Dutch coast'. In: *Working at The Sharp End at Hohen Viecheln*. Ed. by D. Groß, D. Jantzen, H. Lübke and John Meadows. Kiel/Hamburg: Wachholtz Verlag GmbH. Chap. 16, pp. 383–404.

Anderson, P. M., P. J. Bartlein and L. B. Brubaker (1994). 'Late Quaternary History of Tundra Vegetation in Northwestern Alaska'. *Quaternary Research* **41**.3, 306–315.

Appelt, S., L. Fancello, M. Le Bailly, D. Raoult, M. Drancourt and C. Desnues (2014). 'Viruses in a 14th-Century Coprolite'. *Applied and Environmental Microbiology* **80**.9, 2648–2655.

Ardelean, C. F. *et al.* (2020). 'Evidence of human occupation in Mexico around the Last Glacial Maximum'. *Nature* **584**, 87–94.

Armbrecht, L. H. *et al.* (2019). 'Ancient DNA from marine sediments: Precautions and considerations for seafloor coring, sample handling and data generation'. *Earth-Science Reviews* **196**, 102887.

Ball, P. W. and A. A. Reznicek (2008). *Carex Linnaeus, Sp.* URL: `http://www.efloras.org/florataxon.aspx?flora_id=1%5C&taxon_id=105644` (visited on 04/06/2020).

Barnes, M. A. and C. R. Turner (2016). 'The ecology of environmental DNA and implications for conservation genetics'. *Conservation Genetics* **17**.1, 1–17.

Bauer, E. *et al.* (2017). 'Towards a whole-genome sequence for rye (*Secale cereale* L.)' *The Plant Journal* **89**.5, 853–869.

Bayliss, A., F. Healy, A. Whittle and G. Cooney (2011). 'Neolithic narratives: British and Irish enclosures in their timescapes'. In: *Gathering Time: dating the Early Neolithic Enclosures of Southern Britain and Ireland*. Ed. by A. Whittle, F. Healy and A. Bayliss. Oxford: Oxbow Books. Chap. 14, pp. 682–847.

Beer, R., F. Kaiser, K. Schmidt, B. Ammann, G. Carraro, E. Grisa and W. Tinner (2008). 'Vegetation history of the walnut forests in Kyrgyzstan (Central Asia): natural or anthropogenic origin?' *Quaternary Science Reviews* **27**.5, 621–632.

Bell, F. G. (1969). 'The occurrence of southern, steppe and halophyte elements in Weischselian (Last-glacial) floras from southern Britain'. *New Phytologist* **68**.4, 913–922.

Benson, D. A., M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell and E. W. Sayers (2016). 'GenBank'. *Nucleic Acids Research* **45**.D1, D37–D42.

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell and E. W. Sayers (2011). 'GenBank'. *Nucleic Acids Research* **39**.suppl_1, D32–D37.

Betti, L. *et al.* (2020). 'Climate shaped how Neolithic farmers and European hunter-gatherers interacted after a major slowdown from 6,100 BCE to 4,500 BCE'. *Nature Human Behaviour* **4**, 1004–1010.

Birks, H. J. B. (1989). 'Holocene Isochrone Maps and Patterns of Tree-Spreading in the British Isles'. *Journal of Biogeography* **16**.6, 503–540.

Birks, H. J. B. and H. H. Birks (2016). 'How have studies of ancient DNA from sediments contributed to the reconstruction of Quaternary floras?' *New Phytologist* **209**.2, 499–506.

Bliss, L. C. (2000). 'Arctic Tundra and the Polar Desert Biome'. In: *North American Terrestrial Vegetation*. Ed. by M. G. Barbour and W. D. Billings. 2nd ed. Cambridge: Cambridge University Press. Chap. 1, pp. 1–20.

Bollongino, R., O. Nehlich, M. P. Richards, J. Orschiedt, M. G. Thomas, C. Sell, Z. Fajkošová, A. Powell and J. Burger (2013). '2000 Years of Parallel Societies in Stone Age Central Europe'. *Science* **342**.6157, 479–481.

Bon, C., V. Berthonaud, F. Maksud, K. Labadie, J. Poulain, F. Artiguenave, P. Wincker, J.-M. Aury and J.-M. Elalouf (2012). 'Coprolites as a source of information on the genome and diet of the cave hyena'. *Proceedings of the Royal Society B: Biological Sciences* **279**.1739, 2825–2830.

Bondevik, S., S. K. Stormo and G. Skjerdal (2012). 'Green mosses date the Storegga tsunami to the chilliest decades of the 8.2 ka cold event'. *Quaternary Science Reviews* **45**, 1–6.

Bonsall, C., G. Cook, C. Pickard, K. McSweeney and L. Bartosiewicz (2009). 'Dietary trends at the Mesolithic–Neolithic transition in North-west Europe'. In: *Chronology and Evolution within the Mesolithic of North-West Europe: Proceedings of an International Meeting, Brussels, May 30th-June 1st 2007*. Ed. by P. Crombé, M. V. Strydonck, J. Sergant, M. Boudin and M. Bats. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 539–562.

Boulanger-Lapointe, N., E. Lévesque, C. Baittinger and N. M. Schmidt (2016). 'Local variability in growth and reproduction of *Salix arctica* in the High Arctic'. *Polar Research* **35**.1, 24126–24136.

Brewer, S., R. Cheddadi, J. L. de Beaulieu and M. Reille (2002). 'The spread of deciduous *Quercus* throughout Europe since the last glacial period'. *Forest Ecology and Management* **156**.1, 27–48.

Brewer, S. *et al.* (2017). 'Late-glacial and Holocene European pollen data'. *Journal of Maps* **13**.2, 921–928.

Briggs, J. (2011). 'Mistletoe (*Viscum album*): A brief review of its local status with recent observations on its insect associations and conservation problems.' *Proceedings of the Cotteswold Naturalists' Field Club* **45**.2, 181–193.

Brown, A., J. Russell, R. Scaife, L. Tizzard, J. Whittaker and S. F. Wyles (2018). 'Lateglacial/early Holocene palaeoenvironments in the southern North Sea Basin: new data from the Dudgeon offshore wind farm'. *Journal of Quaternary Science* **33**, 597–610.

Burke, S. V., W. P. Wysocki, F. O. Zuloaga, J. M. Craine, J. C. Pires, P. P. Edger, D. Mayfield-Jones, L. G. Clark, S. A. Kelchner and M. R. Duvall (2016). 'Evolutionary relationships in Panicoid grasses based on plastome phylogenomics (Panicoideae; Poaceae)'. *BMC Plant Biology* **16**.1, 140–150.

Capo, E. *et al.* (2021). 'Lake Sedimentary DNA Research on Past Terrestrial and Aquatic Biodiversity: Overview and Recommendations'. *Quaternary* **4**.1, 6–66.

Caudullo, G., W. Tinner and D. de Rigo (2016). '*Picea abies* in Europe: distribution, habitat, usage and threats.' In: *European Atlas of Forest Tree Species*. Ed. by J. San-Miguel-Ayanz, D. de Rigo, G. Caudullo, T. Houston Durrant and A. Mauri. Luxembourg: Publications Office of the European Union, pp. 114–116.

Chefaoui, R. M., C. M. Duarte and E. A. Serrão (2017). 'Palaeoclimatic conditions in the Mediterranean explain genetic diversity of *Posidonia oceanica* seagrass meadows'. *Scientific Reports* **7**.1, 2732.

Clark, J. G. D. (1936). *The Mesolithic Settlement of Northern Europe: A Study of the Food-gathering Peoples of Northern Europe during the Early Post-Glacial Period*. Cambridge: Cambridge University Press.

Coles, B. J. (1998). 'Doggerland: a Speculative Survey'. *Proceedings of the Prehistoric Society* **64**, 45–81.

— (2000). 'Doggerland: the cultural dynamics of a shifting coastline'. *Geological Society, London, Special Publications* **175**.1, 393–401.

Colledge, S., J. Conolly and S. Shennan (2005). 'The Evolution of Neolithic Farming from SW Asian Origins to NW European Limits'. *European Journal of Archaeology* **8**.2, 137–156.

Coolen, M. J. L. and J. Overmann (2007). '217 000-year-old DNA sequences of green sulfur bacteria in Mediterranean sapropels and their implications for the reconstruction of the paleoenvironment'. *Environmental Microbiology* **9**.1, 238–249.

Corinaldesi, C., F. Beolchini and A. Dell'Anno (2008). 'Damage and degradation rates of extracellular DNA in marine sediments: implications for the preservation of gene sequences'. *Molecular Ecology* **17**, 3939–3951.

Cox, C. B. and P. D. Moore (1993). *Biogeography: an ecological and evolutionary approach*. 5th ed. Oxford: Blackwell Scientific.

Cribdon, B., R. Ware, O. Smith, V. Gaffney and R. G. Allaby (2020). 'PIA: More Accurate Taxonomic Assignment of Metagenomic Data Demonstrated on sedaDNA From the North Sea'. *Frontiers in Ecology and Evolution* **8**, 84.

Crombé, P. (2019). 'Mesolithic projectile variability along the southern North Sea basin (NW Europe): Hunter-gatherer responses to repeated climate change at the beginning of the Holocene'. *PLOS ONE* **14**.7, 1–19.

Dabney, J., M. Meyer and S. Paabo (2013a). 'Ancient DNA damage'. *Cold Spring Harb Perspect Biol* **5**.7, a012567.

Dabney, J. *et al.* (2013b). 'Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments'. *Proceedings of the National Academy of Sciences* **110**.39, 15758–15763.

De Rigo, D., C. M. Enescu, T. Houston Durrant, W. Tinner and G. Caudullo (2016). '*Juglans regia* in Europe: distribution, habitat, usage and threats.' In: *European Atlas of Forest Tree Species*. Ed. by J. San-Miguel-Ayanz, D. de Rigo, G. Caudullo,

T. Houston Durrant and A. Mauri. Luxembourg: Publications Office of the European Union, p. 103.

De Witte, L. C. and J. Stöcklin (2011). 'Horizontal growth in arctic-alpine clonal plants is not affected by climatic variability among regions'. *Plant Ecology & Diversity* **4**.4, 329–340.

Dell'Anno, A. and R. Danovaro (2005). 'Extracellular DNA Plays a Key Role in Deep-Sea Ecosystem Functioning'. *Science* **309**.5744, 2179.

Der Sarkissian, C. *et al.* (2016). 'Ancient DNA analysis identifies marine mollusc shells as new metagenomic archives of the past'. *Molecular Ecology Resources* **17**.5, 835–853.

Dupont, C., A. Tresset, N. Desse-Berset, Y. Gruet, G. Marchand and R. Schulting (2009). 'Harvesting the Seashores in the Late Mesolithic of Northwestern Europe: A View From Brittany'. *Journal of World Prehistory* **22**.2, 93–111.

Eaton, E., G. Caudullo and D. de Rigo (2016). '*Tilia platyphyllos* and other limes in Europe: distribution, habitat, usage and threats'. In: *European Atlas of Forest Tree Species*. Ed. by J. San-Miguel-Ayanz, D. de Rigo, G. Caudullo, T. Houston Durrant and A. Mauri. Luxembourg: Publications Office of the European Union, pp. 184–185.

Fang, X. and H. G. Stefan (1998). 'Temperature variability in lake sediments'. *Water Resources Research* **34**.4, 717–729.

Federhen, S. (2012). 'The NCBI Taxonomy database'. *Nucleic Acids Research* **40**.D1, D136–D143.

Feldman, M. and A. A. Levy (2015). 'Origin and Evolution of Wheat and Related Triticeae Species'. In: *Alien Introgression in Wheat: Cytogenetics, Molecular Biology, and Genomics*. Ed. by M. Molnár-Láng, C. Ceoloni and J. Doležel. Cham: Springer International Publishing. Chap. 2, pp. 21–76.

Ficetola, G. F. *et al.* (2018). 'DNA from lake sediments reveals long-term ecosystem changes after a biological invasion'. *Science Advances* **4**.5, eaar4292.

Fischer, A., J. Olsen, M. Richards, J. Heinemeier, Á. E. Sveinbjörnsdóttir and P. Bennike (2007). 'Coast–inland mobility and diet in the Danish Mesolithic and Neolithic: evidence from stable isotope values of humans and dogs'. *Journal of Archaeological Science* **34**.12, 2125–2150.

Fitch, S. (2018). 'Out of the blue: exploring Lost Frontiers in Doggerland'. In: *12th International Conference of Archaeological Prospection*. Ed. by B. Jennings, C. Gaffney, T. Sparrow and S. Gaffney. University of Bradford: Archaeopress, pp. 79–80.

Fort, J. (2015). 'Demic and cultural diffusion propagated the Neolithic transition across different regions of Europe'. *Journal of The Royal Society Interface* **12**.106, 20150166.

Fulton, T. L. and B. Shapiro (2019). 'Setting Up an Ancient DNA Laboratory BT - Ancient DNA: Methods and Protocols'. In: *Ancient DNA: Methods and Protocols*. Ed. by B. Shapiro, A. Barlow, P. D. Heintzman, M. Hofreiter, J. L. A. Paijmans and A. E. R. Soares. New York, NY: Springer New York. Chap. 1, pp. 1–13.

Gaffney, V. *et al.* (2020). 'Multi-Proxy Characterisation of the Storegga Tsunami and Its Impact on the Early Holocene Landscapes of the Southern North Sea.' *Geosciences* **10**.7, 270.

Gaffney, V. *et al.* (2017). ‘Doggerland and the Lost Frontiers Project (2015–2020)’. In: *Under the Sea: Archaeology and Palaeolandscapes of the Continental Shelf*. Ed. by G. N. Bailey, J. Harff and D. Sakellariou. Cham: Springer International Publishing. Chap. 20, pp. 305–319.

Galbraith, H., R. Jones, R. Park, J. Clough, S. Herrod-Julius, B. Harrington and G. Page (2002). ‘Global Climate Change and Sea Level Rise: Potential Losses of Intertidal Habitat for Shorebirds’. *Waterbirds* **25**.2, 173–183.

Garnier, S., N. Ross, B. Rudis, M. Sciaini and C. Scherer (2018). *viridis*. Version 0.5.1. URL: `https://cran.r-project.org/package=viridis`.

Garrow, D. and F. Sturt (2011). ‘Grey waters bright with Neolithic argonauts? Maritime connections and the Mesolithic–Neolithic transition within the ‘western seaways’ of Britain, c. 5000–3500 BC’. *Antiquity* **85**.327, 59–72.

Geary, B. R., E.-J. Hopla, I. Boomer, D. Smith, P. Marshall, S. Fitch, S. Griffiths and D. R. Tappin (2017). ‘Multi-proxy palaeoecological approaches to submerged landscapes: a case study from ‘Doggerland’, in the southern North Sea’. In: *The Archaeological and Forensic Applications of Microfossils: A Deeper Understanding of Human History*. Ed. by M. Williams, T. Hill, I. Boomer and I. P. Wilkinson. The Geological Society, pp. 35–53.

Giguet-Covex, C. *et al.* (2014). ‘Long livestock farming history and human landscape shaping revealed by lake sediment DNA’. *Nature Communications* **5**, 3211.

Gilbert, M. T. P., H.-J. Bandelt, M. Hofreiter and I. Barnes (2005). ‘Assessing ancient DNA studies’. *Trends in Ecology & Evolution* **20**.10, 541–544.

Gilbert, M. T. P. *et al.* (2008). ‘DNA from Pre-Clovis Human Coprolites in Oregon, North America’. *Science* **320**.5877, 786.

Gordon, A. and G. J. Hannon (2010). *FASTX-Toolkit*. Version 0.0.13. URL: `http://hannonlab.cshl.edu/fastx_toolkit/`.

Graham, R. W. *et al.* (2016). ‘Timing and causes of mid-Holocene mammoth extinction on St. Paul Island, Alaska’. *Proceedings of the National Academy of Sciences* **113**.33, 9310–9314.

Gregory, T. R. (2018). *Animal genome size database*. Version 2.0. URL: `http://www.genomesize.com/` (visited on 01/04/2020).

Haile, J., R. N. Holdaway, K. Oliver, M. Bunce, M. T. P. Gilbert, R. Nielsen, K. Munch, S. Y. W. Ho, B. Shapiro and E. Willerslev (2007). ‘Ancient DNA chronology within sediment deposits: are paleobiological reconstructions possible and is DNA leaching a factor?’ *Mol Biol Evol* **24**.4, 982–989.

Haile, J. *et al.* (2009). ‘Ancient DNA reveals late survival of mammoth and horse in interior Alaska’. *Proceedings of the National Academy of Sciences* **106**.52, 22352–22357.

Hansen, A. J., D. L. Mitchell, C. Wiuf, L. Paniker, T. B. Brand, J. Binladen, D. A. Gilichinsky, R. Rønn and E. Willerslev (2006). ‘Crosslinks Rather Than Strand Breaks Determine Access to Ancient DNA Sequences From Frozen Sediments’. *Genetics* **173**.2, 1175–1179.

Herbig, A., F. Maixner, K. I. Bos, A. Zink, J. Krause and D. H. Huson (2016). ‘MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman’. *bioRxiv*, 050559.

Holdredge, C. and M. D. Bertness (2011). ‘Litter legacy increases the competitive advantage of invasive *Phragmites australis* in New England wetlands’. *Biological Invasions* **13**.2, 423–433.

Huang, J., D. E. Giannasi and R. A. Price (2005). ‘Phylogenetic relationships in Ephedra (Ephedraceae) inferred from chloroplast and nuclear DNA sequences’. *Molecular Phylogenetics and Evolution* **35**.1, 48–59.

Hughes, R. G. (2004). ‘Climate change and loss of saltmarshes: consequences for birds’. *Ibis* **146**, 21–28.

Huson, D. H. (2019). *User Manual for MEGAN V6.17.0.* URL: `http://ab.inf.uni-tuebingen.de/data/software/megan6/download/manual.pdf`.

Huson, D. H., S. Beier, I. Flade, A. Górska, M. El-Hadidi, S. Mitra, H.-J. Ruscheweyh and R. Tappu (2016). ‘MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data’. *PLoS Computational Biology* **12**.6, e1004957.

Huson, D. H., A. F. Auch, J. Qi and S. C. Schuster (2007). ‘MEGAN analysis of metagenomic data’. *Genome research* **17**.3, 377–386.

Inagaki, F. *et al.* (2015). ‘Exploring deep microbial life in coal-bearing sediment down to ∼2.5 km below the ocean floor’. *Science* **349**.6246, 420–424.

Janssens, S. B. *et al.* (2020). ‘A large-scale species level dated angiosperm phylogeny for evolutionary and ecological analyses’. *Biodiversity Data Journal* **8**, e39677.

Johnson, S. S. *et al.* (2007). ‘Ancient bacteria show evidence of DNA repair’. *Proceedings of the National Academy of Sciences* **104**.36, 14401–14405.

Jónsson, H., A. Ginolhac, M. Schubert, P. L. F. Johnson and L. Orlando (2013). ‘mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters’. *Bioinformatics* **29**.13, 1682–1684.

Key, F. M., C. Posth, J. Krause, A. Herbig and K. I. Bos (2017). ‘Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication’. *Trends in Genetics* **33**.8, 508–520.

Kircher, M., S. Sawyer and M. Meyer (2012). ‘Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform’. *Nucleic Acids Research* **40**.1, e3.

Kistler, L., R. Ware, O. Smith, M. Collins and R. G. Allaby (2017). ‘A new model for ancient DNA decay based on paleogenomic meta-analysis’. *Nucleic Acids Res* **45**.11, 6310–6320.

Kolde, R. (2018). *pheatmap*. Version 1.0.12. URL: `https://www.rdocumentation.org/packages/pheatmap/versions/1.0.12`.

Krüger, S., W. Dörfler, O. Bennike and S. Wolters (2017). ‘Life in Doggerland – palynological investigations of the environment of prehistoric hunter-gatherer societies in the North Sea Basin’. *E&G Quaternary Science Journal* **66**.1, 3–13.

Kuch, M., N. Rohland, J. L. Betancourt, C. Latorre, S. Steppan and H. N. Poinar (2002). ‘Molecular analysis of a 11 700-year-old rodent midden from the Atacama Desert, Chile’. *Molecular Ecology* **11**.5, 913–924.

Leary, J. I. M. (2009). ‘Perceptions of and responses to the Holocene flooding of the North Sea lowlands’. *Oxford Journal of Archaeology* **28**.3, 227–237.

Lee, M. K., Y. I. Lee, H. S. Lim, J. I. Lee, J. H. Choi and H. I. Yoon (2011). ‘Comparison of radiocarbon and OSL dating methods for a Late Quaternary sediment core from Lake Ulaan, Mongolia’. *Journal of Paleolimnology* **45**.2, 127–135.

Leitch, I., E. Johnston, J. Pellicer, O. Hidalgo and M. Bennett (2019). *Plant DNA C-values database*. Version 7.1. URL: `https://cvalues.science.kew.org/` (visited on 01/04/2020).

Leonard, J. A., O. Shanks, M. Hofreiter, E. Kreuz, L. Hodges, W. Ream, R. K. Wayne and R. C. Fleischer (2007). 'Animal DNA in PCR reagents plagues ancient DNA research'. *Journal of Archaeological Science* **34**.9, 1361–1366.

Lindahl, T. and B. Nyberg (1972). 'Rate of depurination of native deoxyribonucleic acid'. *Biochemistry* **11**.19, 3610–3618.

Liu, Q., X. Li, X. Zhou, M. Li, F. Zhang, T. Schwarzacher and J. S. Heslop-Harrison (2019). 'The repetitive DNA landscape in Avena (Poaceae): chromosome and genome evolution defined by major repeat classes in whole-genome sequence reads'. *BMC Plant Biology* **19**.1, 226.

Lu, K.-Q. *et al.* (2019). 'New pollen classification of Chenopodiaceae for exploring and tracing desert vegetation evolution in eastern arid central Asia'. *Journal of Systematics and Evolution* **57**.2, 190–199.

Marchand, G. (2007). 'Neolithic fragrances: Mesolithic-Neolithic interactions in western France'. *Proceedings of the British Academy* **144**, 225–242.

Martin, C., G. Ménot, N. Thouveny, O. Peyron, V. Andrieu-Ponel, V. Montade, N. Davtian, M. Reille and E. Bard (2020). 'Early Holocene Thermal Maximum recorded by branched tetraethers and pollen in Western Europe (Massif Central, France)'. *Quaternary Science Reviews* **228**, 106109–106129.

Meyer, M. and M. Kircher (2010). 'Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing'. *Cold Spring Harbor Protocols* **2010**.6.

Meylemans, E., J. Bastiaens, M. Boudin, K. Deforce, A. Ervynck, Y. Perdaen, J. Sergant, A. Storme and P. Crombé (2018). 'The oldest cereals in the coversand area along the North Sea coast of NW Europe, between ca. 4800 and 3500 cal BC, at the wetland site of 'Bazel-Sluis' (Belgium)'. *Journal of Anthropological Archaeology* **49**, 1–7.

National Center for Biotechnology Information (US) (2021). *GenBank and WGS Statistics*. URL: `https://www.ncbi.nlm.nih.gov/genbank/statistics/` (visited on 03/01/2021).

Olsen, J. L. *et al.* (2016). 'The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea'. *Nature* **530**.7590, 331–335.

Ondov, B. D., N. H. Bergman and A. M. Phillippy (2011). 'Interactive metagenomic visualization in a Web browser'. *BMC Bioinformatics* **12**.1, 385.

Orlando, L. *et al.* (2021). 'Ancient DNA analysis'. *Nature Reviews Methods Primers* **1**.14.

Packer, J. G., L. A. Meyerson, H. Skálová, P. Pyšek and C. Kueffer (2017). 'Biological Flora of the British Isles: *Phragmites australis*'. *Journal of Ecology* **105**.4, 1123–1162.

Packham, J., D. Harding, G. Hilton and R. Stuttard (1992). 'Origins of British woodlands and their floras'. In: *Functional Ecology of Woodlands and Forests*. Ed. by J. Packham. Cambridge: Kluwer Academic Publishers. Chap. 5.6, pp. 179–185.

Packham, J. R. and A. J. Willis (1997). *Ecology of Dunes, Salt Marsh and Shingle*. London: Springer Science & Business Media.

Pan, Y., S. Yan, H. Behling and G. Mu (2013). 'Transport of airborne *Picea schrenkiana* pollen on the northern slope of Tianshan Mountains (Xinjiang, China) and its implication for paleoenvironmental reconstruction'. *Aerobiologia* **29**.2, 161–173.

Pansu, J., C. Giguet-Covex, G. F. Ficetola, L. Gielly, F. Boyer, L. Zinger, F. Arnaud, J. Poulenard, P. Taberlet and P. Choler (2015). 'Reconstructing long-term human impacts on plant communities: an ecological approach based on lake sediment DNA'. *Molecular Ecology* **24**.7, 1485–1498.

Parducci, L., K. D. Bennett, G. F. Ficetola, I. G. Alsos, Y. Suyama, J. R. Wood and M. W. Pedersen (2017). 'Ancient plant DNA in lake sediments'. *New Phytologist* **214**.3, 924–942.

Pedersen, M. W., A. Ginolhac, L. Orlando, J. Olsen, K. Andersen, J. Holm, S. Funder, E. Willerslev and K. H. Kjær (2013). 'A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa'. *Quaternary Science Reviews* **75**.Supplement C, 161–168.

Pedersen, M. W. *et al.* (2015). 'Ancient and modern environmental DNA'. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**.1660, 20130383.

Peeters, J. H. M. and G. Momber (2014). 'The southern North Sea and the human occupation of northwest Europe after the Last Glacial Maximum'. *Netherlands Journal of Geosciences* **93**.1-2, 55–70.

Pianka, E. R. (1973). 'The Structure of Lizard Communities'. *Annual Review of Ecology and Systematics* **4**.1, 53–74.

Pividori, M., F. Giannetti, A. Barbati and G. Chirici (2016). 'European Forest Types: tree species matrix'. In: *European Atlas of Forest Tree Species.* Ed. by J. San-Miguel-Ayanz, D. de Rigo, G. Caudullo, T. Houston Durrant and A. Mauri. Luxembourg: Publications Office of the European Union, p. 35.

Platt, D. E. *et al.* (2017). 'Mapping Post-Glacial expansions: The Peopling of Southwest Asia'. *Scientific reports* **7**, 40338.

Pollegioni, P. *et al.* (2017). 'Rethinking the history of common walnut (*Juglans regia* L.) in Europe: Its origins and human interactions'. *PLOS ONE* **12**.3, e0172541.

Preece, R. C. and S. P. Day (1994). 'Special Paper: Comparison of Post-Glacial Molluscan and Vegetational Successions from a Radiocarbon-Dated Tufa Sequence in Oxfordshire'. *Journal of Biogeography* **21**.5, 463–478.

Raniello, R. and G. Procaccini (2002). 'Ancient DNA in the seagrass *Posidonia oceanica*'. *Marine Ecology Progress Series* **227**, 269–273.

Rawlence, N. J., D. J. Lowe, J. R. Wood, J. M. Young, J. Churchman G, Y. Huang and A. Cooper (2014). 'Using palaeoenvironmental DNA to reconstruct past environments: progress and prospects'. *Journal of Quaternary Science* **29**.7, 610–626.

Reid, C. (1913). *Submerged Forests.* Cambridge: Cambridge University Press.

Renaud, G., M. Schubert, S. Sawyer and L. Orlando (2019). 'Authentication and Assessment of Contamination in Ancient DNA'. In: *Ancient DNA: Methods and Protocols.* Ed. by B. Shapiro, A. Barlow, P. D. Heintzman, M. Hofreiter, J. L. A. Paijmans and A. E. R. Soares. 2nd ed. New York, NY: Springer New York. Chap. 17, pp. 163–194.

Rivkina, E. M., E. I. Friedmann, C. P. McKay and D. A. Gilichinsky (2000). 'Metabolic Activity of Permafrost Bacteria below the Freezing Point'. *Applied and Environmental Microbiology* **66**.8, 3230–3233.

Robb, J. (2013). 'Material Culture, Landscapes of Action, and Emergent Causation: A New Model for the Origins of the European Neolithic'. *Current Anthropology* **54**.6, 657–683.

Rohland, N. and D. Reich (2012). 'Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture'. *Genome research* **22**.5, 939–946.

Rowley-Conwy, P. (2014). 'Foragers and Farmers in Mesolithic/Neolithic Europe, 5500–3900 cal. BC: Beyond the Anthropological Comfort Zone.' In: *Wild Things: Recent Advances in Palaeolithic and Mesolithic Research*. Ed. by F. W. F. Foulds, H. C. Drinkall, A. R. Perri and D. T. G. Clinnick. Oxford: Oxbow Books. Chap. 15, pp. 185–201.

Salter, S. J., M. J. Cox, E. M. Turek, S. T. Calus, W. O. Cookson, M. F. Moffatt, P. Turner, J. Parkhill, N. J. Loman and A. W. Walker (2014). 'Reagent and laboratory contamination can critically impact sequence-based microbiome analyses'. *BMC biology* **12**.12, 87.

San-Miguel-Ayanz, J. *et al.* (2016). *European Atlas of Forest Tree Species*. Ed. by J. Miguel-Ayanz, D. de Rigo, G. Caudullo, T. Houston Durrant and A. Mauri. Luxembourg: Publications Office of the European Union.

Sayers, E. W., M. Cavanaugh, K. Clark, K. D. Pruitt, C. L. Schoch, S. T. Sherry and I. Karsch-Mizrachi (2021). 'GenBank'. *Nucleic Acids Research* **49**.D1, D92–D96.

Schneider-Binder, E. and F. Kuhlke (2015). 'Habitats with Sea Grape (*Ephedra distachya*) on the Dunes of Letea (Danube Delta, Romania)'. *Transylvanian Review of Systematical and Ecological Research* **17**.2, 45–56.

Schubert, M., S. Lindgreen and L. Orlando (2016). 'AdapterRemoval v2: rapid adapter trimming, identification, and read merging'. *BMC Research Notes* **9**.1, 88.

Schubert, M. *et al.* (2014). 'Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX'. *Nature Protocols* **9**.5, 1056–1082.

Shah, N., M. G. Nute, T. Warnow and M. Pop (2018). 'Misunderstood parameter of NCBI BLAST impacts the correctness of bioinformatics workflows'. *Bioinformatics* **35**.9, 1613–1614.

Shapiro, B., A. Barlow, P. D. Heintzman, M. Hofreiter, J. L. A. Paijmans and A. E. R. Soares, eds. (2019). *Ancient DNA Methods and Protocols*. 2nd ed. New York, NY: Springer Science & Business Media.

Short, F., T. Carruthers, W. Dennison and M. Waycott (2007). 'Global seagrass distribution and diversity: A bioregional model'. *Journal of Experimental Marine Biology and Ecology* **350**.1, 3–20.

Silva, F. and M. Vander Linden (2017). 'Amplitude of travelling front as inferred from 14C predicts levels of genetic admixture among European early farmers'. *Scientific Reports* **7**.1, 11985.

Skoglund, P., B. H. Northoff, M. V. Shunkov, A. P. Derevianko, S. Pääbo, J. Krause and M. Jakobsson (2014). 'Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal'. *Proceedings of the National Academy of Sciences* **111**.6, 2229–2234.

Slon, V. *et al.* (2017). 'Neandertal and Denisovan DNA from Pleistocene sediments'. *Science* **356**.6338, 605–608.

Smith, O., R. G. Allaby, G. Momber, R. Bates, P. Garwood, S. Fitch, M. Pallen and V. Gaffney (2015). 'Sedimentary DNA from a submerged site reveals wheat in the British Isles 8000 years ago'. *Science* **347**.6225, 998–1001.

Sobek, S., E. Durisch-Kaiser, R. Zurbrügg, N. Wongfun, M. Wessels, N. Pasche and B. Wehrli (2009). 'Organic carbon burial efficiency in lake sediments controlled by oxygen exposure time and sediment source'. *Limnology and Oceanography* **54**.6, 2243–2254.

Sommer, R. S., C. Lindqvist, A. Persson, H. Bringsøe, A. G. J. Rhodin, N. Schneeweiss, P. Široký, L. Bachmann and U. W. E. Fritz (2009). 'Unexpected early extinction of the European pond turtle (*Emys orbicularis*) in Sweden and climatic impact on its Holocene range'. *Molecular Ecology* **18**.6, 1252–1262.

Sommer, R. S., A. Persson, N. Wieseke and U. Fritz (2007). 'Holocene recolonization and extinction of the pond turtle, *Emys orbicularis* (L., 1758), in Europe'. *Quaternary Science Reviews* **26**.25, 3099–3107.

Sønstebø, J. H. *et al.* (2010). 'Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate'. *Molecular Ecology Resources* **10**.6, 1009–1018.

Stahlschmidt, M. C. *et al.* (2019). 'Ancient Mammalian and Plant DNA from Late Quaternary Stalagmite Layers at Solkota Cave, Georgia'. *Scientific Reports* **9**.1, 6628–6637.

Stivrins, N., J. Soininen, I. Tõnno, R. Freiberg, S. Veski and V. Kisand (2018). 'Towards understanding the abundance of non-pollen palynomorphs: A comparison of fossil algae, algal pigments and sedaDNA from temperate lake sediments'. *Review of Palaeobotany and Palynology* **249**, 9–15.

Sturt, F., D. Garrow and S. Bradley (2013). 'New models of North West European Holocene palaeogeography and inundation'. *Journal of Archaeological Science* **40**.11, 3963–3976.

Svizzero, S. (2015). 'The long-term decline in terms of trade and the neolithisation of Northern Europe'. *Scandinavian Economic History Review* **63**.3, 260–279.

Thomsen, P. F., S. Elias, M. T. P. Gilbert, J. Haile, K. Munch, S. Kuzmina, D. G. Froese, A. Sher, R. N. Holdaway and E. Willerslev (2009). 'Non-Destructive Sampling of Ancient Insect DNA'. *PLoS One* **4**.4, e5048.

Torti, A., M. A. Lever and B. B. Jørgensen (2015). 'Origin, dynamics, and implications of extracellular DNA pools in marine sediments'. *Marine Genomics* **24**, 185–196.

Tresset, A. and J.-D. Vigne (2011). 'Last hunter-gatherers and first farmers of Europe'. *Comptes Rendus Biologies* **334**.3, 182–189.

Van der Plicht, J., L. W. S. W. Amkreutz, M. J. L. T. Niekus, J. H. M. Peeters and B. I. Smit (2016). 'Surf'n Turf in Doggerland: Dating, stable isotopes and diet of Mesolithic human remains from the southern North Sea'. *Journal of Archaeological Science: Reports* **10**, 110–118.

Van der Valk, T. *et al.* (2021). 'Million-year-old DNA sheds light on the genomic history of mammoths'. *Nature* **591**, 265–269.

Velsko, I. M., L. A. F. Frantz, A. Herbig, G. Larson and C. Warinner (2018). 'Selection of Appropriate Metagenome Taxonomic Classifiers for Ancient Microbiome Research'. *mSystems* **3**.4, e00080–18.

Walker, J., V. Gaffney, S. Fitch, M. Muru, A. Fraser, M. Bates and R. Bates (2020). 'A great wave: the Storegga tsunami and the end of Doggerland?' *Antiquity* **94**.378, 1409–1425.

Walker, K. (2015). *Dryas octopetala L.. Mountain Avens. Species Account.* URL: `https://bsbi.org/wp-content/uploads/dlm_uploads/Dryas_octopetala_species_account.pdf%20https://bsbi.org/species-accounts` (visited on 02/06/2020).

Warinner, C., C. Speller and M. J. Collins (2015). 'A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome'. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**.1660, 20130376.

Weninger, B. *et al.* (2008). 'The catastrophic final flooding of Doggerland by the Storegga Slide tsunami'. *Documenta Praehistorica* **35**, 1–24.

Weyrich, L. S. *et al.* (2017). 'Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus'. *Nature* **544**, 357–372.

Wicker, T. *et al.* (2017). 'The repetitive landscape of the 5100 Mbp barley genome'. *Mobile DNA* **8**, 22.

Wicker, T. *et al.* (2018). 'Impact of transposable elements on genome structure and evolution in bread wheat'. *Genome Biology* **19**, 103.

Willerslev, E. and A. Cooper (2005). 'Ancient DNA'. *Proceedings: Biological Sciences* **272**.1558, 3–16.

Willerslev, E., A. J. Hansen, J. Binladen, T. B. Brand, M. T. P. Gilbert, B. Shapiro, M. Bunce, C. Wiuf, D. A. Gilichinsky and A. Cooper (2003). 'Diverse Plant and Animal Genetic Records from Holocene and Pleistocene Sediments'. *Science* **300**.5620, 791–795.

Willerslev, E. *et al.* (2007). 'Ancient Biomolecules from Deep Ice Cores Reveal a Forested Southern Greenland'. *Science* **317**.5834, 111–114.

Willerslev, E. *et al.* (2014). 'Fifty thousand years of Arctic vegetation and megafaunal diet'. *Nature* **506**.7486, 47–51.

Wolfe, K. H., P. M. Sharp and W.-H. Li (1989). 'Rates of synonymous substitution in plant nuclear genes'. *Journal of Molecular Evolution* **29**.3, 208–211.

Wolters, S., M. Zeiler and F. Bungenstock (2010). 'Early Holocene environmental history of sunken landscapes: pollen, plant macrofossil and geochemical analyses from the Borkum Riffgrund, southern North Sea'. *International Journal of Earth Sciences* **99**.8, 1707–1719.

Zhang, Z., S. Schwartz, L. Wagner and W. Miller (2000). 'A Greedy Algorithm for Aligning DNA Sequences'. *Journal of Computational Biology* **7**.1-2, 203–214.

Zhou, Z., N. Luhmann, N.-F. Alikhan, C. Quince and M. Achtman (2018). 'Accurate Reconstruction of Microbial Strains from Metagenomic Sequencing Using Representative Reference Genomes'. In: *Research in Computational Molecular Biology. RECOMB 2018. Lecture Notes in Computer Science*. Ed. by B. J. Raphael. Vol. **10812**. Cham: Springer International Publishing, pp. 225–240.

Zinger, L. *et al.* (2019). 'DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions'. *Molecular Ecology* **28**, 1857–1862.

Appendix A

Full PDF of Cribdon *et al.* (2020)

The following paper was published as Cribdon *et al.* (2020) and can be downloaded at https://doi.org/10.3389/fevo.2020.00084.

frontiers
in Ecology and Evolution

ORIGINAL RESEARCH
published: 03 April 2020
doi: 10.3389/fevo.2020.00084

PIA: More Accurate Taxonomic Assignment of Metagenomic Data Demonstrated on sedaDNA From the North Sea

Becky Cribdon[1], Roselyn Ware[1], Oliver Smith[1†], Vincent Gaffney[2] and Robin G. Allaby[1*]

[1] School of Life Sciences, University of Warwick, Coventry, United Kingdom, [2] School of Archaeological and Forensic Sciences, University of Bradford, Bradford, United Kingdom

Assigning metagenomic reads to taxa presents significant challenges. Existing approaches address some issues, but are mostly limited to metabarcoding or optimized for microbial data. We present PIA (Phylogenetic Intersection Analysis): a taxonomic binner that works from standard BLAST output while mitigating key effects of incomplete databases. Benchmarking against MEGAN using sedaDNA suggests that, while PIA is less sensitive, it can be more accurate. We use known sequences to estimate the accuracy of PIA at up to 96% when the real organism is not represented in the database. For ancient DNA, where taxa of interest are frequently over-represented domesticates or absent, poorly-known organisms, more accurate assignment is critical, even at the expense of sensitivity. PIA offers an approach to objectively filter out false positive hits without the need to manually remove taxa and so make presuppositions about past environments and their palaeoecologies.

Keywords: ancient DNA, BLAST, MEGAN, metagenomics, sedaDNA, taxonomic assignment

OPEN ACCESS

Edited by:
Michael Knapp,
University of Otago, New Zealand

Reviewed by:
John Kenneth Pearman,
Cawthron Institute, New Zealand
Kieren James Mitchell,
The University of Adelaide, Australia

Correspondence:
Robin G. Allaby
r.g.allaby@warwick.ac.uk

†Present address:
Oliver Smith,
Micropathology Ltd., University of Warwick Science Park, Coventry, United Kingdom

Specialty section:
This article was submitted to Paleoecology,
a section of the journal
Frontiers in Ecology and Evolution

Received: *22 October 2019*
Accepted: *12 March 2020*
Published: *03 April 2020*

Citation:
Cribdon B, Ware R, Smith O, Gaffney V and Allaby RG (2020) PIA: More Accurate Taxonomic Assignment of Metagenomic Data Demonstrated on sedaDNA From the North Sea. Front. Ecol. Evol. 8:84. doi: 10.3389/fevo.2020.00084

INTRODUCTION

Next-generation sequencing allows detailed metagenomic analysis of a wide range of ancient samples. Studies have attempted to recreate biological communities from material including coprolites (Bon et al., 2012; Appelt et al., 2014), dental calculus (Warinner et al., 2015; Weyrich et al., 2017), ice cores (Willerslev et al., 2007), sediment (Birks and Birks Hilary, 2015; Smith et al., 2015), stalagmites (Stahlschmidt et al., 2019), rodent middens (Kuch et al., 2002) and mollusc shells (Der Sarkissian et al., 2016). Our understanding of contamination and best laboratory practice has made good progress (Gilbert et al., 2005; Shapiro et al., 2019) and methods for authenticating ancient DNA sequences are developing (Key et al., 2017; Renaud et al., 2019). However, identifying ancient metagenomic sequences is still a challenge, particularly for shotgun data.

Shotgun sequencing has three key advantages over metabarcoding for ancient metagenomics. First, it can capture information from anywhere in the genome, greatly increasing sensitivity. Every DNA molecule extracted from a sample has the potential to be identified, provided that reference databases are adequate. Second, read count and genome size could be used to calculate biogenomic mass: a proxy of biomass (Gaffney et al., 2020). Third, metabarcoding is far less likely to record DNA damage signals. Damage accumulates in DNA over time (Kistler et al., 2017), so is important for authentication of ancient reads, and occurs most rapidly on the single-stranded overhangs at the ends of molecules. A characteristic damage signal is C-T deamination; changes to the base sequence make it less likely that metabarcoding primers will anneal, so damaged molecules are less likely to be sequenced. Furthermore, primer regions are typically removed during analysis, so even if the very ends of molecules are amplified, they will not be considered. Shotgun sequencing can potentially sequence whole molecules, especially when fragments are short, as is the case for ancient DNA. This preserves any damage signal intact. Overall, shotgun data has the potential to supply highly sensitive and informative metagenomic data.

However, because sequences can come from anywhere in the genome, accurately assigning shotgun reads to taxa requires a much larger reference database than for metabarcoding. The GenBank database is the most comprehensive (Benson et al., 2016), but even this is highly incomplete. Only a tiny fraction of organisms have had their full genomes sequenced and most are not represented at all. Reads from unrepresented organisms may go unassigned. Worse, the uneven representation of taxa that are in a database can create two additional problems that may lead to incorrect assignments.

The first problem is the over-representation of some taxa. This was recently identified as an issue for BLAST (Zhang et al., 2000), the "gold standard" of taxonomic binning (Herbig et al., 2016), by Shah et al. (2018). When BLAST searches against a database, it starts at the top and returns the first *n* hits that pass a quality filter, not the best *n* hits. If an over-represented taxon is a reasonable match, BLAST could return *n* hits and finish before it has a chance to identify closer but less represented taxa further down the database. Better matches may be missing from the list of hits. Even if BLAST does check the whole database, the list of hits may be disproportionately full of over-represented taxa. Taxonomic assignment methods that consider this list may then assign with too much weight to these taxa.

The second problem with an uneven database is "oasis" taxa in "sparse" areas. Consider a sparsely-populated area of the database with just one or a few taxa represented, not including the real taxon (**Figure 1B**). A specific sequence is unlikely to hit anything and will probably be left unassigned. But a conservative sequence may hit that one or few taxa, not necessarily because they are a good match, but because there is nothing else closer. The list of BLAST hits for that read will not be empty, but will have very low diversity. This can give the illusion of a confident match. Taxonomic binners that use a phylogenetic intersection or "lowest common ancestor" approach, robust to conservative sequences, can produce false positives because of oasis taxa.

BLAST and BLAST-like algorithms have a minimum quality filter that affects how similar a reference sequence must be to count as a hit and how much empty space there must be around a read for it to go unassigned (**Figure 1**, "hit radius"). But as with many aspects of taxonomic assignment, this filter has a trade-off between accuracy and sensitivity. A very strict filter would increase the resistance of reads to not-very-similar oases, but make them less attracted to more similar sequences that could be informative. This is especially an issue for aDNA, where even a read from an organism that is in the database may not share an identical sequence because of DNA damage or mutations over time. The minimum quality filter cannot protect from oasis taxa alone.

One of the main arguments in favor of metabarcoding is its use of confined, curated databases that aim to be functionally complete for the study taxon in the study area, such as the Arctic flora database in Sønstebø et al. (2010). Uneven representation is limited if all taxa are represented to some degree. It is currently realistic to sequence a barcode region of a several hundred species for a study, as in Sønstebø et al. But because shotgun sequencing can access the whole genome, a complete shotgun database must have the full genome of all organisms, which will not happen in the foreseeable future. Metabarcoding databases are typically far more "complete" in that more of the study taxa are represented. However, this still assumes that an environment can even be well-studied enough for a complete list of taxa. This is debatable, especially for ancient ecosystems. Despite metabarcoding databases being easier to fill, arguably neither can ever be truly complete. Metabarcoding does not fully address uneven representation in databases. Both metabarcoding and shotgun approaches would benefit from an alternative solution.

A method that accepts shotgun data while also improving the database is SPARSE (Zhou et al., 2018). It rebuilds a given database as hierarchical clusters of similar sequences. If a taxon is represented by several very similar genomes, these genomes will be combined into a single cluster. The final SPARSE database has every present taxon represented by one genome, addressing the problem of over-represented taxa. However, SPARSE is designed for microbial data in relatively well-studied systems, where the database is both relatively well-populated and small enough to be rebuilt on a typical lab server. It does not address the problem of oasis taxa in sparse areas, nor would it be easily applicable to studies of organisms with larger genomes.

A popular standard tool for metagenomic studies not limited to microbes is MEGAN (Huson et al., 2007, 2016). This analyses output from various reference-matching programs, including BLAST. Its sister program, MALT (Herbig et al., 2016), aims to generate comparable output to BLAST at greatly increased speed before assigning taxonomy in the same way as MEGAN. This shared method is the LCA (Lowest Common Ancestor) algorithm (Huson et al., 2016). The default naive LCA is best suited to taxonomic binning. For each read, hits are first quality-filtered against multiple criteria. Good hits are assumed to belong not to the single organism they were sequenced from, but the "lowest common ancestor" (ancestral node) of all associated taxa. Being associated with multiple taxa suggests that the hit sequence is conservative, so should be assigned to a higher taxon. The more conserved the sequence, the more diverse the associated taxa, so the higher the taxon to which the hit is assigned. Following the same logic, the *read* is then assigned to the lowest common ancestor of its list of processed hits.

The LCA is robust to overrepresented taxa in the list of hits. The lowest common ancestor is calculated on presence/absence, not number of occurrences. However, accurate assignment still depends on the list containing accurate hits to begin with, which overrepresented taxa can prevent (Shah et al., 2018).

The LCA also addresses unrepresented taxa: even if the real taxon is not in the database, the list of hits should include relatives, so the read should be assigned to an "ancestor" that encompasses the real taxon. The more sparse the database,

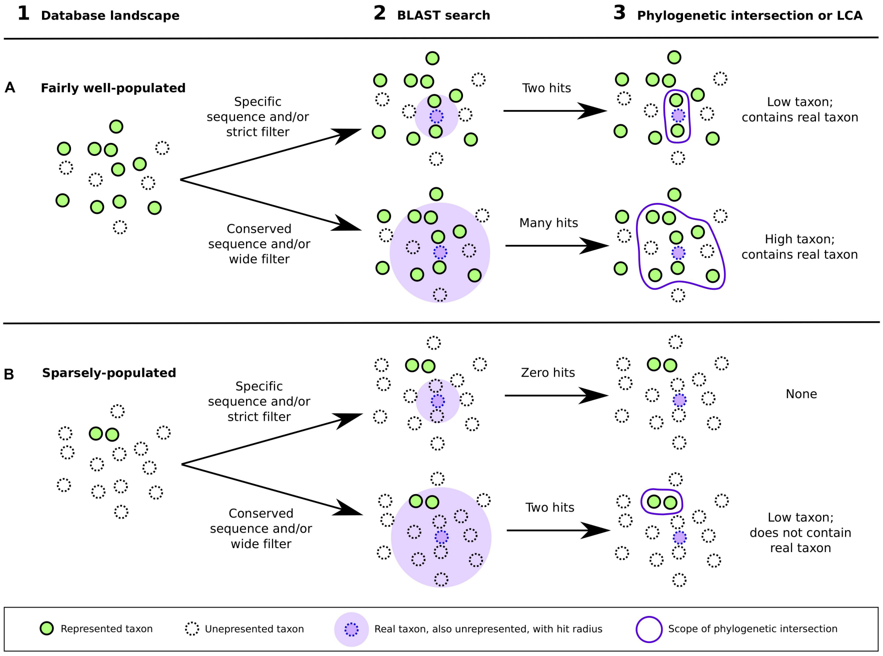


FIGURE 1 | Database landscape diagrams of an **(A)** well-populated and **(B)** sparsely-populated region. Column (1) is the state of the database region. (2) Shows a BLAST search for a sequence from an unrepresented taxon; for a reference sequence to be a hit, it must be inside the hit radius. Column (3) is what phylogenetic intersection or LCA ("lowest common ancestor") those BLAST hits would produce: what the read would be assigned to. In a well-populated region **(A)**, the more conservative the sequence (or the wider the quality filter), the higher the taxon the read would be assigned to. The intersection should contain diverse relatives of the real taxon, so is likely to contain the real taxon. In a sparsely-populated region **(B)**, a specific sequence (or strict filter) is unlikely to be assigned. But a more conservative sequence (or wider filter) may hit the one or few represented taxa in the region: oases. These are probably not close relatives of the real taxon. However, the resulting low-diversity list of hits gives the intersection an illusion of confidence.

the more diverse the list of hits, so the higher the taxon the read is assigned to. In very sparse regions, this means that reads are likely to be under- or unassigned but not incorrectly over-assigned (Huson et al., 2007). However, we argue that the LCA approach may incorrectly assign these reads if they are influenced by oasis taxa. If, for instance, a sparse region were occupied by clumps of taxa rather than an even spread of relatives around the unrepresented taxon (**Figure 1B**), the list of hits may be dominated by one of those taxon clumps, resulting in a relatively specific "ancestor" close to the oasis but not necessarily the real taxon.

MEGAN does have a further check against false positives: the *min-support* filter (Huson et al., 2007; Huson, 2019). Once all reads have been assigned, resulting taxa are only reported if they contain a minimum number of reads. If a read was assigned to a taxon that does not meet this threshold, it is pushed up the taxonomy until it reaches a taxon that does. This excludes very rare taxa, which Huson et al. argue are more likely to be false positives. However, we argue that oasis taxa could escape this check. Being the only represented taxon in that database region, an oasis could potentially pull in reads that would otherwise be assigned to multiple local taxa. The fewer other taxa around, the stronger the oasis effect, and the greater the number of reads incorrectly assigned to that taxon. Oasis taxa can systematically generate false positives that are not necessarily rare.

In this paper, we present Phylogenetic Intersection Analysis (PIA) as a taxonomic binner which, like MEGAN, works from gold-standard BLAST output and is not designed specifically for microbial data, yet goes further to address the shortcomings of BLAST and databases. It also filters BLAST hits by a strict quality threshold. It also accounts for over-represented taxa by only counting each hit taxon once. It also avoids over-assigning conservative hits and sequences by finding a lowest common ancestor, here called a phylogenetic intersection to avoid ambiguity when dealing with ancient sequences that may genuinely be ancestral. However, there are two key differences between MEGAN and PIA. First is a difference with finding the intersection. MEGAN accepts an LCA calculated from just one taxon (i.e., that taxon itself), but if PIA does not have at least two taxa, it discards the read. It assumes that the real taxon is not in the database, so will not assign directly to a taxon in the database. It only assigns to a higher taxon, assuming that the real taxon lies within that phylogenetic range. This avoids over-assigning unrepresented reads to close relatives. Second is a diversity check that measures the extent of population in the region of the database. Reads assigned in sparse regions, vulnerable to the influence of oasis taxa, are discarded. PIA discards the majority of reads, but those that remain are robustly assigned. The resulting assignations are reliable despite low read counts.

This study evaluates PIA by benchmarking its performance against MEGAN with empirical and simulated data. The empirical data was generated as part of the Europe's Lost Frontiers project. This aims to reconstruct submerged palaeolandscapes around the United Kingdom, particularly Doggerland, which now lies under the North Sea. One arm of the project is multi-proxy analysis of sediment cores. This study uses our sedaDNA data from core ELF039, chosen because most samples had a relatively high data yield and the geological context suggested a potentially interesting story. For more information, see Gaffney et al. (2020).

ALGORITHM

A very early version of PIA was originally presented in Smith et al. (2015). Although the central approach has not changed, it has been substantially rewritten and refined. Scripts are available from https://github.com/Allaby-lab/PIA.

The Input BLAST File

The two inputs for PIA are a FASTA of query sequences and a corresponding BLAST file. The BLAST file must be in format six (tabular) with all standard columns followed by an additional column containing taxonomic IDs associated with the reference sequence hit. This column is how PIA assigns hits to taxa. We also use the "-max_target_seqs" parameter to limit the number of hits returned per query sequence, recognizing that the hits returned will be the first *n* to meet a quality threshold (Shah et al., 2018). Although PIA aims to reduce the impact of overrepresented taxa in databases once the BLAST is complete, it is important that this BLAST takes enough hits to reach underrepresented taxa. "-max_target_seqs" should be as high as practical. We suggest 500 as a default. Finally, note that BLAST can be run with *x* number of threads. Many of our larger samples took days to BLAST despite using several threads. This is by far the most computationally expensive part of the pipeline.

A typical pre-PIA BLAST command:

```
blastn -db [nucleotide database] -num_threads [x] -query
[input FASTA] -out [output] -max_target_seqs 500 -outfmt
"6 std staxids"
```

The resulting BLAST file (**Figure 2**) lists hits first by query sequence, so all hits to a query are together, and then by descending Expect value (*E*), so better matches are generally further up the list. However, within *E* value, the order is simply the order in which the hits occur in the database.

PIA

The PIA algorithm itself is computationally light enough to be run on a laptop with small sample files (FASTA $\sim <3$ MB). The index-building step required before first use should take no more than a few minutes. Time to analyze the seven samples used in this study on one thread ranged from approximately 10 s to 10 min. PIA can also be multi-threaded for larger samples, for which we recommend a server.

Figure 3 illustrates the PIA algorithm. PIA considers one read at a time. Reads with no BLAST hits are discarded. For reads with hits, PIA first calculates the coverage of the top hit:

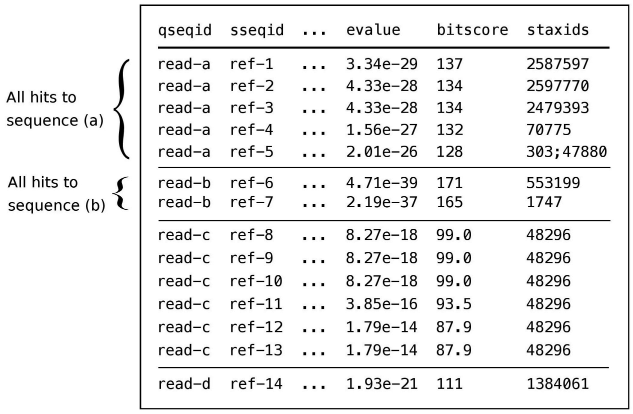
$$\%\ coverage = \frac{match\ length}{read\ length} \times 100$$

If the coverage does not meet a threshold (default 95%), the read is discarded. The taxonomic assignment of the read is strongly influenced by the top hit, so it only accepts a very close match.

PIA then considers each hit in order of the BLAST file. First, the hit is assigned to a taxon. If a hit is associated with multiple taxa, PIA assumes that this indicates a conservative sequence and assigns the hit to the phylogenetic intersection of those taxa. The assigned taxon is then evaluated. If there has already been a hit to the taxon, the hit is discarded. Because hits are listed in order of *E* value, this means that only the best hit for each taxon is retained. This taxon check aims to mitigate the problem of overrepresented taxa. Provided that the BLAST found enough hits to reach underrepresented taxa in the database at all, this check gives them equal weight to overrepresented taxa. Every taxon is reduced to a single hit.

The second check performed on each hit is the *E* value. If there has already been a hit that passed the taxon check with this *E* value, those hits are grouped together. Once all hits for this read have been taxon-checked and grouped by *E* value, the *E* value groups are collapsed to a single "hit" per *E* value. This "hit" is the phylogenetic intersection of the group members. If a read is found to be equally similar to sequences from several different organisms, PIA again assumes that this indicates a conservative sequence. Finally, if these new "hits" are to previously seen taxa, then as before, only the hit with the best *E* value is retained.

Once the list of BLAST hits for the read has been reduced to one (best) hit per taxon, PIA assigns the read to the phylogenetic intersection of the top and second-top hits. If only one hit remains, there cannot be an intersection, so the read is discarded. Finding the intersection firstly avoids over-assigning conservative sequences. Secondly, it avoids over-assigning reads from unrepresented taxa to represented relatives. PIA assumes that the real taxon is not in the database, so it will not assign directly to any organism in the database. The intersection is only taken between the top two hits because, after the taxon check and grouping by *E* value, those two hits may already be to distantly-related and/or high taxa.



```
qseqid  sseqid  ...  evalue    bitscore  staxids
read-a  ref-1   ...  3.34e-29  137       2587597
read-a  ref-2   ...  4.33e-28  134       2597770
read-a  ref-3   ...  4.33e-28  134       2479393
read-a  ref-4   ...  1.56e-27  132       70775
read-a  ref-5   ...  2.01e-26  128       303;47880
read-b  ref-6   ...  4.71e-39  171       553199
read-b  ref-7   ...  2.19e-37  165       1747
read-c  ref-8   ...  8.27e-18  99.0      48296
read-c  ref-9   ...  8.27e-18  99.0      48296
read-c  ref-10  ...  8.27e-18  99.0      48296
read-c  ref-11  ...  3.85e-16  93.5      48296
read-c  ref-12  ...  1.79e-14  87.9      48296
read-c  ref-13  ...  1.79e-14  87.9      48296
read-d  ref-14  ...  1.93e-21  111       1384061
```

FIGURE 2 | Example partial BLAST output structure in format "6 std staxids". The standard (std) fields are the first columns, starting with query sequence (qseqid) and ending with Expect (*E*) value (evalue) and score (bitscore). Additional fields, here the taxonomic IDs (staxids) associated with the reference, are at the end. Each row is a hit between the query sequence and a reference sequence from the database. Hits are ordered first by query sequence, then by *E* value from lowest to highest.

The final step is the diversity check, which filters reads by taxonomic diversity score:

$$\textit{Taxonomic diversity score} = \frac{t-1}{c}$$

Where t is the number of different taxa in the original list of BLAST hits and c is a predefined cap on the number of hit taxa to consider. The score measures how populated this area of the database is. A well-populated region will have more hits. If the region is sparsely-populated, there may be a disproportionately high number of hits to oasis taxa. Reads which seem to match an organism in a too sparsely-populated area are discarded.

METHODS

Analysis of Empirical sedaDNA Data

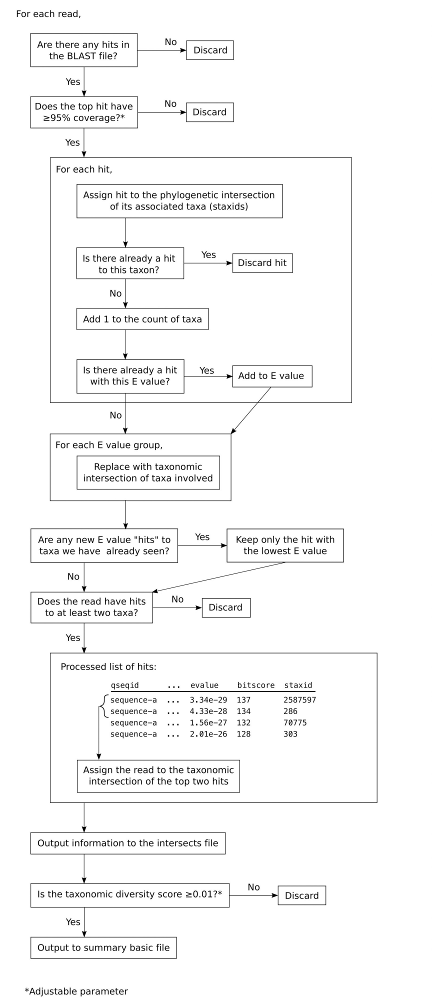
PIA and MEGAN were compared in a parallel analysis of seven samples from the Europe's Lost Frontiers project (Gaffney et al., 2020). These samples are from sediment core ELF039 which was taken from a palaeochannel approximately 50 km north of the present Norfolk coast. No dates are available for that core at the time of writing, but the channel is interpreted as a river valley that underwent marine inundation during the early Holocene. The samples were shotgun sequenced on a NextSeq 550 as part of our work using sedaDNA for palaeoenvironmental reconstruction. We typically focus on plants because of their high biomass in most environments, increasing the chance of DNA deposition, and the abundance of ecological and distribution information available. Accordingly, this study made use of reads from Viridiplantae.

Raw FASTQ files were adapter-trimmed and collapsed in AdapterRemoval 2.2.2 (Lindgreen, 2012), converted to FASTA, and had duplicates removed using fastx_collapser from the FAST-X Toolkit 0.0.13 (Gordon and Hannon, 2010). Then an initial BLAST was performed against the full nucleotide GenBank database (downloaded on 05-09-2019) using blastn 2.6.0 (Zhang et al., 2000) with -outfmt "6 std staxids" and -num_alignments 10. Output format six is tabular, reducing file size, and reference sequence taxonomic IDs were included to allow full parsing by MEGAN. In format 6, -num_alignments states the maximum number of hits per query. Ten was sufficient for this stage. An RMA file was generated from that BLAST output using the MEGAN5 command line interface with default settings (Huson et al., 2016). Reads assigned to Viridiplantae or below were extracted to a new FASTA. This FASTA was then BLASTed more thoroughly, with -max_target_seqs set to 500 to give up to approximately 500 hits per read.

For the MEGAN analysis, an RMA file was again generated from this final BLAST output using the default settings. All nodes were exported to a text file in the format "taxonID_to_count". The BLAST output and corresponding FASTA were also run through PIA. A custom script[1] (see **Supplementary Material**) was then used to filter both sets of output by a negative control: taxa with a control:sample hit ratio of at least 0.02 were discarded from the sample data. The control is the sum of all negative controls in the wider sequencing run of 142 samples from the same project. The seven filtered sample files were concatenated together and visualized with Krona (Ondov et al., 2011; see **Supplementary Material**).

---

[1] https://github.com/Allaby-lab/PIA-accessories



```
qseqid      ...  evalue    bitscore  staxid
sequence-a  ...  3.34e-29  137       2587597
sequence-a  ...  4.33e-28  134       286
sequence-a  ...  1.56e-27  132       70775
sequence-a  ...  2.01e-26  128       303
```

FIGURE 3 | Flowchart illustrating the PIA algorithm. There are three key checks that may result in a read being discarded: sufficient coverage of the top BLAST hit, at least two hits remaining after processing, and a high enough taxonomic diversity score. Reads that pass are assigned to the intersection of the top two remaining BLAST hits.

Accuracy Testing With Simulated Data

Benchmarking against MEGAN suggested that PIA may successfully increase the accuracy of taxonomic assignations at the cost of sensitivity. To test the accuracy more objectively, we ran both MEGAN and PIA on two test datasets of known GenBank sequences. For each dataset, the control condition used the original BLAST database from the benchmarking analysis (downloaded on 05-09-2019). An "exclusion" condition excluded all taxa in the test dataset from the BLAST database. This aimed to simulate the unrepresented taxa, common in metagenomic data, that PIA is designed to analyze. In each condition, we tracked the assignations of individual sequences and compared them to the actual source organisms. Most stages involved custom scripts available from https://github.com/Allaby-lab/PIA-accessories and detailed in the **Supplementary Material**.

Each test dataset comprised 250 GenBank sequences downloaded through the NCBI website. For the first dataset, sequences were first filtered to Embryophyta and to a length of 30–150 bp to reflect typical aDNA. We then iterated through "All other taxa" from the "Results by taxon" option until taxa were represented by no more than 44 relevant sequences. Metagenomic data is likely to contain poorly-represented organisms. Single sequences from 245 taxa were downloaded as a FASTA with GIs included. An additional five 30–150 bp sequences were added from well-represented domesticates: *Hordeum vulgare, Musa acuminata*, *Triticum dicoccon*, *Triticum aestivum*, and *Zea mays*. These were run through BLASTn to check that they did match their taxa labels, as model organism sequences are frequently assigned to incorrect taxa. The second dataset was constructed in a similar way, but first filtered to Mammalia instead of Embryophyta. The low-frequency taxa were represented by up to 47 relevant sequences and the five high-frequency taxa were *Camelus bactrianus*, *Camelus dromedarius*, *Balaenoptera bonaerensis*, *Chlorocebus aethiops*, and *Papio anubis*. Finally, each FASTA file was re-formatted to single-line using fasta_formatter from the FAST-X toolkit 0.0.13 (Gordon and Hannon, 2010). The final FASTAs are included as **Supplementary Data Sheets S2, S3** in the **Supplementary Material**.

The FASTAs were run through BLAST with the same settings as in benchmarking. The exclusion condition only differed in the reduced database. For every taxon, a list of GIs for all sequences from that taxon was downloaded from GenBank. These lists were concatenated into a master GI list. The BLAST option "-negative_gilist" was used to exclude this list from the database. For each BLAST file, the MEGAN and PIA analyses were performed with the same settings as in benchmarking. See the **Supplementary Material** for details.

It became apparent after analysis that two Mammalia sequences may be affected by human contamination: GI 2198752 (accession no. U84666.1, *Cavia porcellus* Y5 scRNA gene, partial sequence) and GI 13508496 (accession no. AY028924.1, *Mammut americanum* 16S ribosomal RNA gene, partial sequence; mitochondrial gene for mitochondrial product). We ran BLAST on both sequences to check, changing "-max_target_seqs 500" to "-num_alignments 1" to produce easily readable output with the default limit of 500 hits. Other settings were the same as in benchmarking.

Finally, a small separate test of GenBank data was used to evaluate the performance of PIA on highly divergent taxa.

Because of the diversity check, we expect PIA to unnecessarily discard reads assigned to taxa with few living relatives because their region of the database will always appear incomplete. We ran BLAST and PIA on the available GenBank sequences from two monotypic orders: Ginkgoales (containing the gymnosperm *Ginkgo biloba*; 22,600 sequences) and Microbiotheria (containing the marsupial *Dromiciops gliroides*; 417 sequences). This used the same settings as in benchmarking.

RESULTS

Analysis of Empirical sedaDNA Data

Taxonomic assignations of early Holocene sedaDNA from a submerged palaeochannel in the North Sea by MEGAN and PIA are compared in **Figures 4A,B**. The most frequent taxa are labeled in full. Of these, taxa not native to Europe are highlighted in bold (see below). The original interactive HTML chart is included as **Supplementary Data Sheet S4** in the **Supplementary Material**.

The taxonomic profiles of the MEGAN and PIA outputs are broadly similar (**Figures 4A,B**). **Figure 4** begins at Mesangiospermae, to which the vast majority of reads are assigned by both methods. Most reads are assigned to *Zostera marina* (eelgrass), related taxa in Potamogetonaceae or to its parent order Alismatales, suggesting a wetland or fully aquatic environment with at least some saltwater influence. There is also a sizeable signal from grasses (Poaceae). In the largest remaining segment, Pentapetalae (**Figure 4B**), both profiles show a diverse range of taxa found in northwest Europe today. This includes Rosaceae (strawberry, bramble, apple, drupe trees), *Salix* (willow), *Populus* (poplar), and Fagales (birch, oak).

However, the numbers of reads making up these taxa differ significantly. Though proportionally similar, the MEGAN profile was built from 88,497 reads compared to just 27,547 accepted by PIA. The MEGAN profile also has higher taxonomic richness, containing 374 taxa versus 210 (**Table 1**). Those MEGAN taxa are also generally more specific. MEGAN assigned far more reads to genus or lower. Overall, the results are consistent with MEGAN placing more emphasis on sensitivity than PIA.

Because the samples originate from northwest Europe in the early Holocene, we would expect DNA sequences to be comparable to European taxa today. The samples have been filtered by negative controls which should have removed most assignations to common modern contaminant taxa present in reagents. We therefore assume any assignations to non-European taxa to be false positives.

Many of the most frequent non-European taxa assigned to by MEGAN are domesticated grasses such as *Oryza*, *Setaria italica* and *Sorghum bicolor* (**Figure 4A**). In Pentapetalae (**Figure 4B**), most of the terminal taxa in the MEGAN output – those genera and species that suggest a higher sensitivity than PIA – are non-European and therefore likely false positives. **Table 1** quantifies all assignations: 40.11% of taxa in the MEGAN profile are suspect compared to 20.95% for PIA. In total, MEGAN assigned 12.78% of reads to non-European taxa and PIA assigned just 0.52%. The false positive taxa have lower counts on average, suggesting that the minimum support filter in MEGAN is a valid approach, but in this case PIA was more effective at removing this sort of false positive.

It appears that the lower sensitivity of PIA is associated with higher accuracy. To investigate this more objectively, we ran PIA on test sequences of known origin.

Accuracy Testing With Simulated Data

Embryophyta

Individual reads, their source organism and all four assignations are listed in the first worksheet of **Supplementary Table S2**. **Table 2** provides a summary. We considered an assignation correct if it was to the actual taxon or one of its parent taxa. For example, if PIA assigned a read from *Betula* to the family Betulaceae, it would be a correct assignment at family level. Family level is typically precise enough to be useful for environmental reconstruction in plants. An assignment to Viridiplantae would be correct at kingdom level. An assignment to Poaceae would be incorrect.

In the control condition, MEGAN assigned 91% of sequences and PIA 52%, mirroring the higher sensitivity of MEGAN observed in the analysis of real data. Both were highly accurate at 97 and 100%, respectively. MEGAN was somewhat more precise, with 62% of assignments correct to family level or below, compared to 53.49% for PIA. Overall, MEGAN showed a much greater ability to assign sequences at the cost of a very small drop in accuracy compared to PIA.

The exclusion condition, where the source taxa had been removed from the database, shows a similar pattern of results with generally worse performance by both tools. However, MEGAN appears to suffer more. The "Change" columns in **Table 2** show that MEGAN assigns proportionally fewer sequences at all, correctly, and with precision than PIA. Notably, accuracy of MEGAN falls to 80% but that of PIA remains at a healthy 96%.

Despite the exclusion database generally presenting more of a challenge, there were a small number of sequences that were assigned better than with the complete database. PIA did not assign the *Lapageria rosea* and *Lupinus luteus* sequences in the control condition but matched MEGAN's broad Mesangiospermae assignment for the exclusion. Both MEGAN and PIA assigned the *Metasequoia glyptostroboides* and *Magnolia x soulangeana* sequences more precisely in the exclusion condition, although not particularly so. This unexpected behavior may be due to peculiarities of the database around those sequences.

Mammalia

Full results are listed in the second worksheet of **Supplementary Table S2**. **Table 3** provides a summary. In the control condition, the Mammalia dataset showed a similar pattern to Embryophyta. MEGAN assigned more reads and with more precision; both programs were very accurate. The exclusion condition resulted in worse performance for both programs, again with a greater impact on MEGAN. However, the decrease in accuracy was even more pronounced than for Embryophyta. MEGAN only assigned

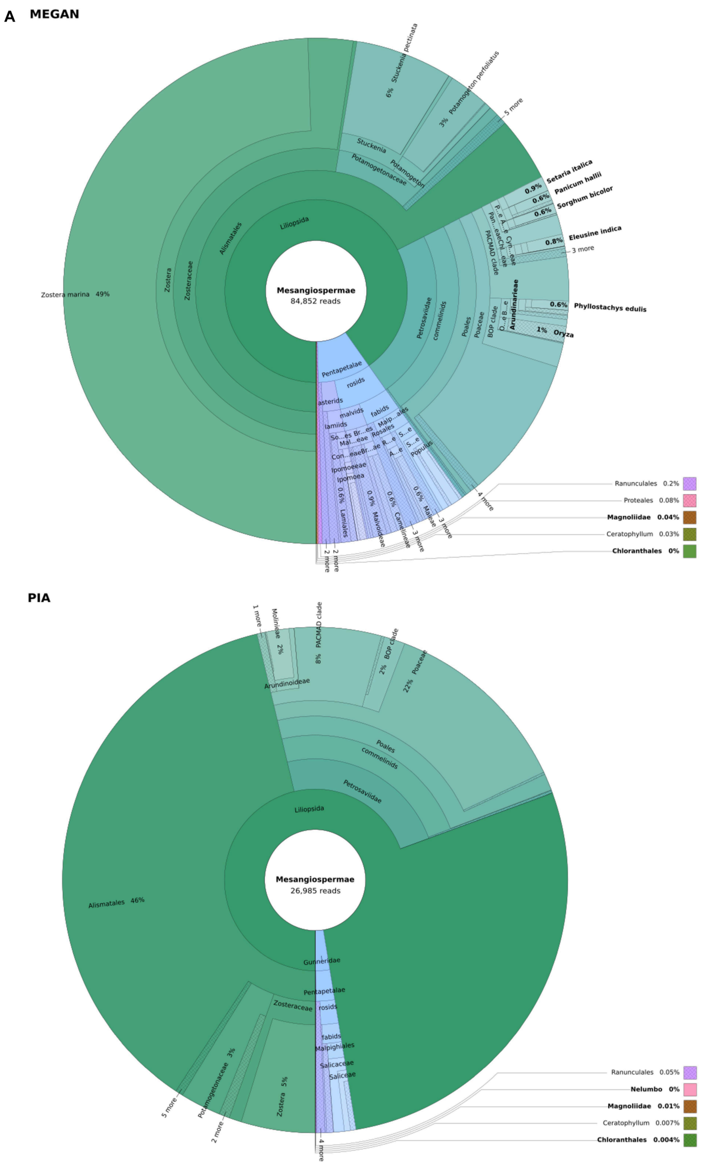


FIGURE 4 | Continued

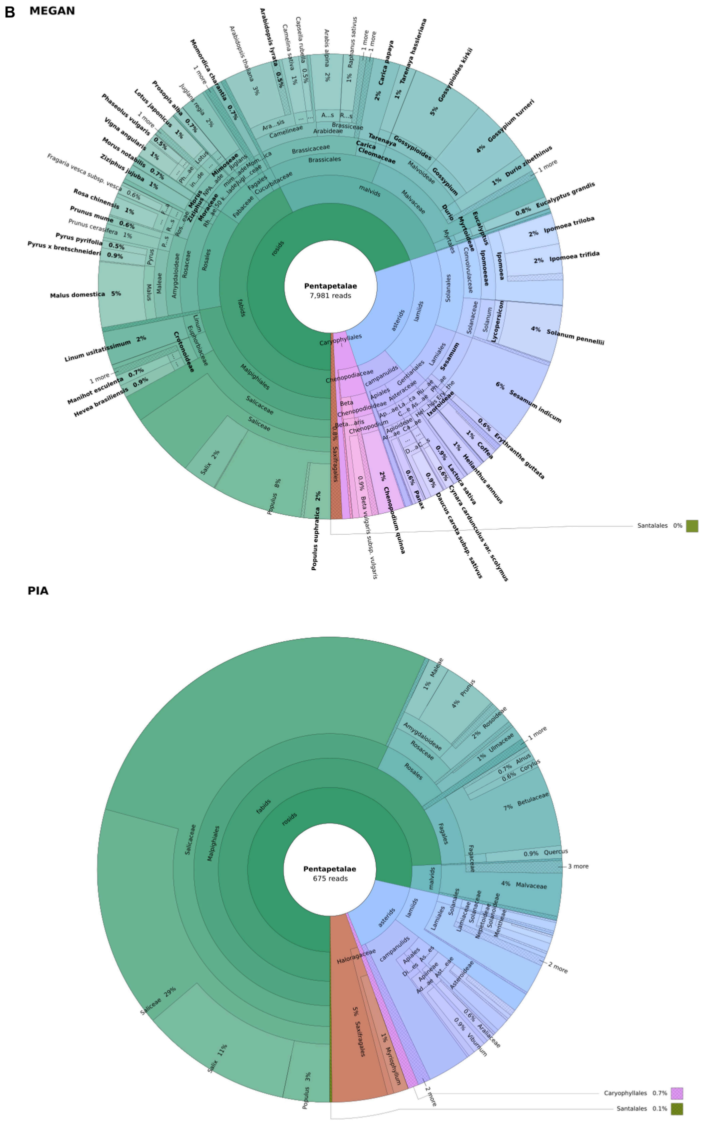


FIGURE 4 | Taxonomic profiles of the combined MEGAN and PIA outputs for the seven sediment samples, after filtering each by negative controls. **(A)** shows Mesangiospermae, which includes the vast majority of reads. **(B)** zooms in on Pentapetalae, the largest segment in panel **(A)** that cannot easily be seen. Taxa not native to Europe, which are suspected to be false positive assignments for this data, are highlighted in bold. Colors indicate taxon frequency. See **Supplementary Data Sheet S4** in the **Supplementary Material** for the original interactive HTML chart, which was produced using Krona (Ondov et al., 2011).

TABLE 1 | Numbers of European and non-European taxa hit and the numbers of reads assigned to each category in the MEGAN and PIA benchmarking output.

	MEGAN	PIA
Total taxa	374	210
European	224 (59.89%)	166 (79.05%)
Non-European	150 (40.11%)	44 (20.95%)
Total reads	88,497	27,547
To European taxa	77,189 (87.22%)	27,405 (99.48%)
To non-European taxa	11,308 (12.78%)	142 (0.52%)

Reads assigned to non-European taxa are suspected to be false positives for this data.

60% of sequences accurately. PIA assigned 83% accurately, which while better, is far from the 96% accuracy seen for Embryophyta.

Note that these accuracy results are likely a slight underestimate, as the two questionable sequences (to *Cavia porcellus* and *Mammut americanum*), do indeed appear to be mislabeled. Both BLAST outputs are dominated by *Homo sapiens* and other primates. MEGAN and PIA generally assigned them either to high mammal taxa or close parent taxa of humans, both of which are reasonable if the sequences are actually human.

As with Embryophyta, a small number of sequences were assigned better with their taxa excluded from the database. MEGAN assigned the *Stenella attenuata* sequence incorrectly in the control but broadly correct after exclusion. PIA assigned the *Kogia sima* sequence more precisely after exclusion, though only by one level. Finally, the only time the *Halichoerus grypus* sequence was assigned was by PIA after exclusion, and it did so correctly to family.

Monotypic Taxa

Phylogenetic Intersection Analysis assigned 5% of reads from Ginkgoales and with only 77% accuracy. For Microbiotheria, PIA assigned 37% of reads; 100% were accurate but the most precise was only to Metatheria. The proportion of reads assigned to each was considerably lower than the ∼50–60% from the mixed test datasets above.

DISCUSSION

Ancient metagenomics has much potential, but taxonomic assignation of reads can be improved. Databases are highly uneven, resulting in the joint problems of over-represented taxa filling up hit lists at the expense of poorly-represented but closer matches, and oasis taxa in sparsely-populated areas drawing in reads and giving an illusion of confident assignation. There are methods that partly address these problems in some circumstances, but we demonstrate here that PIA performs strongly, providing an objective approach to remove false positives from data sets.

Benchmarking on plant sedaDNA data against a standard tool, MEGAN, showed that PIA produces a comparatively low-resolution taxonomic profile. Far fewer reads are assigned and those that are rarely make it to genus. However, we argue that

TABLE 2 | Percentages of the 250 sequences assigned by MEGAN and PIA in the Embryophyta accuracy test.

Embryophyta	Control BLAST		Exclusion BLAST		Change	
	MEGAN	PIA	MEGAN	PIA	MEGAN	PIA
Assigned	91.20%	51.60%	76.00%	45.60%	−15.20%	−06.00%
Incorrect	03.07%	00.00%	20.00%	04.39%	16.93%	04.39%
Correct	96.93%	100.00%	80.00%	95.61%	−16.93%	−04.39%
Correct to above family	35.09%	46.51%	46.84%	60.53%	11.75%	14.02%
Correct to family or below	61.84%	53.49%	33.16%	35.09%	−28.68%	−18.40%

The control condition BLASTed against the full GenBank nucleotide database (downloaded on 05-09-2019). The exclusion condition omitted the source taxa from the database. Of those reads assigned, percentages assigned incorrectly or correctly are given. The final two rows detail whether correctly-assigned reads were assigned to higher taxa or to at least family. These rows sum to the total percent correct.

TABLE 3 | Percentages of the 250 sequences assigned by MEGAN and PIA in the Mammalia accuracy test.

Mammalia	Control BLAST		Exclusion BLAST		Change	
	MEGAN	PIA	MEGAN	PIA	MEGAN	PIA
Assigned	93.60%	57.60%	76.40%	52.40%	−17.20%	−05.20%
Incorrect	02.99%	00.00%	40.31%	16.79%	37.32%	16.79%
Correct	97.01%	100.00%	59.69%	83.21%	−37.32%	−16.79%
Correct to above family	28.21%	45.14%	41.36%	49.62%	13.36%	4.48%
Correct to family or below	68.80%	54.86%	18.32%	33.59%	−50.48%	−21.27%

The control condition BLASTed against the full GenBank nucleotide database (downloaded on 05-09-2019). The exclusion condition omitted the source taxa from the database. Of those reads assigned, percentages assigned incorrectly or correctly are given. The final two rows detail whether correctly-assigned reads were assigned to higher taxa or to at least family. These rows sum to the total percent correct.

much of the sensitivity of MEGAN in this context is oversensitivity. Both methods describe core ELF039 as coming from a primarily wetland environment, with a clear signal from fresh and saltwater plants in Alismatales and the riverine *Salix*, along with some signal from grasses in Poaceae and woodland trees in Fagales. Yet the MEGAN profile assigned nearly 13% of reads to clearly questionable taxa, such as the tropical *Sorghum bicolor*, Australasian *Eucalyptus* and American *Carica papaya*, that if taken at face value would present a radical departure from the established palaeoecology of Europe. Once such taxa are removed as "known" false positives, the MEGAN analysis only retrieves a few more taxa than PIA (**Figure 4B**), which add little to the palaeoecological reconstruction and likely still contain false positives. One example is *Arabidopsis thaliana*, a known model organism not expected to feature greatly in the Mesolithic. In our context, the additional accuracy of PIA appears to outweigh the increased sensitivity of MEGAN.

The accuracy test on simulated data returned similar results. With a full BLAST database, MEGAN assigned nearly twice as many sequences with greater precision and only marginally lower accuracy than PIA. However, when the source taxa were excluded from the database, exacerbating the problems caused by incomplete databases and better representing real metagenomic data, the improvements of MEGAN over PIA diminished and the difference in accuracy became substantial. For Embryophyta sequences, PIA maintained a very high accuracy of 96%, whereas that of MEGAN fell to 80%.

Both programs performed less well with the Mammalia dataset, but PIA still returned 83% accuracy after exclusion of source taxa compared to 60% from MEGAN. We suspect that this difference may simply be due to the fact that there are far fewer species of mammal than embryophyte, so removing 250 mammal taxa will have removed proportionally more of the relevant database than removing the same number from Embryophyta. Both PIA and MEGAN performed very well in the control condition, so it is unlikely to be directly due to the mammal sequences themselves. Instead, we suggest that the exclusion condition simulated a more incomplete database for Mammalia than Embryophyta. PIA still outperformed MEGAN. However, it is clear that while PIA copes better with incomplete databases, it is not a perfect solution.

Additionally, two specific limitations of PIA are apparent from its algorithm. First, PIA cannot assign to leaf taxa. It can only assign to a species if there are subspecies in the database, for example. PIA does not fully take advantage of sequences with very good taxonomic resolution. If better resolution is desired, it may be helpful to first identify reads to higher taxa more accurately using PIA, then further analyze any sequences assigned to taxa of interest using a different approach.

The second limitation is a result of the taxonomic diversity check. PIA discards assignments to taxa in sparse areas of the database because these areas are vulnerable to the influence of oasis taxa. However, this assumes that sparsity is due to incompleteness. There are divergent taxa with very few living relatives that will occupy a naturally sparse database region. PIA is less likely to accept assignments to these taxa. To demonstrate this, we ran PIA on the available GenBank sequences from Ginkgoales and Microbiotheria, which are orders containing a single species. PIA assigned fewer reads from these taxa than from the mixed Embryophyta or Mammalia datasets. Such divergent taxa are unusual, but are less likely to be recovered by PIA. Again, PIA shows a lack of sensitivity that may limit its application in some studies.

However, even with these caveats, we have demonstrated that the improved ability of PIA to address the challenges of an incomplete reference database can result in highly accurate taxonomic assignation of metagenomic shotgun data. PIA produced fewer false positives than the standard approach. The more likely false positives are to occur, the more necessary it becomes to manually sort taxa into plausible and implausible, which requires subjective presuppositions about the source of the data. This is particularly problematic for ancient metagenomics where little is known about the study environment. PIA offers an objective alternative with an estimated 96% accuracy for plants.

DATA AVAILABILITY STATEMENT

The datasets analyzed for this study can be found in the European Nucleotide Archive under the project code PRJEB33717. See **Supplementary Table S1** for sample accession codes.

AUTHOR CONTRIBUTIONS

RA, OS, RW, and BC wrote and designed the PIA. BC performed benchmarking and accuracy testing with some input from RW. BC was the primary author of the manuscript with review and editing by RA and RW. VG was the Principal Investigator of the project through which the sedaDNA dataset was obtained.

FUNDING

The sedaDNA dataset used for benchmarking was generated as part of the Europe's Lost Frontiers project. This received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (ERC funded project no. 670518 LOST FRONTIER, https://europa.eu/european-union/index_en, https://lostfrontiers.teamapp.com/).

ACKNOWLEDGMENTS

Thanks to the Europe's Lost Frontiers Project for supplying the sedaDNA data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.00084/full#supplementary-material

REFERENCES

Appelt, S., Fancello, L., Le Bailly, M., Raoult, D., Drancourt, M., Desnues, C., et al. (2014). Viruses in a 14th-Century Coprolite. *Appl. Environ. Microbiol.* 80, 2648–2655. doi: 10.1128/AEM.03242-13

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., et al. (2016). GenBank. *Nucleic Acids Res.* 45, D37–D42. doi: 10.1093/nar/gkw1070

Birks, H. J. B., and Birks Hilary, H. (2015). How have studies of ancient DNA from sediments contributed to the reconstruction of Quaternary floras? *New Phytol.* 209, 499–506. doi: 10.1111/nph.13657

Bon, C., Berthonaud, V., Maksud, F., Labadie, K., Poulain, J., Artiguenave, F., et al. (2012). Coprolites as a source of information on the genome and diet of the cave hyena. *Proc. R. Soc. B Biol. Sci.* 279, 2825–2830. doi: 10.1098/rspb.2012.0358

Der Sarkissian, C., Pichereau, V., Dupont, C., Ilsøe, P. C., Perrigault, M., Butler, P., et al. (2016). Ancient DNA analysis identifies marine mollusc shells as new metagenomic archives of the past. *Mol. Ecol. Resour.* 17, 835–853. doi: 10.1111/1755-0998.12679

Gaffney, V., Fitch, S., Bates, M., Ware, R. L., Kinnaird, T., Gearey, B., et al. (2020). Multi-proxy evidence for the impact of the Storegga Slide Tsunami on the early Holocene landscapes of the southern North Sea. *BioRxiv* [Preprint]. doi: 10.1101/2020.02.24.962605

Gilbert, M. T. P., Bandelt, H.-J., Hofreiter, M., and Barnes, I. (2005). Assessing ancient DNA studies. *Trends Ecol. Evol.* 20, 541–544. doi: 10.1016/j.tree.2005.07.005

Gordon, A., and Hannon, G. J. (2010). *Fastx-toolkit*. Cold Spring Harbor, NY: Hannon Laboratory.

Herbig, A., Maixner, F., Bos, K. I., Zink, A., Krause, J., Huson, D. H., et al. (2016). MALT: fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *bioRxiv* [Preprint]. doi: 10.1101/050559

Huson, D. H. (2019). *User Manual for MEGAN V6.17.0. 0–74.* Available online at: http://ab.inf.uni-tuebingen.de/data/software/megan6/download/manual.pdf (accessed September 26, 2019)

Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Res.* 17, 377–386. doi: 10.1101/gr.5969107

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., et al. (2016). MEGAN community edition – interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* 12:e1004957. doi: 10.1371/journal.pcbi.1004957

Key, F. M., Posth, C., Krause, J., Herbig, A., and Bos, K. I. (2017). Mining metagenomic data sets for ancient DNA: recommended protocols for authentication. *Trends Genet.* 33, 508–520. doi: 10.1016/j.tig.2017.05.005

Kistler, L., Ware, R., Smith, O., Collins, M., and Allaby, R. G. (2017). A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Res.* 45, 6310–6320. doi: 10.1093/nar/gkx361

Kuch, M., Rohland, N., Betancourt, J. L., Latorre, C., Steppan, S., and Poinar, H. N. (2002). Molecular analysis of a 11 700-year-old rodent midden from the Atacama Desert. *Chile. Mol. Ecol.* 11, 913–924. doi: 10.1046/j.1365-294x.2002.01492.x

Lindgreen, S. (2012). AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res. Notes* 5:337. doi: 10.1186/1756-0500-5-337

Ondov, B. D., Bergman, N. H., and Phillippy, A. M. (2011). Interactive metagenomic visualization in a web browser. *BMC Bioinformatics* 12:385. doi: 10.1186/1471-2105-12-385

Renaud, G., Schubert, M., Sawyer, S., and Orlando, L. (2019). "Authentication and assessment of contamination in ancient DNA," in *Ancient DNA: Methods and Protocols*, ed. B. Shapiro, A. Barlow, P. D. Heintzman, M. Hofreiter, J. L. A. Paijmans, and A. E. R. Soares (New York, NY: Springer), 163–194. doi: 10.1007/978-1-4939-9176-1_17

Shah, N., Nute, M. G., Warnow, T., and Pop, M. (2018). Misunderstood parameter of NCBI BLAST impacts the correctness of bioinformatics workflows. *Bioinformatics* 35, 1613–1614. doi: 10.1093/bioinformatics/bty833

Shapiro, B., Barlow, A., Heintzman, P. D., Hofreiter, M., Paijmans, J. L. A., and Soares, A. E. R. (eds). (2019). *Ancient DNA Methods and Protocols*, 2nd Edn. New York, NY: Humana Press. doi: 10.1007/978-1-4939-9176-1

Smith, O., Momber, G., Bates, R., Garwood, P., Fitch, S., Pallen, M., et al. (2015). Sedimentary DNA from a submerged site reveals wheat in the British Isles 8000 years ago. *Science* 347, 998–1001. doi: 10.1126/science.1261278

Sønstebø, J. H., Gielly, L., Brysting, A. K., Elven, R., Edwards, M., Haile, J., et al. (2010). Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. *Mol. Ecol. Resour.* 10, 1009–1018. doi: 10.1111/j.1755-0998.2010.02855.x

Stahlschmidt, M. C., Collin, T. C., Fernandes, D. M., Bar-Oz, G., Belfer-Cohen, A., Gao, Z., et al. (2019). Ancient mammalian and plant DNA from late quaternary stalagmite layers at Solkota Cave. *Georgia. Sci. Rep.* 9:6628. doi: 10.1038/s41598-019-43147-0

Warinner, C., Speller, C., and Collins, M. J. (2015). A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. *Philos. Trans. R. Soc. B Biol. Sci.* 370:20130376. doi: 10.1098/rstb.2013.0376

Weyrich, L. S., Duchene, S., Soubrier, J., Arriola, L., Llamas, B., Breen, J., et al. (2017). Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature* 544, 357–361. doi: 10.1038/nature21674

Willerslev, E., Cappellini, E., Boomsma, W., Nielsen, R., Hebsgaard, M. B., Brand, T. B., et al. (2007). Ancient biomolecules from deep ice cores reveal a forested southern Greenland. *Science* 317, 111–114. doi: 10.1126/science.1141758

Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. A. (2000). Greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7, 203–214. doi: 10.1089/10665270050081478

Zhou, Z., Luhmann, N., Alikhan, N.-F., Quince, C., and Achtman, M. (2018). "Accurate reconstruction of microbial strains from metagenomic sequencing using representative reference genomes," in *Research in Computational Molecular Biology*, ed. B. J. Raphael (Cham: Springer International Publishing), 225–240. doi: 10.1007/978-3-319-89929-9_15

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix B

DNA yields by sample

Table B.1: DNA yields for each sample per replicate.

	DNA yield (ug/ml) of replicate 1		DNA yield (ug/ml) of replicate 2	
Sample	After extraction	Before sequencing	After extraction	Before sequencing
ELF0022_140	<0.05	0.090	<0.05	<1.00
ELF0027_160	0.367	2.780	0.182	7.810
ELF0027_174	0.559	2.350	0.251	1.110
ELF0027_177	0.727	1.170	0.098	26.270
ELF0031 _33	0.252	6.370	0.395	17.830
ELF0031 _43	0.218	1.900	0.241	1.320
ELF0031 _56	0.799	12.900	2.070	9.910
ELF0031A_72	0.134	5.650	<0.05	7.480
ELF0031A_88	0.069	9.550	0.076	0.810
ELF0031A_107	0.124	2.920	0.127	<1.00
ELF0031A_123	0.059	1.560	<0.05	<1.00
ELF0031A_152	0.083	2.850	<0.05	0.680
ELF0031A_177	<0.05	9.910	<0.05	1.900
ELF0031A_202	0.377	17.200	0.280	8.950
ELF0031A_219	<0.05	0.958	<0.05	<1.00
ELF0031A_281	<0.05	0.358	<0.05	<1.00
ELF0031A_310	<0.05	1.070	<0.05	<1.00
ELF0032A_47	0.436	2.710	0.696	1.420
ELF0032A_74	0.638	2.790	0.550	1.730
ELF0032A_95	0.471	0.470	0.061	1.250
ELF0032A_117	2.360	0.456	0.080	4.630
ELF0032A_153	0.332	0.262	0.215	1.977
ELF0032A_177	0.560	0.222	0.238	2.233
ELF0033_46	0.358	3.930	0.563	16.880
ELF0033_75	0.339	1.670	1.530	10.210
ELF0033_118	0.510	0.059	0.729	4.180
ELF0033_155	0.681	1.030	1.090	5.240
ELF0033_176	0.981	1.100	0.521	2.160
ELF0033_183	0.510	0.493	0.130	10.090
ELF0033_187	4.760	1.580	1.860	8.040
ELF0033_195	1.190	2.070	0.481	<1.00
ELF0033_203	3.780	1.870	2.980	<1.00
ELF0033A_50	0.554	6.360	0.550	9.150
ELF0033A_70	0.132	4.010	0.424	4.850
ELF0033A_98	0.227	8.710	0.789	12.700
ELF0033A_111	0.124	6.570	5.110	7.060
ELF0033A_126	0.236	4.120	0.214	2.150
ELF0033A_158	2.720	8.490	1.510	2.060
ELF0034_61	0.540	4.990	0.847	12.830
ELF0034_79	0.155	0.371	0.087	0.840
ELF0034_94	0.142	1.440	0.279	8.270
ELF0034_132	0.202	11.700	3.600	13.500
ELF0034_157	0.186	0.635	0.184	<1.00
ELF0034_177	0.156	0.372	0.274	0.950
ELF0034_185	0.141	0.453	0.332	2.340
ELF0034_202	0.059	<1.00	1.010	<1.00
ELF0034_219	<0.05	<1.00	1.200	<1.00
ELF0034A_63	<0.05	0.285	0.134	3.397
ELF0034A_81	0.183	1.320	0.234	<1.00
ELF0034A_126	0.061	9.970	0.718	2.270
ELF0034A_146	0.111	4.750	0.249	11.910
ELF0034A_166	0.123	5.560	<0.05	<1.00
ELF0034A_172	0.194	0.434	0.796	4.620
ELF0034A_183	0.469	7.050	1.110	8.490
ELF0034A_195	0.116	3.470	0.204	18.700
ELF0034A_225	0.195	3.700	0.600	8.380
ELF0034A_261	0.063	5.190	0.171	30.030
ELF0034A_282	0.092	3.650	0.158	<1.00
ELF039_145	0.431	0.610	<0.05	2.913
ELF039_250	0.231	0.307	0.174	<1.00
ELF039_321	4.330	0.159	1.220	<1.00
ELF039_355	5.930	0.231	0.094	1.350
ELF039_384	2.280	0.192	0.071	<1.00
ELF039_415	4.400	0.550	<0.05	2.603
ELF039_460	0.381	0.184	<0.05	1.930
ELF039_485	<0.05	0.124	<0.05	<1.00

Continued on next page

Table B.1 continued

	DNA yield (ug/ml) of replicate 1		DNA yield (ug/ml) of replicate 2	
Sample	After extraction	Before sequencing	After extraction	Before sequencing
ELF040A_95	0.448	0.544	0.157	1.073
ELF040A_112	2.170	0.210	0.066	0.683
ELF040A_192	0.251	0.295	<0.05	1.657
ELF040A_208	0.235	0.324	0.332	1.383
ELF040A_298	0.177	2.460	0.155	4.937
ELF040A_350	0.369	5.550	0.266	4.567
ELF040A_487	0.098	6.250	0.230	3.307
ELF041_87	0.169	1.070	0.196	<1.00
ELF041_110	0.132	0.206	0.188	0.790
ELF041_180	0.213	1.360	0.201	<1.00
ELF041_295	0.160	0.094	<0.05	<1.00
ELF042_65	0.127	1.680	0.172	1.703
ELF042_151	0.206	14.600	0.180	1.413
ELF042_250	0.346	<0.05	0.117	0.913
ELF042_350	0.091	0.083	0.081	1.863
ELF044_90	<0.05	1.300	1.430	2.927
ELF044_137	1.610	4.420	0.960	2.970
ELF044A_97	0.523	0.111	<0.05	13.370
ELF045_90	<0.05	<0.05	0.450	10.100
ELF045_145	0.522	0.333	0.861	14.400
ELF045_252	0.230	0.214	0.191	<1.00
ELF045_346	0.375	0.229	0.298	1.470
ELF045_450	0.261	0.869	0.355	<1.00
ELF045_522	0.297	0.184	0.228	<1.00
ELF046A_270	<0.05	17.400	<0.05	<1.00
ELF047_70	0.287	0.450	0.164	1.450
ELF047_150	0.607	1.490	0.334	10.630
ELF047_241	0.158	0.200	0.279	0.850
ELF047_274	0.161	1.040	0.065	<1.00
ELF047_325	0.750	3.020	0.378	12.600
ELF047_386	0.796	1.510	0.156	10.800
ELF047A_50	0.142	0.154	0.267	<1.00
ELF047A_150	0.358	<0.05	0.281	<1.00
ELF047A_256	0.200	0.664	0.090	<1.00
ELF047A_354	0.127	0.446	0.319	<1.00
ELF049_295	0.623	0.548	0.429	19.170
ELF049_361	<0.05	0.118	<0.05	<1.00
ELF050_250	<0.05	<0.05	<0.05	<1.00
ELF050_354	<0.05	0.055	<0.05	<1.00
ELF050_450	<0.05	0.197	<0.05	<1.00
ELF050_595	<0.05	0.552	<0.05	<1.00
ELF051_96	0.219	0.433	0.213	<1.00
ELF051_120	0.101	1.210	0.140	2.810
ELF051_151	0.261	0.272	0.088	4.000
ELF051_196	0.426	0.673	0.237	13.700
ELF051_255	0.176	0.467	0.190	10.400
ELF051_292	0.432	1.150	0.120	<1.00
ELF053_179	0.073	6.490	<0.05	<1.00
ELF053_214	<0.05	0.642	<0.05	<1.00
ELF053_275	<0.05	1.450	<0.05	<1.00
ELF053_289	<0.05	0.551	<0.05	<1.00
ELF053_336	0.068	25.500	<0.05	<1.00
ELF054_58	0.682	1.140	1.100	3.390
ELF054_140	1.730	8.600	0.344	16.167
ELF054_182	0.374	25.700	0.124	4.330
ELF054_268	0.332	9.190	0.414	8.777
ELF054_291	0.480	2.580	0.328	1.250
ELF054_315	0.284	2.610	1.130	3.820
ELF054_330	0.370	0.165	1.740	1.663
ELF054_356	0.051	1.600	1.600	2.030
ELF059 _170	1.530	2.680	5.510	2.393
ELF059 _210	0.850	1.610	0.747	8.887
ELF059 _230	3.520	0.060	5.180	4.730
ELF059 _280	1.900	0.784	3.090	7.920
ELF059 _337	0.779	0.374	1.570	9.800
ELF059 _359	3.130	2.390	1.540	10.200
ELF059 _378	0.195	0.696	<0.05	19.800
ELF059A_135	0.079	0.521	0.256	7.837
ELF059A_190	1.510	0.118	2.110	7.763
ELF059A_250	2.550	0.288	0.103	5.973
ELF059A_320	2.470	9.210	0.290	7.310
ELF059A_355	0.071	0.208	1.070	5.587
ELF060_250	0.608	0.222	2.540	8.550
ELF060_350	0.672	5.420	0.604	5.187
ELF060_420	0.362	5.670	0.232	6.360
ELF060_465	0.823	3.520	1.420	6.940

Appendix C

Read counts at key stages of initial data analysis

Table C.1: Read counts of the 149 samples, by replicate and in total, at three key stages of initial data analysis. The raw count is the number of reads per R1 FASTQ file straight after demultiplexing. The pre-processed count is just before the initial BLAST. The PIA-assigned, negative-control-filtered counts for Viridiplantae and Metazoa (no Primates) are equal to those in the read count plots for each core in Chapter 4 (Taxonomic results). Note that the two additional replicates for sample ELF059A_355 are included on an additional row (ELF059_355*) to fit the table, and that second replicates from several samples were excluded after negative-control-filtering as described in Chapter 2 (Main materials and methods), section 2.2.2.

Table C.1

Sample	Raw reads (R1)			Pre-processed reads			PIA-assigned, negative-control-filtered Viridiplantae reads			PIA-assigned, negative-control-filtered Metazoa (no primates) reads		
	Replicate 1	Replicate 2	Total	Replicate 1	Replicate 2	Total	Replicate 1	Replicate 2	Total	Replicate 1	Replicate 2	Total
ELF001A_120	2174712	5948705	8123417	243064	549757	792821	8	16	24	377	11	388
ELF001A_126	4088858	8240682	12329540	460267	1149545	1609812	29	85	114	4	1	5
ELF001A_138	5326654	10026476	15353130	3593230	8103525	11696755	289	554	843	4	7	11
ELF001A_149	82	436	518	53	100	153	0	0	0	0	0	0
ELF003_332	108417400	48544454	156961854	59662214	29879928	89542142	13246	2671	15917	81	46	127
ELF022_140	23223	243208	266431	3829	97351	101180	0	2	2	0	3	3
ELF027_160	899985	7374689	8274674	404246	3012304	3416550	111	664	775	0	6	6
ELF027_174	1167532	35258776	36426308	611647	13289804	13901451	1977	100978	102955	0	22	22
ELF027_177	154328	1030188	1184516	98202	287906	386108	30	39	69	0	0	0
ELF031_33	99060	807632	906692	43211	233183	276394	0	0	0	0	0	0
ELF031_43	211755	830942	1042697	126982	384748	511730	6	10	16	0	3	3
ELF031_56	983487	4257672	5241159	628222	1725639	2353861	14	36	50	0	3	3
ELF031A_58	24232197	81300569	105532766	16160162	48658848	64819010	1379	6392	7771	47	144	191
ELF031A_72	237048	2909081	3146129	202000	2120383	2322383	298	7907	8205	0	3	3
ELF031A_88	269569	1671610	1941179	232988	914854	1147842	294	506	800	0	0	0
ELF031A_107	161082	898646	1059728	100871	426330	527201	61	188	249	0	14	14
ELF031A_123	144948	1055090	1200038	89403	439588	528991	70	705	775	1	2	3
ELF031A_152	153710	904659	1058369	115330	469245	584575	47	296	343	0	3	3
ELF031A_177	172622	994985	1167607	145648	739641	885289	55	221	276	0	2	2
ELF031A_202	219232	1039589	1258821	121989	615085	737074	12	44	56	0	1	1
ELF031A_219	92770	72638	165408	7093	13257	20350	0	0	0	0	0	0
ELF031A_281	115777	274616	390393	14411	69480	83891	5	1	6	0	1	1
ELF031A_310	83483	67068	150551	1312	9004	10316	0	0	0	0	0	0
ELF032A_47	723040	11586044	12309084	291906	4956861	5248767	9	2829	2838	1	6	7
ELF032A_74	2062198	38798425	40860623	834630	19117474	19952104	62	5376	5438	0	11	11
ELF032A_95	4737184	7259900	11997084	2979161	3402023	6381184	1486	2195	3681	3	3	6
ELF032A_117	20498732	8682787	29181519	7770623	4469467	12240090	1210	194	1404	4	1	5
ELF032A_153	10367316	6294431	16661747	3782185	3192582	6974767	2280	104	2384	1	1	2
ELF032A_177	15443892	825566	16269458	5288321	327724	5616045	6387	23	6410	75	2	77
ELF033_46	1013395	4839336	5852731	650839	3047247	3698086	4	28	32	0	2	2
ELF033_75	293550	6984709	7278259	231187	4894128	5125315	0	29	29	0	6	6
ELF033_118	3363086	6454356	9817442	2924939	4215722	7140661	323	2967	3290	3	6	9
ELF033_155	1625	12237539	12239164	1019	6060575	6061594	0	70	70	0	6	6
ELF033_176	3061614	3419928	6481542	2648816	2432386	5081202	133	80	213	1	2	3
ELF033_183	500984	1028965	1529949	389319	695687	1085006	14	46	60	0	0	0
ELF033_187	731564	1589189	2320753	444328	1044970	1489298	130	140	270	1	0	1
ELF033_195	323178	2954	326132	149538	1768	151306	54	0	54	0	0	0
ELF033_203	332444	1560	334004	153530	428	153958	48	0	48	0	0	0
ELF033A_50	864964	1634207	2499171	606989	1161268	1768257	1	7	8	0	2	2
ELF033A_70	636485	3396103	4032588	376523	2688240	3064763	1	25	26	0	0	0
ELF033A_98	687490	1466391	2153881	516264	1042697	1558961	18	31	49	0	3	3
ELF033A_111	1362171	3185314	4547485	758368	1205052	1963420	26	43	69	0	1	1
ELF033A_126	130170	872662	1002832	94826	640164	734990	1	6	7	0	0	0
ELF033A_158	1015023	507166	1522189	609633	403064	1012697	429	20	449	2	0	2
ELF034_61	234095	939426	1173521	151805	715992	867797	1	0	1	0	0	0
ELF034_79	113184	791265	904449	46699	473918	520617	0	0	0	0	0	0
ELF034_94	195340	2137085	2332425	138702	1886633	2025335	0	0	0	0	1	1
ELF034_132	381289	2007242	2388531	232043	1217935	1449978	1	1	2	0	1	1
ELF034_157	135265	124672	259937	44973	23779	68752	1	0	1	0	1	1
ELF034_177	87509	3668978	3756487	19711	3048025	3067736	1	3	4	0	0	0

Continued on next page

Table C.1 continued

	Raw reads (R1)			Pre-processed reads			PIA-assigned, negative-control-filtered Viridiplantae reads			PIA-assigned, negative-control-filtered Metazoa (no primates) reads		
Sample	Replicate 1	Replicate 2	Total	Replicate 1	Replicate 2	Total	Replicate 1	Replicate 2	Total	Replicate 1	Replicate 2	Total
ELF034_185	226688	1303599	1530287	46504	871671	918175	0	4	4	0	0	0
ELF034_202	1081787	1943711	3025498	394017	711439	1105456	22	10	32	1	1	2
ELF034_219	369598	725050	1094648	94319	219068	313387	0	15	15	1	4	5
ELF034A_63	21835189	14748981	36584170	9864707	7097452	16962159	1818	52	1870	4	0	4
ELF034A_81	2915846	2292643	5208489	985198	582879	1568077	496	32	528	6	10	16
ELF034A_126	559779	16246814	16806593	346771	8591132	8937903	1	492	493	1	6	7
ELF034A_146	668610	1516208	2184818	503643	1080228	1583871	0	3	3	1	1	2
ELF034A_166	584348	124903	709251	464395	20345	484740	2	10	12	0	0	0
ELF034A_172	103057	5652314	5755371	83540	3871098	3954638	3	2	5	0	0	0
ELF034A_183	279716	2415827	2695543	208121	1770571	1978692	5	2	7	0	0	0
ELF034A_195	181919	1830349	2012268	118732	1378322	1497054	1	1	2	0	0	0
ELF034A_225	241368	824484	1065852	114768	545893	660661	0	0	0	0	0	0
ELF034A_261	224806	1185928	1410734	149288	597084	746372	0	1	1	0	0	0
ELF034A_282	330164	292	330456	174971	112	175083	3	0	3	0	0	0
ELF039_145	115808	1075	116883	6510	193	6703	0	0	0	0	0	0
ELF039_250	3775080	20627807	24402887	2974932	7154184	10129116	5715	116759	122474	15	10	25
ELF039_321	27162425	34094703	61257128	11142003	8440155	19582158	7464	149722	157186	26	6	32
ELF039_341	41916668	11484101	53400769	4210751	2369671	6580422	176	58	234	41	6	47
ELF039_355	10825758	493933	11319691	7375527	306660	7682187	5331	16	5347	18	0	18
ELF039_384	6210040	178782	6388822	4637373	65074	4702447	366	3	369	14	0	14
ELF039_415	24077358	51639	24128997	11091338	29517	11120855	5819	1	5820	25	1	26
ELF039_460	21177442	22329	21199771	9508665	9236	9517901	2834	0	2834	31	0	31
ELF039_485	517407	250348	767755	197641	43239	240880	17	1	18	10	0	10
ELF040A_95	14630222	115	14630337	7259713	36	7259749	2242	0	2242	13	0	13
ELF040A_112	5433965	22	5433987	3781132	8	3781140	132	0	132	8	0	8
ELF040A_192	6119249	42681	6161930	4015943	17526	4033469	967	4	971	6	1	7
ELF040A_208	2932732	783532	3716264	2028476	291676	2320152	10	34	44	1	2	3
ELF040A_298	383179	181648	564827	275578	93660	369238	12	1450	1462	0	1	1
ELF040A_350	162738	164150	326888	116914	55843	172757	0	1449	1449	0	0	0
ELF040A_487	117155	12006	129161	86258	5618	91876	1	4	5	0	0	0
ELF041_87	135835	933212	1069047	61101	219820	280921	108	49	157	2	4	6
ELF041_110	58093	1382513	1440606	26493	322600	349093	7	617	624	1	22	23
ELF041_180	189005	998646	1187651	91488	298881	390369	166	1418	1584	0	5	5
ELF041_295	204608	1495051	1699659	69693	311969	381662	17	58	75	0	0	0
ELF042_65	4185682	944475	5130157	1915003	262550	2177553	8857	69	8926	23	2	25
ELF042_151	242851	841642	1084493	124509	220540	345049	340	1773	2113	1	0	1
ELF042_250	990984	1166420	2157404	735242	263174	998416	640	223	863	7	2	9
ELF042_350	599722	999339	1599061	437526	353603	791129	818	1868	2686	6	6	12
ELF044_90	5918708	NA	5918708	2898288	NA	2898288	7041	NA	7041	6	NA	6
ELF044_137	2498990	NA	2498990	1357108	NA	1357108	638	NA	638	0	NA	0
ELF044A_97	2569404	1281629	3851033	2090546	500663	2591209	13458	4806	18264	8	0	8
ELF045_90	16136	1747083	1763219	13078	597489	610567	8	1757	1765	0	1	1
ELF045_145	38545335	1727314	40272649	19246709	824883	20071592	28626	388	29014	24	1	25
ELF045_252	3123994	35660	3159654	1128843	14120	1142963	3810	0	3810	7	0	7
ELF045_346	4009194	120901	4130095	1284779	65377	1350156	1041	0	1041	5	0	5
ELF045_450	1053616	1795370	2848986	419665	463172	882837	458	61	519	3	4	7
ELF045_522	8864031	1355068	10219099	4235868	300381	4536249	186	1880	2066	69	4	73
ELF046A_270	834665	NA	834665	383221	NA	383221	0	NA	0	8	NA	8
ELF047_70	374463	1281946	1656409	294587	488656	783243	22	46	68	2	0	2
ELF047_150	2448651	1647606	4096257	1839101	439652	2278753	2695	2834	5529	7	1	8
ELF047_241	146363	635877	782240	111953	93525	205478	468	2235	2703	0	0	0
ELF047_274	107263	2157310	2264573	70807	775836	846643	0	407	407	0	4	4

Continued on next page

Table C.1 continued

	Raw reads (R1)			Pre-processed reads			PIA-assigned, negative-control-filtered Viridiplantae reads			PIA-assigned, negative-control-filtered Metazoa (no primates) reads		
Sample	Replicate 1	Replicate 2	Total	Replicate 1	Replicate 2	Total	Replicate 1	Replicate 2	Total	Replicate 1	Replicate 2	Total
ELF047_325	947656	1698175	2645831	643151	773925	1417076	2435	2105	4540	3	0	3
ELF047_386	877659	2775433	3653092	548459	1054496	1602955	578	1311	1889	0	1	1
ELF047A_50	2778558	2383273	5161831	1032414	667884	1700298	1928	41	1969	6	0	6
ELF047A_150	340714	277156	617870	131943	102474	234417	2	2	4	8	6	14
ELF047A_256	315767	3009628	3325395	152662	868562	1021224	1	19	20	1	6	7
ELF047A_354	1274201	2230196	3504397	543084	561397	1104481	8	9	17	7	4	11
ELF049_295	6957895	1858102	8815997	4598535	1095181	5693716	1034	216	1250	4	1	5
ELF049_361	306421	1258617	1565038	134813	389466	524279	47	93	140	2	3	5
ELF050_250	243257	149989	393246	64294	43155	107449	1	0	1	1	0	1
ELF050_354	492353	4234	496587	134893	606	135499	3	0	3	4	0	4
ELF050_450	191629	153025	344654	61646	30008	91654	9	0	9	2	0	2
ELF050_595	150831	109769	260600	68127	23114	91241	3	0	3	2	0	2
ELF051_96	577237	907941	1485178	364825	342108	706933	10	57	67	1	0	1
ELF051_120	1145268	512864	1658132	749755	397642	1147397	2	9	11	0	0	0
ELF051_151	826527	808587	1635114	678460	584561	1263021	7	17	24	2	2	4
ELF051_196	9855544	2486805	12342349	8789360	2148081	10937441	458	65	523	13	4	17
ELF051_255	12588734	1263593	13852327	10254213	937967	11192180	2737	50	2787	23	2	25
ELF051_292	8042652	3078365	11121017	2866727	1399470	4266197	7860	654	8514	3	6	9
ELF053_179	222977	105017	327994	168388	4559	172947	0	0	0	3	0	3
ELF053_214	222630	23899	246529	108967	3052	112019	3	0	3	9	0	9
ELF053_275	2060870	NA	2060870	1041205	NA	1041205	1	NA	1	22	NA	22
ELF053_289	328266	681214	1009480	148485	160164	308649	13	7	20	4	1	5
ELF053_336	641041	553907	1194948	209789	131279	341068	0	4	4	5	1	6
ELF054_58	12915417	1448319	14363736	5058108	961994	6020102	2211	36	2247	15	1	16
ELF054_140	1719234	1602492	3321726	679533	741714	1421247	81	6	87	2	1	3
ELF054_182	288959	902874	1191833	149462	644029	793491	7	460	467	0	1	1
ELF054_268	435132	1278860	1713992	238307	729947	968254	3	32	35	0	1	1
ELF054_291	4726612	NA	4726612	2705240	NA	2705240	25	NA	25	35	NA	35
ELF054_315	2729077	NA	2729077	2098323	NA	2098323	23	NA	23	0	NA	0
ELF054_330	2826912	NA	2826912	2203252	NA	2203252	12	NA	12	3	NA	3
ELF054_356	8347570	NA	8347570	4353156	NA	4353156	2	NA	2	2	NA	2
ELF059_210	7205859	3473551	10679410	4469746	1733046	6202792	4266	2008	6274	17	7	24
ELF059_230	2192054	13991581	16183635	1359779	3438159	4797938	90	1325	1415	3	2	5
ELF059_270	11105897	4207702	15313599	3949504	1705709	5655213	2111	479	2590	11	2	13
ELF059_280	1351989	5594414	6946403	1094459	1755270	2849729	28	80	108	3	2	5
ELF059_337	791478	4827240	5618718	504396	1343752	1848148	102	1120	1222	1	0	1
ELF059_359	2035602	1326634	3362236	1233485	479724	1713209	4439	499	4938	7	2	9
ELF059_378	530681	4284455	4815136	341892	1611807	1953699	25	113	138	1	4	5
ELF059A_135	14369331	2125313	16494644	8875496	1087881	9963377	24995	2340	27335	45	5	50
ELF059A_190	29959218	2004037	31963255	14210182	768211	14978393	18173	83	18256	42	2	44
ELF059A_250	33375305	3842781	37218086	15358572	1482420	16840992	20209	1126	21335	26	2	28
ELF059A_320	475747	2463321	2939068	344801	908998	1253799	75	33	108	0	1	1
ELF059A_355	9855433	146511	10001944	6662812	50111	6712923	3989	2	3991	9	1	10
ELF059A_355*	92491251	120	92491371	54574560	63	54574623	51431	0	51431	114	0	114
ELF060_250	4652490	4318014	8970504	2449333	1095551	3544884	1821	482	2303	0	0	0
ELF060_350	1062221	4868524	5930745	708850	2401665	3110515	2262	4004	6266	0	1	1
ELF060_420	697552	2482458	3180010	447690	1292602	1740292	797	5587	6384	0	1	1
ELF060_465	863247	6042879	6906126	538238	2171871	2710109	368	1532	1900	0	0	0

Appendix D

Taxa found in the Viridiplantae data

Table D.1: List of taxa found in the Viridiplantae data, sorted by full taxonomy. Includes NCBI taxonomic ID, full taxon name, common name, ecological category, informal taxon group (e.g. algae; liverworts), native status in Great Britain and Europe, and mean estimated C-value (haploid genome size). Native status was not necessary for taxa outside Embryophyta. Mean C-value was calculated for taxa in Viridiplantae only; due to the imprecise nature of initial taxonomic assignment, some taxa in this dataset are outside Viridiplantae. Non-native status is highlighted in gold. Mean C-values are shaded by size: values <1 are in blue and generate an increase from raw read count to biogenomic mass, and values >1 are in red and generate a decrease. Magnitude is shown by saturation. CV values ≥ 1 suggest imprecision and are highlighted in gold.

Table D.1

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
131567	cellular organisms		Mixed	Mixed	1	1	NA	NA
2157	Archaea		Mixed	Mixed	1	1	NA	NA
2	Bacteria		Mixed	Bacteria	1	1	NA	NA
48479	environmental samples		Mixed	Bacteria	1	1	NA	NA
1196022	unclassified Rhodothermaceae		Mixed	Bacteria	NA	NA	NA	NA
1224	Proteobacteria		Mixed	Bacteria	1	1	NA	NA
28211	Alphaproteobacteria		Mixed	Bacteria	1	1	NA	NA
356	Rhizobiales	Rhizobia order	Mixed	Bacteria	1	1	NA	NA
46913	Devosia		Mixed	Bacteria	NA	NA	NA	NA
119045	Methylobacteriaceae		Mixed	Bacteria	NA	NA	NA	NA
41297	Sphingomonadaceae		Mixed	Bacteria	1	1	NA	NA
28216	Betaproteobacteria		Mixed	Bacteria	1	1	NA	NA
135622	Alteromonadales		Mixed	Bacteria	1	1	NA	NA
28228	Colwellia		Salt/brackish aquatics	Bacteria	NA	NA	NA	NA
91347	Enterobacterales		Mixed	Bacteria	NA	NA	NA	NA
1783257	PVC group		Mixed	Bacteria	1	1	NA	NA
203682	Planctomycetes		Mixed aquatics	Bacteria	NA	NA	NA	NA
666505	Phycisphaerae		Salt/brackish aquatics	Bacteria	NA	NA	NA	NA
2483368	Sedimentisphaera		Salt/brackish aquatics	Bacteria	NA	NA	NA	NA
473814	unclassified Planctomycetes		Mixed aquatics	Bacteria	NA	NA	NA	NA
1783272	Terrabacteria group		Mixed	Mixed	1	1	NA	NA
1162	Nostocaceae		Mixed aquatics	Bacteria	NA	NA	NA	NA
2132	Spiroplasma		Mixed	Bacteria	1	1	NA	NA
2250122	Candidatus Bipolaricaulis		Mixed	Bacteria	NA	NA	NA	NA
95818	Candidatus Saccharibacteria		Mixed	Bacteria	1	1	NA	NA
2759	Eukaryota		Mixed	Mixed	1	1	NA	NA
61964	environmental samples		Mixed	Eukaryotes	NA	NA	NA	NA
33090	Viridiplantae		Mixed	Mixed	1	1	5.600	1.659
3041	Chlorophyta		Algae	Algae	1	1	0.762	1.036
35429	Chlorodendraceae		Freshwater aquatics	Algae	1	1	0.475	0.372
3164	Tetraselmis		Mixed aquatics	Algae	1	1	0.475	0.372
3166	Chlorophyceae		Freshwater aquatics	Algae	1	1	0.762	1.036
3042	Chlamydomonadales		Mixed	Algae	1	1	0.762	1.036
3051	Chlamydomonadaceae		Algae	Algae	1	1	0.762	1.036
3052	Chlamydomonas		Mixed	Algae	NA	1	0.762	1.036
2034146	unclassified Chlamydomonas		Algae	Algae	NA	NA	0.762	1.036
51727	Chloromonas		Freshwater aquatics	Algae	1	1	0.762	1.036
77634	Chlamydomonadales incertae sedis		Algae	Algae	NA	NA	0.762	1.036
44649	Chlorococcum		Mixed	Algae	NA	1	0.762	1.036
3044	Dunaliella		Mixed aquatics	Algae	1	1	0.762	1.036
51321	Hafniomonas		Freshwater aquatics	Algae	1	1	0.762	1.036
2126383	unclassified Balticola		Algae	Algae	NA	NA	0.762	1.036
52030	Phacotaceae		Algae	Algae	NA	NA	0.762	1.036
47777	Tetrabaenaceae		Algae	Algae	1	1	0.762	1.036
3065	Volvocaceae		Freshwater aquatics	Algae	1	1	0.762	1.036
128464	Lobomonas		Freshwater aquatics	Algae	1	1	0.762	1.036
47284	Pleodorina		Algae	Algae	NA	NA	0.762	1.036
3066	Volvox		Freshwater aquatics	Algae	1	1	0.762	1.036
47792	Volvulina		Algae	Algae	NA	NA	0.762	1.036
34144	Chlorophyceae incertae sedis		Algae	Algae	1	1	0.762	1.036
3191	Spermatozopsis		Algae	Algae	NA	NA	0.762	1.036

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
1521739	Pleurastrosarcina		Algae	Algae	0	0	0.762	1.036
2546211	OCC clade		Algae	Algae	NA	1	0.762	1.036
51326	Chaetopeltidaceae		Mixed aquatics	Algae	1	1	0.762	1.036
31299	Chaetophorales		Algae	Algae	NA	1	0.762	1.036
2066085	unclassified Diplosphaera		Terrestrial	Algae	NA	1	0.762	1.036
35490	Oedogoniales		Freshwater aquatics	Algae	1	1	0.762	1.036
35491	Sphaeropleales		Freshwater aquatics	Algae	1	1	0.762	1.036
50036	Bracteacoccus		Freshwater aquatics	Algae	NA	1	0.762	1.036
3103	Hydrodictyaceae		Freshwater aquatics	Algae	1	1	0.762	1.036
3104	Pediastrum		Algae	Algae	NA	NA	0.762	1.036
427904	Pseudopediastrum		Algae	Algae	NA	NA	0.762	1.036
91354	Mychonastes		Mixed	Algae	1	1	0.762	1.036
42695	Pseudomuriella		Algae	Algae	NA	NA	0.762	1.036
1656252	Rotundella		Algae	Algae	0	0	0.762	1.036
3086	Scenedesmaceae		Mixed aquatics	Algae	1	1	0.762	1.036
91202	Desmodesmus		Freshwater aquatics	Algae	1	1	0.762	1.036
305820	Hariotina		Freshwater aquatics	Algae	NA	1	0.762	1.036
35466	Selenastraceae		Algae	Algae	1	1	0.762	1.036
2302913	Chloropicaceae		Algae	Algae	NA	NA	0.762	1.036
2302916	Chloroparvula		Salt/brackish aquatics	Algae	NA	NA	0.762	1.036
2565278	unclassified Chloroparvula		Algae	Algae	NA	NA	0.762	1.036
2302914	Chloropicon		Algae	Algae	NA	NA	0.762	1.036
1035538	Mamiellophyceae		Salt/brackish aquatics	Algae	NA	1	0.020	0.707
13792	Mamiellales		Salt/brackish aquatics	Algae	NA	1	0.020	0.707
1525212	Bathycoccaceae		Salt/brackish aquatics	Algae	1	1	0.010	NA
70447	Ostreococcus		Salt/brackish aquatics	Algae	1	1	0.010	NA
2268852	unclassified Ostreococcus		Salt/brackish aquatics	Algae	NA	1	0.010	NA
41873	Mamiellaceae		Salt/brackish aquatics	Algae	1	1	0.030	NA
38832	Micromonas		Salt/brackish aquatics	Algae	1	1	0.030	NA
2201466	Palmophyllaceae		Algae	Algae	NA	NA	0.762	1.036
35422	Pedinophyceae		Algae	Algae	1	1	0.762	1.036
3158	Pedinomonas		Algae	Algae	1	1	0.762	1.036
2306716	unclassified Pedinomonas		Algae	Algae	NA	NA	0.762	1.036
41878	Pycnococcaceae		Algae	Algae	1	1	0.762	1.036
38834	Pyramimonadales		Salt/brackish aquatics	Algae	NA	1	0.080	NA
36882	Pyramimonas		Freshwater aquatics	Algae	1	1	0.080	NA
75966	Trebouxiophyceae		Mixed	Algae	NA	1	0.133	1.552
35460	Chlorellales		Algae	Algae	NA	1	0.027	0.433
35461	Chlorellaceae		Algae	Algae	1	1	0.033	0.382
191392	Auxenochlorella		Algae	Algae	NA	1	0.033	0.382
2511126	Chlorella clade		Algae	Algae	NA	1	0.040	NA
3071	Chlorella		Mixed	Algae	NA	1	0.043	0.118
247496	Dictyosphaerium		Algae	Algae	NA	NA	0.040	NA
126838	Micractinium		Algae	Algae	NA	NA	0.040	NA
2064757	unclassified Nannochloris		Freshwater aquatics	Algae	NA	1	0.020	NA
3110	Prototheca		Algae	Algae	0	0	0.033	0.382
63681	Pseudochlorella		Algae	Algae	NA	NA	0.033	0.382
3070	Oocystaceae		Freshwater aquatics	Algae	1	1	0.027	0.433
163319	Planctonema		Freshwater aquatics	Algae	1	1	0.027	0.433
2070461	Oocystoideae		Algae	Algae	NA	NA	0.027	0.433
202678	Lagerheimia		Freshwater aquatics	Algae	1	1	0.027	0.433
2511161	Elliptochloris clade		Algae	Algae	NA	1	0.133	1.552
38881	Botryococcus braunii		Mixed aquatics	Algae	1	1	0.170	NA

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
41891	Coccomyxa		Algae	Algae	NA	1	0.133	1.552
2064756	unclassified Stichococcus		Mixed aquatics	Algae	NA	1	0.080	NA
2030936	Koliellaceae		Mixed	Algae	NA	NA	0.230	NA
135266	Prasiolaceae		Mixed	Algae	NA	1	0.230	NA
2507902	Trebouxiaceae		Algae	Algae	1	1	0.133	1.552
13786	Trebouxia		Mixed	Algae	NA	1	0.133	1.552
75981	Trebouxiophyceae incertae sedis		Algae	Algae	NA	NA	0.030	0.943
41299	Choricystis		Mixed	Algae	1	1	0.030	0.943
2306081	unclassified Choricystis		Algae	Algae	NA	1	0.030	0.943
1403567	Lemmermannia		Freshwater aquatics	Algae	NA	1	0.030	0.943
249344	Picochlorum		Salt/brackish aquatics	Algae	NA	1	0.027	0.781
550772	Heveochlorella		Terrestrial	Algae	NA	NA	0.610	NA
33103	Ulvophyceae		Mixed aquatics	Algae	1	1	0.875	0.924
2546215	OUU clade		Algae	Algae	1	1	0.274	0.418
31306	Ulotrichales		Algae	Algae	NA	1	0.228	0.660
205393	Ulotrichales incertae sedis		Algae	Algae	NA	1	0.145	0.244
3113	Ulvales		Algae	Algae	1	1	0.294	0.334
3118	Ulva		Salt/brackish aquatics	Algae	1	1	0.260	0.418
2546214	TCBD clade		Algae	Algae	NA	NA	1.014	0.825
33104	Bryopsidales		Salt/brackish aquatics	Algae	1	1	0.702	1.332
3128	Bryopsis	Hair algae	Salt/brackish aquatics	Algae	1	1	0.437	0.251
76312	Caulerpa		Salt/brackish aquatics	Algae	NA	1	0.105	0.165
3132	Codium		Salt/brackish aquatics	Algae	1	1	1.520	0.838
121087	Ostreobium		Salt/brackish aquatics	Algae	0	1	0.230	NA
2086555	unclassified Ostreobium		Salt/brackish aquatics	Algae	NA	NA	0.230	NA
35435	Udoteaceae		Algae	Algae	NA	NA	0.702	1.332
35436	Cladophoraceae		Algae	Algae	0	0	0.591	0.918
285966	Polyphysaceae		Salt/brackish aquatics	Algae	0	0	0.989	0.438
35443	Trentepohliales		Algae	Algae	1	1	1.045	0.543
35445	Trentepohliaceae		Algae	Algae	1	1	1.165	0.514
173374	Trentepohlia		Algae	Algae	1	1	0.864	0.459
2137841	unclassified Trentepohlia		Algae	Algae	NA	NA	0.864	0.459
948884	unclassified Trentepohliales		Algae	Algae	NA	NA	1.045	0.543
35493	Streptophyta		Mixed	Mixed	1	1	5.600	1.663
3173	Klebsormidiaceae		Algae	Algae	1	1	0.210	0.471
519230	Interfilum		Algae	Algae	NA	NA	0.210	0.471
3174	Klebsormidium		Mixed	Algae	1	1	0.210	0.471
131221	Streptophytina		Mixed	Mixed	1	1	5.602	1.663
3146	Characeae	Stonewort family	Mixed aquatics	Algae	1	1	11.487	0.625
13778	Chara	Stoneworts	Mixed aquatics	Algae	1	1	11.487	0.625
3148	Nitella	Stoneworts	Freshwater aquatics	Algae	NA	1	11.487	0.625
37933	Tolypella		Algae	Algae	NA	NA	11.487	0.625
3193	Embryophyta	Land plants	Mixed	Mixed	1	1	5.599	1.666
13809	Anthocerotophyta	Hornworts	Bryophytes	Hornworts	1	1	0.249	0.468
402688	Anthocerotopsida		Bryophytes	Hornworts	1	1	0.253	0.465
3208	Bryophyta	Mosses	Bryophytes	Mosses	1	1	0.515	0.499
404260	Bryophytina		Bryophytes	Mosses	1	1	0.512	0.529
3214	Bryopsida		Bryophytes	Mosses	1	1	0.504	0.498
114658	Bryidae		Bryophytes	Mosses	1	1	0.519	0.516
404297	Bryanae		Bryophytes	Mosses	1	1	0.777	0.533
37411	Bryaceae		Bryophytes	Mosses	1	1	0.872	0.162
404315	Hypnanae		Bryophytes	Mosses	1	1	0.457	0.373
13798	Hypnales	Feather mosses	Bryophytes	Mosses	1	1	0.446	0.286

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
52997	Amblystegiaceae		Bryophytes	Mosses	NA	NA	0.388	0.317
37421	Brachytheciaceae		Bryophytes	Mosses	1	1	0.493	0.401
67432	Fontinalaceae	Water moss family	Freshwater aquatics	Mosses	1	1	0.595	0.012
109248	Herzogiella		Bryophytes	Mosses	1	1	0.410	NA
114657	Dicranidae		Bryophytes	Mosses	1	1	0.441	0.366
65550	Grimmiaceae		Bryophytes	Mosses	1	1	0.350	NA
38586	Pottiaceae		Bryophytes	Mosses	1	1	0.360	0.118
114656	Funariidae		Bryophytes	Mosses	1	1	0.460	0.143
3216	Funariaceae		Bryophytes	Mosses	1	1	0.465	0.198
3211	Polytrichaceae	Hair moss family	Bryophytes	Mosses	1	1	0.739	0.753
3212	Polytrichum	Hair moss	Bryophytes	Mosses	1	1	0.480	0.103
13802	Sphagnales	Peat moss order	Bryophytes	Mosses	1	1	0.534	0.337
13804	Sphagnum	Peat moss	Bryophytes	Mosses	1	1	0.534	0.337
3195	Marchantiophyta	Liverworts	Bryophytes	Liverworts	1	1	1.890	1.551
186771	Jungermanniopsida		Bryophytes	Liverworts	1	1	1.854	1.646
186782	Jungermanniidae		Bryophytes	Liverworts	1	1	1.064	1.337
3199	Jungermanniales		Bryophytes	Liverworts	1	1	1.192	1.366
71154	Cephaloziineae		Bryophytes	Liverworts	1	1	0.617	1.036
186798	Porellales		Bryophytes	Liverworts	1	1	0.631	0.495
186799	Jubulineae		Bryophytes	Liverworts	1	1	0.560	0.372
65055	Lejeuneaceae		Bryophytes	Liverworts	1	1	0.210	NA
139836	Radula		Bryophytes	Liverworts	1	1	0.462	0.382
186770	Marchantiopsida		Bryophytes	Liverworts	1	1	0.655	0.450
122623	Blasiaceae		Bryophytes	Liverworts	1	1	0.500	NA
186774	Marchantiidae		Bryophytes	Liverworts	1	1	0.665	0.454
28908	Marchantiales		Bryophytes	Liverworts	1	1	0.665	0.470
29585	Marchantiaceae		Bryophytes	Liverworts	1	1	0.597	0.476
58023	Tracheophyta		Mixed	Mixed	1	1	5.780	1.635
78536	Euphyllophyta		Mixed	Mixed	1	1	5.810	1.630
241806	Polypodiopsida	Ferns	Ferns	Ferns	1	1	14.806	0.964
3256	Equisetaceae	Horsetails	Ferns	Ferns	1	1	21.729	0.295
3257	Equisetum	Horsetails	Ferns	Ferns	1	1	21.729	0.295
1521257	Ophioglossidae		Ferns	Ferns	1	1	35.878	0.935
1521262	Polypodiidae	Leptosporangiate ferns	Ferns	Ferns	1	1	12.175	0.657
3268	Polypodiales	Polypod ferns	Ferns	Ferns	1	1	12.435	0.665
1203511	Aspleniineae	Spleenwort suborder	Ferns	Ferns	1	1	11.105	0.520
32071	Asplenium	Spleenworts	Ferns	Ferns	1	1	11.095	0.547
1203520	Athyriaceae		Ferns	Ferns	1	1	13.586	0.459
32109	Athyrium		Ferns	Ferns	1	1	7.200	NA
2184407	Athyrium incertae sedis		Ferns	Ferns	1	1	7.200	NA
2175194	Athyrium sect. Athyrium		Ferns	Ferns	0	0	7.200	NA
65717	Deparia	False spleenworts	Ferns	Ferns	0	0	20.205	0.066
29614	Diplazium	Twinsorus ferns	Ferns	Ferns	0	0	12.955	0.492
29600	Blechnaceae	Chain fern family	Ferns	Ferns	1	1	12.266	0.338
1203500	Cystopteridaceae	Bladder fern family	Ferns	Ferns	1	1	7.136	0.215
32111	Cystopteris	Bladder ferns	Ferns	Ferns	1	1	7.358	0.275
1203514	Diplaziopsidaceae		Ferns	Ferns	0	0	11.105	0.520
29616	Thelypteridaceae		Ferns	Ferns	1	1	10.226	0.476
2014961	Phegopteridoideae		Ferns	Ferns	1	1	7.200	NA
2014962	Thelypteridoideae		Ferns	Ferns	1	1	11.976	0.385
29617	Thelypteris	Maiden ferns	Ferns	Ferns	1	1	7.590	NA
29618	Thelypteris palustris	Marsh fern	Ferns	Ferns	1	1	7.590	NA
32084	Dennstaedtiaceae	Bracken family	Ferns	Ferns	1	1	9.716	0.357

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
32100	Pteridium	Brackens	Ferns	Ferns	1	1	8.370	0.122
32101	Pteridium aquilinum	Bracken	Ferns	Ferns	1	1	7.410	NA
1203512	Polypodiineae		Ferns	Ferns	1	1	16.092	0.646
29607	Dryopteridaceae	Wood ferns	Ferns	Ferns	1	1	17.950	0.708
2014984	Dryopteridoideae		Ferns	Ferns	1	1	12.669	0.460
3287	Dryopteris	Buckler and male ferns	Ferns	Ferns	1	1	13.083	0.456
2014983	Elaphoglossoideae		Ferns	Ferns	0	0	26.409	0.600
3275	Polypodiaceae		Ferns	Ferns	1	1	12.993	0.290
1580210	Grammitidoideae		Ferns	Ferns	0	0	12.993	0.290
2017694	Microsoroideae		Ferns	Ferns	0	0	11.440	NA
2017695	Polypodioideae		Ferns	Ferns	1	1	12.739	0.301
38352	Polypodium	Polypodies	Ferns	Ferns	1	1	12.972	0.319
13819	Pteridaceae		Ferns	Ferns	1	1	9.056	0.827
2003546	Cheilanthoideae	Heart ferns	Ferns	Ferns	0	1	5.100	0.353
29597	Coniogramme	Bamboo ferns	Ferns	Ferns	0	0	9.700	0.296
2003544	Pteridoideae		Ferns	Ferns	0	1	8.076	0.345
13817	Adiantum	Maidenhair ferns	Ferns	Ferns	1	1	6.033	0.200
29627	Antrophyum		Ferns	Ferns	0	0	11.610	1.078
13814	Marsileaceae	Water fern family	Freshwater aquatics	Ferns	1	1	4.080	NA
32186	Salviniaceae	Water fern family	Freshwater aquatics	Ferns	0	1	1.525	0.700
13823	Lygodium	Climbing ferns	Ferns	Ferns	0	0	10.870	0.404
58024	Spermatophyta	Seed plants	Mixed	Mixed	1	1	5.589	1.648
1437180	Acrogymnospermae	Gymnosperms	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	18.490	0.398
3297	Cycadales	Cycad order	Trees/shrubs	Trees/shrubs	0	0	21.290	0.229
3395	Cycas		Trees/shrubs	Trees	0	0	13.413	0.051
3298	Zamiaceae		Trees/shrubs	Trees/shrubs	0	0	21.877	0.193
3303	Zamia		Trees/shrubs	Shrubs	0	0	19.836	0.087
1445966	Gnetidae	Gnetophytes	Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	9.812	0.567
3387	Ephedra	Joint-pine	Xerophytes	Shrubs	0	1	14.918	0.241
3380	Gnetum		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	3.367	0.288
3313	Pinidae	Conifers	Trees/shrubs	Trees/shrubs	1	1	17.942	0.430
1446378	Araucariales	Araucarians	Trees/shrubs	Trees/shrubs	0	0	11.780	0.430
25664	Araucariaceae	Araucarian family	Trees/shrubs	Trees/shrubs	0	0	16.392	0.186
25666	Araucaria	Monkey-puzzle genus	Trees/shrubs	Trees/shrubs	0	0	16.943	0.172
3362	Podocarpaceae	Podocarp family	Trees/shrubs	Trees/shrubs	0	0	8.420	0.393
1446379	Cupressales		Trees/shrubs	Trees	1	1	13.384	0.365
3367	Cupressaceae	Cypress family	Trees/shrubs	Trees	1	1	12.785	0.333
25623	Taxaceae	Yew family	Trees/shrubs	Trees	1	1	16.232	0.421
50178	Cephalotaxus	Plum yew	Trees/shrubs	Trees/shrubs	0	0	25.350	NA
25628	Taxus	Yews	Trees/shrubs	Trees/shrubs	1	1	11.056	0.125
3318	Pinaceae	Pine family	Trees/shrubs	Trees	1	1	23.856	0.242
3321	Cedrus	Cedars	Trees/shrubs	Trees	0	0	16.130	0.011
3325	Larix	Larch	Trees/shrubs	Trees	0	1	12.976	0.097
3328	Picea	Spruces	Trees/shrubs	Trees	0	1	18.851	0.131
3337	Pinus	Pines	Trees/shrubs	Trees	1	1	26.720	0.159
139271	Pinus	Hard pines	Trees/shrubs	Trees	1	1	26.720	0.159
139272	Strobus	Soft pines	Trees/shrubs	Trees	0	1	30.457	0.093
3356	Pseudotsuga		Trees/shrubs	Trees	0	0	19.050	NA
3398	Magnoliopsida	Flowering plants	Mixed	Mixed	1	1	4.950	1.782
82956	Austrobaileyales		Trees/shrubs	Trees/shrubs	0	0	9.270	0.408
16733	Schisandraceae		Trees/shrubs	Trees/shrubs	0	0	9.193	0.412
13097	Illicium	Star anise	Trees/shrubs	Trees/shrubs	0	0	10.883	0.595
1437183	Mesangiospermae	Core angiosperms	Mixed	Mixed	1	1	4.958	1.782

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
4000	Buxaceae	Box family	Trees/shrubs	Trees	1	1	1.061	0.312
4427	Ceratophyllum	Coontails	Freshwater aquatics	Herbs	1	1	0.690	NA
16737	Chloranthaceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	3.600	0.194
13669	Sarcandra		Herbs	Herbs	0	0	4.300	NA
71240	eudicotyledons	Eudicots	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.124	1.674
91827	Gunneridae		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.124	1.674
1437201	Pentapetalae		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.123	1.675
71274	asterids		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.429	1.053
91882	campanulids		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	3.248	0.826
4036	Apiales	Carrot order	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.143	0.629
364270	Apiineae		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.143	0.629
48035	Cicuta	Water-hemlock	Freshwater aquatics	Herbs	1	1	1.780	0.207
54827	Berula erecta	Lesser water-parsnip	Freshwater aquatics	Herbs	1	1	0.440	NA
241777	Azorelloideae		Herbs	Herbs	0	0	2.194	0.609
52507	Azorella		Herbs	Herbs	0	0	2.194	0.609
241778	Apioideae		Herbs	Herbs	1	1	2.279	0.599
241789	Scandiceae	Carrot tribe	Herbs	Herbs	1	1	2.022	0.730
241780	apioid superclade		Herbs	Herbs	1	1	2.852	0.411
4038	Daucus	Carrots	Herbs	Herbs	1	1	1.587	0.937
241792	Selineae	Angelica tribe	Herbs	Herbs	1	1	2.865	0.358
1589887	Arracacia clade		Herbs	Herbs	0	0	2.865	0.358
241787	Oenantheae	Water-parsnip tribe	Herbs	Herbs	1	1	1.263	0.763
241800	Scandicinae		Herbs	Herbs	1	1	1.496	0.469
241793	Apieae	Celeries	Herbs	Herbs	1	1	1.947	0.334
4039	Daucus carota	Wild carrot	Herbs	Herbs	1	1	2.585	1.157
4037	Apiaceae	Umbellifer family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.194	0.609
1431669	Arthrophyllum		Trees/shrubs	Trees/shrubs	0	0	1.981	0.732
4051	Hedera	Ivies	Herbs	Herbs	1	1	2.658	0.578
4050	Araliaceae	Ivy family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.981	0.732
91883	Aquifoliales	Holly order	Trees/shrubs	Trees/shrubs	1	1	1.372	0.425
4295	Ilex	Hollies	Trees/shrubs	Trees/shrubs	1	1	1.372	0.425
4209	Asterales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	3.508	0.777
49929	Alseuosmiaceae		Trees/shrubs	Shrubs	0	0	3.508	0.777
4210	Asteraceae	Daisy family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	3.683	0.830
4219	Artemisia	Mugworts	Xerophytes	Trees/shrubs and herbs	1	1	5.382	0.480
72924	Diplostephium		Trees/shrubs	Trees/shrubs	0	0	1.930	0.662
18794	Senecio	Ragworts and groundsels	Herbs	Herbs	1	1	5.667	0.814
795077	Senecioninae	Ragwort subtribe	Herbs	Herbs	1	1	4.447	0.741
13328	Achillea	Yarrows	Herbs	Herbs	1	1	3.904	0.496
886729	Matricariinae		Herbs	Herbs	1	1	3.796	0.434
4231	Helianthus		Herbs	Herbs	0	0	5.570	0.444
13422	Chrysanthemum	Chrysanthemums	Herbs	Herbs	0	1	7.580	0.444
886697	Mediterranean clade		Herbs	Herbs	1	1	9.822	0.535
886730	Anthemidinae	Tansy subtribe	Herbs	Herbs	1	1	4.942	0.430
911294	Inulinae		Herbs	Herbs	1	1	1.589	0.622
41479	Aster		Herbs	Herbs	1	1	3.284	0.796
41589	Inula		Herbs	Herbs	1	1	1.698	0.477
795080	Tussilagininae	Coltsfoot subtribe	Herbs	Herbs	1	1	3.102	0.737
41574	Erigeron	Fleabanes	Herbs	Herbs	1	1	1.512	0.301
4211	Ambrosia	Ragweeds	Herbs	Herbs	0	1	2.286	0.456
99105	Tanacetum	Tansies	Herbs	Herbs	1	1	7.273	0.383
13516	Eupatorium	Hemp-agrimony	Herbs	Herbs	1	1	2.825	0.271
102813	Helenieae	Sneezeweed tribe	Herbs	Herbs	0	0	7.320	NA

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
4246	Arnica	Mountain tobacco	Herbs	Herbs	1	1	1.570	NA
4239	Ligularia	Leopard plants	Herbs	Herbs	0	1	3.102	0.737
102804	Asteroideae		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	3.991	0.778
102810	Anthemideae	Chamomile tribe	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	5.471	0.577
911341	Heliantheae alliance		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	3.678	0.751
886714	Artemisiinae		Trees/shrubs and herbs	Shrubs/herbs	1	1	5.317	0.447
102812	Senecioneae	Ragwort tribe	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	4.338	0.745
102814	Heliantheae	Sunflower tribe	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	5.334	0.598
102808	Gnaphalieae		Trees/shrubs and herbs	Shrubs/herbs	1	1	1.034	0.604
102809	Astereae	Aster tribe	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.930	0.662
886584	South American lineages	Aster tribe	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.930	0.662
877976	North American clade		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.845	0.623
102815	Eupatorieae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	2.197	0.386
886583	Australasian lineages		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	2.600	0.462
102818	Cardueae	Thistle tribe	Herbs	Herbs	1	1	2.102	0.769
219103	Carduoideae	Thistle subfamily	Herbs	Herbs	1	1	2.102	0.769
41503	Centaurea	Starthistles	Herbs	Herbs	1	1	1.422	0.335
742011	Centaureinae	Knapweed subtribe	Herbs	Herbs	1	1	1.244	0.468
742010	Carduinae	Thistle subtribe	Herbs	Herbs	1	1	1.810	0.854
41549	Cirsium	Plume thistles	Herbs	Herbs	1	1	1.589	0.664
742007	Carlininae		Herbs	Herbs	1	1	3.860	0.214
219120	Cichorioideae	Chicory subfamily	Herbs	Herbs	1	1	3.518	0.616
219121	Cichorieae	Chicory tribe	Herbs	Herbs	1	1	3.564	0.599
102745	Hieracium	Hawkweed	Herbs	Herbs	1	1	4.820	0.297
745087	Hieraciinae	Hawkweed subtribe	Herbs	Herbs	1	1	4.763	0.333
745067	Crepidinae	Hawk's-beard subtribe	Herbs	Herbs	1	1	3.535	0.683
745069	Hypochaeridinae	Cat's ear subtribe	Herbs	Herbs	1	1	1.849	0.591
745063	Hyoseridinae	Sow-thistle subtribe	Herbs	Herbs	1	1	1.563	0.260
49743	Taraxacum	Dandelions	Herbs	Herbs	1	1	1.523	0.446
50190	Sonchus	Sow-thistles	Herbs	Herbs	1	1	1.497	0.338
4381	Campanulaceae	Bellflower family	Herbs	Herbs	1	1	2.117	0.802
239444	Githopsis	Bluecups	Herbs	Herbs	0	0	2.117	0.802
378889	Monopsis		Herbs	Herbs	0	0	2.117	0.802
40568	Campanula	Bellflowers	Herbs	Herbs	1	1	1.809	0.463
4382	Lobelia	Lobelias	Herbs	Herbs	1	1	3.685	1.042
16472	Goodeniaceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.520	NA
24579	Menyanthaceae		Freshwater aquatics	Herbs	1	1	0.700	NA
4199	Dipsacales		Trees/shrubs and herbs	Shrubs/herbs	1	1	2.291	1.253
4206	Adoxaceae	Elder family	Trees/shrubs	Shrubs	1	1	7.564	0.628
4204	Viburnum		Trees/shrubs	Shrubs	1	1	3.806	0.098
49606	Lonicera	Honeysuckles	Trees/shrubs	Shrubs	1	1	1.171	0.643
4200	Caprifoliaceae	Honeysuckle family	Trees/shrubs and herbs	Shrubs/herbs	1	1	1.763	1.010
41934	Cornales		Trees/shrubs	Shrubs	1	1	1.631	0.437
42219	Cornaceae	Dogwood family	Trees/shrubs	Shrubs	1	1	1.683	0.515
4281	Cornus	Dogwoods	Trees/shrubs	Shrubs	1	1	1.683	0.515
23097	Hydrangeaceae	Hydrangea family	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.707	0.388
23109	Hydrangea	Hydrangea	Trees/shrubs	Trees/shrubs	0	0	1.776	0.348
41945	Ericales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.815	1.130
3623	Actinidiaceae	Kiwifruit family	Trees/shrubs	Trees/shrubs	0	0	1.264	0.525
3624	Actinidia	Kiwifruits	Trees/shrubs	Shrubs	0	0	1.264	0.525
35939	Impatiens	Balsams	Herbs	Herbs	1	1	1.581	0.600
1758105	Impatiens sect. Impatiens		Herbs	Herbs	1	1	0.780	NA
13492	Diospyros	Ebony and persimmon trees	Trees/shrubs	Trees/shrubs	0	0	1.476	0.495

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
4345	Ericaceae	Heather family	Trees/shrubs	Trees/shrubs	1	1	2.997	2.003
217033	Arbutoideae	Bearberry subfamily	Trees/shrubs	Trees/shrubs	1	1	1.250	NA
217035	Ericoideae	Heather subfamily	Trees/shrubs	Trees/shrubs	1	1	0.602	0.145
45918	Empetrum	Crowberries	Trees/shrubs	Shrubs	1	1	0.965	0.462
217043	Empetreae	Crowberry tribe	Trees/shrubs	Shrubs	1	1	0.650	NA
217045	Phyllodoceae	Mountain heather family	Trees/shrubs	Trees/shrubs	1	1	0.670	NA
4346	Rhododendron	Rhododendron	Trees/shrubs	Shrubs	0	1	0.740	NA
217032	Pyroloideae	Wintergreen subfamily	Herbs	Herbs	1	1	11.420	0.921
217039	Monotropeae	Monotropa tribe	Herbs	Herbs	1	1	29.900	NA
217038	Pyroleae	Wintergreen tribe	Herbs	Herbs	1	1	6.800	0.340
13650	Pyrola	Wintergreen	Herbs	Herbs	1	1	4.810	0.003
217037	Vaccinioideae	Cranberry subfamily	Trees/shrubs	Trees/shrubs	1	1	0.721	0.321
217059	Gaultherieae	Salal tribe	Trees/shrubs	Trees/shrubs	0	0	1.250	NA
217062	Vaccinieae	Cranberry tribe	Trees/shrubs	Trees/shrubs	1	1	0.662	0.221
13749	Vaccinium	Cranberry genus	Trees/shrubs	Shrubs	1	1	0.674	0.208
3642	Lecythidaceae		Trees/shrubs	Trees/shrubs	0	0	1.140	NA
79564	Corythophora		Trees/shrubs	Trees/shrubs	0	0	1.140	NA
4335	Primulaceae	Primrose family	Herbs	Herbs	1	1	2.867	0.813
85163	Androsace	Rock jasmine	Herbs	Herbs	0	1	2.867	0.813
59977	Lysimachia	Loosestrifes	Herbs	Herbs	1	1	2.773	0.688
49647	Primula	Primroses	Herbs	Herbs	1	1	1.247	1.044
1609961	Chrysophylloideae		Trees/shrubs	Trees/shrubs	0	0	0.540	NA
85165	Planchonella		Trees/shrubs	Trees	0	0	0.540	NA
85249	Pouteria	Eggfruit trees	Trees/shrubs	Trees	NA	NA	0.540	NA
20008	Styracaceae		Trees/shrubs	Trees/shrubs	0	0	0.775	0.210
167982	Bruinsmia		Trees/shrubs	Trees	0	0	0.775	0.210
27065	Theaceae	Camellia family	Trees/shrubs	Trees/shrubs	0	0	3.634	0.558
4441	Camellia		Trees/shrubs	Trees/shrubs	0	0	3.634	0.558
91888	lamiids		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.460	1.323
1538097	Boraginales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.536	1.070
21571	Boraginaceae	Borage family	Herbs	Herbs	1	1	1.477	1.214
1874405	Lithospermeae	Gromwell tribe	Herbs	Herbs	1	1	0.994	0.749
1874400	Cynoglossoideae		Herbs	Herbs	1	1	0.904	0.428
1874406	Myosotideae	Forget-me-not tribe	Herbs	Herbs	1	1	0.930	0.715
1874418	Amsinckiinae	Fiddleneck subtribe	Herbs	Herbs	0	0	1.200	0.118
181188	Cynoglossum		Herbs	Herbs	1	1	0.910	NA
4055	Gentianales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.107	0.864
4056	Apocynaceae	Dogbane family	Herbs	Herbs	1	1	0.988	0.789
21199	Asclepias	Milkweeds	Herbs	Herbs	0	0	0.420	0.253
167488	Asclepiadeae	Milkweed tribe	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.398	0.262
1498477	Asclepiadinae	Milkweed subtribe	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.420	0.253
167487	Rauvolfioideae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.076	0.609
167498	Vinceae	Periwinkle tribe	Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.280	0.535
21472	Gentianaceae	Gentian family	Herbs	Herbs	1	1	3.285	0.688
303185	Gentianeae	Gentian tribe	Herbs	Herbs	1	1	3.513	0.428
2546018	Swertiinae	Dwarf gentian tribe	Herbs	Herbs	1	1	3.513	0.428
24966	Rubiaceae	Bedstraw family	Herbs	Herbs	1	1	0.990	0.713
169619	Cinchonoideae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.244	0.464
13442	Coffea	Coffees	Trees/shrubs	Trees/shrubs	0	0	0.623	0.190
169617	Rubioideae		Herbs	Herbs	1	1	1.392	0.570
25168	Galium	Bedstraw	Herbs	Herbs	1	1	1.374	0.356
169660	Rubieae	Madder tribe	Herbs	Herbs	1	1	1.418	0.326
169663	Spermacoceae		Trees/shrubs and herbs	Shrubs/herbs	0	0	3.003	0.514

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
4321	Icacinaceae		Trees/shrubs	Trees/shrubs	0	0	1.090	NA
4143	Lamiales	Dead-nettle order	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.139	1.050
216691	Acanthoideae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.410	0.679
24079	Bignoniaceae	Bigonia family	Trees/shrubs	Trees/shrubs	0	0	1.009	0.456
423302	Bignonieae	Bigonia tribe	Trees/shrubs	Trees/shrubs	0	0	1.543	0.284
26122	Gesneriaceae	African violet family	Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.015	0.585
214607	Trichosporeae		Herbs	Herbs	0	1	0.953	0.177
1477611	Didymocarpinae		Herbs	Herbs	0	0	0.950	0.167
214598	Gesnerioideae	African violet subfamily	Herbs	Herbs	0	0	0.335	0.190
4136	Lamiaceae	Dead-nettle family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.274	0.881
216703	Lamioideae	Dead-nettle subfamily	Herbs	Herbs	1	1	1.576	0.685
21861	Pogostemon		Herbs	Herbs	0	0	1.576	0.685
983535	Stachydeae		Trees/shrubs and herbs	Shrubs/herbs	1	1	1.387	0.361
260602	Lycopus	Gipsyworts	Freshwater aquatics	Herbs	1	1	0.625	0.396
216706	Nepetoideae	Catmint subfamily	Herbs	Herbs	1	1	0.885	0.698
216718	Mentheae	Mint tribe	Herbs	Herbs	1	1	0.815	0.622
21880	Salvia	Sages	Herbs	Herbs	1	1	0.761	0.525
2291027	Salvia incertae sedis		Herbs	Herbs	1	1	0.751	0.628
49990	Thymus	Thymes	Herbs	Herbs	1	1	0.874	0.364
39173	Ocimum	Basils	Herbs	Herbs	0	0	2.500	0.461
21819	Mentha	Mints	Herbs	Herbs	1	1	0.656	0.744
39174	Origanum	Oreganos	Herbs	Herbs	1	1	0.747	0.086
4139	Scutellaria	Skullcaps	Herbs	Herbs	1	1	0.582	0.535
4196	Lentibulariaceae	Bladderwort family	Freshwater aquatics	Herbs	1	1	0.397	0.798
13747	Utricularia	Bladderworts	Freshwater aquatics	Herbs	1	1	0.233	0.475
4144	Oleaceae	Olive family	Trees/shrubs	Trees/shrubs	1	1	1.498	0.416
4147	Jasminum	Jasmines	Trees/shrubs	Shrubs	0	0	1.440	NA
426106	Oleeae	Olive tribe	Trees/shrubs	Trees/shrubs	1	1	1.243	0.487
126412	Forestiera	Swampprivets	Trees/shrubs	Trees/shrubs	0	0	1.243	0.487
13596	Ligustrum	Privet	Trees/shrubs	Shrubs	1	1	1.480	0.048
91896	Orobanchaceae	Broomrape family	Herbs	Herbs	1	1	3.109	0.650
216770	Orobancheae	Broomrape tribe	Herbs	Herbs	1	1	3.539	0.418
2249430	Aphyllon		Herbs	Herbs	0	0	3.539	0.418
320805	Phelipanche	Broomrapes	Herbs	Herbs	1	1	4.357	0.115
216775	Orobanchaceae incertae sedis		Herbs	Herbs	0	0	0.460	NA
43174	Pedicularis	Louseworts	Herbs	Herbs	1	1	2.963	0.174
1325730	Pedicularideae	Lousewort tribe	Herbs	Herbs	1	1	3.100	0.274
216772	Rhinantheae	Yellow-rattle tribe	Herbs	Herbs	1	1	3.121	1.140
46053	Euphrasia	Eyebright	Herbs	Herbs	1	1	1.320	0.723
46059	Rhinanthus	Rattles	Herbs	Herbs	1	1	1.940	0.496
156152	Plantaginaceae	Plantain family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.086	1.006
13380	Callitriche	Water-starwort	Freshwater aquatics	Herbs	1	1	2.173	0.537
26867	Plantago	Plantains	Herbs	Herbs	1	1	2.180	1.098
4173	Veronica	Speedwells	Herbs	Herbs	1	1	0.867	0.534
216795	Veroniceae	Speedwell tribe	Herbs	Herbs	1	1	0.884	0.497
418793	Metteniusaceae		Trees/shrubs	Trees	0	0	1.460	1.323
4069	Solanales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.664	1.239
4118	Convolvulaceae	Bindweed family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.866	2.058
4128	Cuscuta	Dodders	Herbs	Herbs	1	1	7.173	1.383
1824618	Grammica		Herbs	Herbs	0	1	5.386	1.733
1824619	Monogynella	Dodders	Herbs	Herbs	0	1	21.700	0.050
4119	Ipomoea	Morning glories	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.172	0.348
267213	Ipomoeeae	Morning glory tribe	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.206	0.343

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
4070	Solanaceae	Nightshade family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.835	0.931
4085	Nicotiana	Tobacco plants	Trees/shrubs and herbs	Shrubs/herbs	0	0	3.628	0.320
24646	Lycium	Boxthorns	Trees/shrubs	Trees	0	1	2.937	0.613
4107	Solanum	Nightshades	Herbs	Herbs	1	1	1.402	0.520
4071	Capsicum	Peppers	Herbs	Herbs	0	0	4.115	0.173
49274	Lycopersicon	Tomatoes	Herbs	Herbs	0	0	1.000	0.119
424551	Solanoideae	Nightshade subfamily	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.116	0.640
3524	Caryophyllales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.911	0.869
3542	Aizoaceae	Ice plant family	Xerophytes	Trees/shrubs and herbs	0	0	0.400	NA
85189	Delosperma		Xerophytes	Herbs	0	0	0.400	NA
3563	Amaranthaceae	Amaranth family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.498	0.866
3564	Amaranthus	Amaranth	Herbs	Herbs	0	1	0.716	0.358
866800	Cactineae	Cactus suborder	Herbs	Herbs	1	1	2.746	0.646
3593	Cactaceae	Cactus family	Xerophytes	Herbs	0	0	2.709	0.614
186265	Cactoideae		Xerophytes	Herbs	0	0	2.766	0.698
186275	Echinocereeae		Xerophytes	Herbs	0	0	1.764	0.515
153889	Pfeiffera		Xerophytes	Herbs	0	0	1.764	0.515
186273	Hylocereeae		Xerophytes	Herbs	0	0	1.900	NA
3568	Caryophyllaceae	Pink family	Herbs	Herbs	1	1	1.464	0.666
1141488	Alsineae		Herbs	Herbs	1	1	1.730	0.503
146094	Gypsophila	Baby's-breath	Herbs	Herbs	0	1	0.600	0.236
1141492	Sileneae	Campion tribe	Herbs	Herbs	1	1	2.242	0.281
3573	Silene	Campions and catchflies	Herbs	Herbs	1	1	2.223	0.469
1804623	Chenopodiaceae	Goosefoot family	Xerophytes	Trees/shrubs and herbs	1	1	1.256	0.482
3554	Beta	Beets	Halophytes	Herbs	1	1	1.290	0.420
161934	Beta vulgaris	Beet	Halophytes	Herbs	1	1	1.110	0.242
3558	Chenopodium	Goosefoots	Halophytes	Trees/shrubs and herbs	1	1	1.249	0.426
3550	Atriplex	Orache	Halophytes	Herbs	1	1	1.458	0.466
1307796	Chenopodioideae	Goosefoot subfamily	Xerophytes	Trees/shrubs and herbs	1	1	1.222	0.498
1307774	Atripliceae	Atriplex tribe	Xerophytes	Shrubs/herbs	1	1	1.254	0.484
240058	Dysphania		Herbs	Herbs	0	1	0.370	NA
46104	Salicornia	Glassworts	Halophytes	Herbs	1	1	2.170	0.515
1316646	Salicornioideae	Glasswort subfamily	Halophytes	Herbs	1	1	2.170	0.515
2116531	Salicornia subg. Arthrocnemoides		Halophytes	Herbs	1	1	2.170	0.515
2116532	Salicornia subg. Salicornia		Halophytes	Herbs	1	1	1.380	NA
1804622	Suaedoideae	Sea-blite subfamily	Halophytes	Herbs	1	1	1.060	NA
46108	Suaeda	Sea-blites	Halophytes	Herbs	1	1	1.060	NA
4360	Droseraceae	Sundew family	Herbs	Herbs	1	1	0.648	0.372
98042	Limeum		Herbs	Herbs	0	0	1.911	0.869
3590	Molluginaceae	Carpetweed family	Herbs	Herbs	1	1	0.670	0.380
3536	Nyctaginaceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	3.236	0.712
4437	Plumbaginaceae	Leadwort family	Halophytes	Shrubs/herbs	1	1	4.840	0.430
63086	Armeria	Thrifts	Halophytes	Herbs	1	1	4.543	0.050
46093	Limonium	Sea-lavenders	Halophytes	Shrubs/herbs	1	1	4.692	0.514
3615	Polygonaceae	Knotweed family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.811	0.688
1110379	Eriogonoideae	American wild buckwheat subfamily	Herbs	Herbs	0	0	1.540	0.358
1110380	Polygonoideae	Knotweed subfamily	Herbs	Herbs	1	1	1.792	0.687
3616	Fagopyrum	Buckwheats	Herbs	Herbs	0	0	1.316	0.470
61508	Persicaria	Knotweeds	Herbs	Herbs	1	1	2.233	0.430
1110384	Persicarieae	Knotweed tribe	Herbs	Herbs	1	1	2.233	0.430
137670	Bistorta	Bistorts	Herbs	Herbs	1	1	2.233	0.430
1110385	Polygoneae	Knotweed tribe	Trees/shrubs and herbs	Shrubs/herbs	1	1	1.373	0.917
1110386	Rumiceae	Rhubarb family	Herbs	Herbs	1	1	2.415	0.621

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
3618	Rumex	Docks and sorrels	Herbs	Herbs	1	1	2.415	0.621
63083	Tamaricaceae	Saltcedar family	Xerophytes	Trees/shrubs	0	1	1.577	0.041
24942	Dilleniaceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.575	0.956
71275	rosids		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.429	1.325
91835	fabids		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.570	1.362
4305	Celastraceae	Staff vine family	Trees/shrubs	Shrubs	1	1	0.834	1.268
123471	Salaciopsis		Trees/shrubs	Trees/shrubs	0	0	0.834	1.268
123484	Tripterygium		Herbs	Herbs	0	0	0.834	1.268
71239	Cucurbitales	Gourd order	Herbs	Herbs	1	1	0.770	0.660
3681	Begonia		Herbs	Herbs	0	0	0.573	0.518
2203166	Begonia sect. Pritzelia		Herbs	Herbs	0	0	0.302	0.143
3650	Cucurbitaceae	Gourd family	Herbs	Herbs	1	1	0.975	0.655
3655	Cucumis	Cucumber and some melons	Herbs	Herbs	0	0	1.071	0.304
1003877	Benincaseae	Wax gourd tribe	Herbs	Herbs	0	0	0.924	0.286
3653	Citrullus	Watermelon genus	Herbs	Herbs	0	0	0.500	NA
3654	Citrullus lanatus	Watermelon	Herbs	Herbs	0	0	0.500	NA
3656	Cucumis melo	Muskmelons	Herbs	Herbs	0	0	1.000	NA
3660	Cucurbita	Gourds	Herbs	Herbs	0	0	0.407	0.149
182083	Gynostemma		Herbs	Herbs	0	0	1.810	NA
72025	Fabales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.107	1.264
3803	Fabaceae	Legume family	Trees/shrubs and herbs	Shrubs/herbs	1	1	2.079	1.214
3804	Caesalpinioideae	Peacock flower subfamily	Trees/shrubs	Trees	0	0	0.914	0.375
3807	mimosoid clade		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.910	0.373
163092	Cercideae	Redbud subfamily.	Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	0.642	0.078
2115957	Bauhiniinae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.642	0.078
1978182	Detarioideae	Tamarind subfamily	Trees/shrubs	Trees/shrubs	0	0	1.420	0.578
2231382	NPAAA clade		Herbs	Herbs	1	1	2.798	1.160
2233838	Hologalegina		Herbs	Herbs	1	1	3.593	1.015
2233839	IRL clade		Herbs	Herbs	1	1	4.171	0.898
163743	Fabeae		Herbs	Herbs	1	1	6.793	0.492
3904	Vicia	Vetches	Herbs	Herbs	1	1	5.642	0.651
3853	Lathyrus	Vetchlings	Herbs	Herbs	1	1	7.748	0.266
163742	Trifolieae	Clover tribe	Herbs	Herbs	1	1	1.060	0.916
3867	Lotus	Bird's-foot trefoils	Herbs	Herbs	1	1	0.695	0.497
3898	Trifolium	Clovers	Herbs	Herbs	1	1	1.120	0.996
163728	Galegeae		Herbs	Herbs	1	1	1.485	0.548
46347	Glycyrrhiza	Liquorice genus	Herbs	Herbs	0	1	1.485	0.548
3817	Arachis	Peanut genus	Herbs	Herbs	0	0	2.843	0.315
3913	Vigna		Herbs	Herbs	0	0	0.723	0.425
3877	Medicago	Medick	Herbs	Herbs	0	1	0.902	0.510
47034	Hedysarum	Sweetvetches	Herbs	Herbs	0	1	2.750	0.136
163726	Desmodieae		Herbs	Herbs	0	0	1.218	0.620
1462606	Soja		Herbs	Herbs	0	0	1.115	0.019
20802	Oxytropis		Herbs	Herbs	1	1	4.610	NA
3814	Papilionoideae		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.524	1.177
2231393	50 kb inversion clade		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.528	1.176
2233855	indigoferoid/millettioid clade		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.109	0.702
163735	Phaseoleae	Bean tribe	Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.055	0.446
2231385	core genistoids		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.064	0.558
2233857	robinioid clade		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	0.772	0.542
163747	Loteae		Trees/shrubs and herbs	Shrubs/herbs	1	1	0.756	0.563
20400	Astragalus	Milkvetches	Trees/shrubs and herbs	Shrubs/herbs	1	1	1.652	0.753
163725	Dalbergieae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	2.585	0.358

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
2231388	Adesmia clade		Trees/shrubs and herbs	Shrubs/herbs	0	0	2.585	0.358
2231390	Pterocarpus clade		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	2.910	0.311
163730	Hedysareae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.875	0.471
3896	Sophora		Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.355	0.715
3881	Onobrychis	Sainfoins	Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.250	NA
3502	Fagales	Beech order	Trees/shrubs	Trees/shrubs	1	1	0.944	0.469
3514	Betulaceae	Birch family	Trees/shrubs	Trees	1	1	0.926	0.656
3515	Alnus	Alders	Trees/shrubs	Trees	1	1	0.620	0.314
3517	Alnus glutinosa	Black alder	Trees/shrubs	Trees	1	1	0.673	0.433
3504	Betula	Birches	Trees/shrubs	Trees	1	1	1.032	0.623
13450	Corylus	Hazels	Trees/shrubs	Trees/shrubs	1	1	0.440	0.079
78606	Ostryopsis		Trees/shrubs	Shrubs	0	0	0.926	0.656
3503	Fagaceae	Oak family	Trees/shrubs	Trees	1	1	1.016	0.167
21019	Castanea	Chestnuts	Trees/shrubs	Trees	0	1	0.980	NA
114815	Castanopsis	Chinkapins	Trees/shrubs	Trees	0	0	1.088	0.068
21024	Fagus	Beeches	Trees/shrubs	Trees	1	1	0.563	0.072
3511	Quercus	Oaks	Trees/shrubs	Trees	1	1	0.945	0.145
38942	Quercus robur	English oak	Trees/shrubs	Trees	1	1	0.900	NA
16714	Juglandaceae	Walnut family	Trees/shrubs	Trees	0	1	0.733	0.110
16718	Juglans	Walnuts	Trees/shrubs	Trees	0	1	0.640	NA
26766	Myricaceae	Bayberry family	Trees/shrubs	Trees/shrubs	1	1	0.635	0.479
3646	Malpighiales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.531	1.435
22973	Chrysobalanaceae		Trees/shrubs	Trees/shrubs	0	0	1.531	1.435
3977	Euphorbiaceae	Spurge family	Herbs	Herbs	1	1	3.274	1.291
235631	Crotonoideae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.906	0.835
235887	Jatropheae	Nettlespurge family	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.510	0.509
3995	Jatropha	Nettlespurges	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.365	0.252
629714	Hypericaceae	St. John's wort family	Trees/shrubs and herbs	Shrubs/herbs	1	1	0.423	0.442
4004	Linaceae	Flax family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.540	0.567
4005	Linum	Flaxes	Herbs	Herbs	1	1	1.540	0.567
4268	Malpighiaceae	Acerola cherry family	Trees/shrubs	Trees/shrubs	0	0	2.705	0.620
3683	Passifloraceae	Passionflower family	Trees/shrubs	Trees/shrubs	0	0	1.286	0.489
387559	Basananthe		Trees/shrubs	Trees/shrubs	0	0	1.286	0.489
3684	Passiflora		Trees/shrubs	Trees/shrubs	0	0	1.199	0.491
999643	Podostemoideae	Riverweed subfamily	Freshwater aquatics	Herbs	0	0	1.531	1.435
26949	Rafflesiaceae		Herbs	Herbs	0	0	1.531	1.435
40029	Rhizophoraceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.531	1.435
40030	Rhizophora	True mangroves	Halophytes	Herbs	0	0	1.531	1.435
3688	Salicaceae	Willow family	Trees/shrubs	Trees/shrubs	1	1	0.554	0.300
238070	Flacourtieae		Trees/shrubs	Trees/shrubs	0	0	0.554	0.300
238069	Saliceae	Willow tribe	Trees/shrubs	Trees/shrubs	1	1	0.551	0.307
40685	Salix	Willows	Trees/shrubs	Trees/shrubs	1	1	0.578	0.363
3689	Populus	Poplars	Trees/shrubs	Trees	1	1	0.506	0.071
24921	Violaceae	Violet family	Herbs	Herbs	1	1	1.374	0.539
85258	Rinorea		Trees/shrubs	Trees/shrubs	0	0	1.374	0.539
13757	Viola	Violets	Herbs	Herbs	1	1	1.374	0.539
71243	Oxalidales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.085	1.528
4033	Oxalidaceae	Wood sorrel family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.311	1.463
3744	Rosales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	0.933	0.700
3481	Cannabaceae	Hemp family	Herbs	Herbs	1	1	2.284	1.122
3484	Humulus	Hops	Herbs	Herbs	1	1	2.067	0.350
3475	Parasponia		Trees/shrubs	Trees	0	0	2.284	1.122
25996	Elaeagnaceae	Sea-buckthorn family	Trees/shrubs	Trees/shrubs	1	1	1.250	0.775

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
3487	Moraceae	Mulberry family	Trees/shrubs	Trees/shrubs	0	0	0.748	0.286
66379	Broussonetia		Trees/shrubs	Trees/shrubs	0	0	0.700	NA
3493	Ficus	Weeping figs	Trees/shrubs	Trees/shrubs	0	0	0.705	0.197
3608	Rhamnaceae	Buckthorn family	Trees/shrubs	Trees/shrubs	1	1	0.520	0.601
72171	Ziziphus		Trees/shrubs	Trees/shrubs	0	1	1.550	NA
325289	Rhamneae	Buckthorn tribe	Trees/shrubs	Trees/shrubs	1	1	0.422	0.270
3609	Rhamnus	Buckthorns	Trees/shrubs	Trees/shrubs	1	1	0.420	0.298
3745	Rosaceae	Rose family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	0.930	0.532
721807	Sorbarieae		Trees/shrubs	Trees/shrubs	0	0	0.997	0.432
721813	Maleae	Apple tribe	Trees/shrubs	Trees	1	1	1.122	0.326
3749	Malus	Apples	Trees/shrubs	Trees	1	1	1.009	0.377
23159	Crataegus	Hawthorns	Trees/shrubs	Trees	1	1	1.157	0.284
3754	Prunus	Stone fruit trees	Trees/shrubs	Trees	1	1	0.548	1.155
721805	Amygdaleae		Trees/shrubs	Trees	1	1	0.452	0.500
3766	Pyrus	Pears	Trees/shrubs	Trees	0	1	0.605	0.032
721810	Osmaronieae		Trees/shrubs	Trees/shrubs	0	0	1.050	0.606
23205	Prinsepia		Trees/shrubs	Trees	0	0	1.500	NA
23166	Eriobotrya	Loquat genus	Trees/shrubs	Trees/shrubs	0	0	0.800	NA
721808	Spiraeeae	Bridal-wreath tribe	Herbs	Herbs	0	1	0.536	0.564
721791	Dryadoideae	Dryas subfamily	Trees/shrubs	Shrubs	1	1	0.600	NA
48230	Dryas		Trees/shrubs	Shrubs	1	1	0.600	NA
171637	Amygdaloideae		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	0.930	0.532
171638	Rosoideae	Rose subfamily	Trees/shrubs and herbs	Shrubs/herbs	1	1	0.719	0.725
3761	Geum	Avens	Herbs	Herbs	1	1	1.580	0.147
721790	Colurieae	Aven tribe	Herbs	Herbs	1	1	1.580	0.147
721789	Potentilleae	Cinquefoil tribe	Herbs	Herbs	1	1	0.786	0.755
23204	Potentilla	Cinquefoils	Herbs	Herbs	1	1	0.707	0.683
1184124	Fragariinae	Strawberry subtribe	Herbs	Herbs	1	1	0.679	0.957
1184125	Potentilleae incertae sedis		Herbs	Herbs	1	1	0.894	0.628
3746	Fragaria	Strawberries	Herbs	Herbs	1	1	0.332	0.570
3764	Rosa	Roses	Trees/shrubs	Shrubs	1	1	0.757	0.402
23216	Rubus	Blackberry and relatives	Trees/shrubs	Shrubs	1	1	0.451	0.905
23170	Filipendula	Meadowsweets	Herbs	Herbs	1	1	0.400	NA
1176516	Rosoideae incertae sedis		Trees/shrubs and herbs	Shrubs/herbs	1	1	0.531	0.567
721788	Sanguisorbeae		Trees/shrubs and herbs	Shrubs/herbs	1	1	1.048	0.660
1183420	Sanguisorbinae	Burnet subtribe	Trees/shrubs and herbs	Shrubs/herbs	1	1	0.857	0.517
1183439	Agrimoniinae	Agrimony subtribe	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.195	0.454
57937	Sanguisorba	Burnet	Trees/shrubs and herbs	Shrubs/herbs	1	1	0.560	NA
3474	Ulmaceae	Elm family	Trees/shrubs	Trees	1	1	1.992	0.441
24735	Ulmus	Elms	Trees/shrubs	Trees	1	1	2.320	0.240
3499	Urticaceae	Nettle family	Herbs	Herbs	1	1	0.596	0.663
226089	Elatostema		Herbs	Herbs	0	0	0.596	0.663
403666	Zygophyllales		Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.086	0.598
43873	Zygophyllaceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.086	0.598
293144	Larreoideae	Larrea subfamily	Trees/shrubs	Trees/shrubs	0	0	1.105	0.621
91836	malvids		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.102	0.926
3699	Brassicales	Mustard order	Trees/shrubs and herbs	Shrubs/herbs	1	1	0.770	0.862
3700	Brassicaceae	Mustard family	Trees/shrubs and herbs	Shrubs/herbs	1	1	0.735	0.852
72656	Aethionema	Stonecresses	Trees/shrubs and herbs	Shrubs/herbs	0	1	1.045	0.700
981067	Alysseae		Trees/shrubs and herbs	Shrubs/herbs	0	1	1.080	0.240
169067	Alyssum		Trees/shrubs and herbs	Shrubs/herbs	0	1	0.590	0.094
947475	Anchonieae	Stock tribe	Herbs	Herbs	1	1	1.994	0.158
981070	Arabideae		Herbs	Herbs	1	1	0.457	0.631

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
87301	Draba	Whitlow-grasses	Herbs	Herbs	1	1	0.414	0.500
981071	Brassiceae	Mustard tribe	Herbs	Herbs	1	1	1.034	0.811
3705	Brassica	Mustards	Herbs	Herbs	1	1	0.923	0.428
3712	Brassica oleracea	Wild and cultivated cabbages	Herbs	Herbs	1	1	0.923	0.428
3711	Brassica rapa		Herbs	Herbs	0	1	0.800	NA
3701	Arabidopsis	Thale cresses	Herbs	Herbs	1	1	0.325	0.330
980083	Camelineae		Herbs	Herbs	1	1	0.319	0.328
71323	Camelina	False flaxes	Herbs	Herbs	0	1	0.319	0.328
97854	Cochlearia	Scurvy-grasses	Halophytes	Herbs	1	1	0.704	0.582
981120	Cochlearieae	Scurvy-grass tribe	Halophytes	Herbs	1	1	0.950	0.428
19205	Lepidium	Pepperworts	Herbs	Herbs	1	1	0.670	0.568
301454	Cleomaceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.310	NA
85174	Francoaceae		Trees/shrubs	Trees/shrubs	0	0	0.640	NA
4027	Geraniaceae	Geranium family	Herbs	Herbs	1	1	1.237	1.115
21555	Erodium	Stork's-bills	Herbs	Herbs	1	1	1.130	0.088
4028	Geranium	Crane's-bills	Herbs	Herbs	1	1	1.759	0.682
21556	Monsonia		Herbs	Herbs	0	0	1.237	1.115
4030	Pelargonium	Stork's-bills	Herbs	Herbs	0	0	1.119	1.302
41938	Malvales	Mallow order	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.710	0.687
69450	Cistaceae	Rock-rose family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.332	0.372
69453	Helianthemum	Rock-rose	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.300	0.061
3629	Malvaceae	Mallow family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.532	0.742
69119	Pseudobombax	Shaving brush tree	Trees/shrubs	Trees	0	0	1.750	NA
214929	Bombacoideae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.976	0.576
214909	Byttnerioideae	Cacao subfamily	Trees/shrubs	Trees/shrubs	0	0	0.421	0.164
93758	Corchorus		Herbs	Herbs	0	0	0.269	0.417
214915	Helicteroideae	Durian subfamily	Trees/shrubs	Trees/shrubs	0	0	1.532	0.742
3633	Gossypium	Cottons	Trees/shrubs	Shrubs	0	0	1.715	0.391
214907	Malvoideae	Mallow subfamily	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.801	0.633
214911	Tilioideae	Lime subfamily	Trees/shrubs	Trees	1	1	1.250	NA
64580	Tilia	Limes	Trees/shrubs	Trees	1	1	1.250	NA
39987	Thymelaeaceae	Leatherwood family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	2.928	0.519
41944	Myrtales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	1.009	0.927
39992	Terminalia		Trees/shrubs	Trees/shrubs	0	0	3.967	0.402
3928	Lythraceae	Loosestrife family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	0.797	0.488
13128	Lythrum	Loosestrifes	Freshwater aquatics	Herbs	1	1	0.560	NA
39998	Melastomataceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.190	NA
119900	Memecylon		Trees/shrubs	Trees/shrubs	0	0	0.190	NA
3931	Myrtaceae		Trees/shrubs	Trees/shrubs	0	0	0.513	0.667
1699513	Myrtoideae		Trees/shrubs	Trees	0	0	0.543	0.676
1699523	Myrteae		Trees/shrubs	Trees	0	0	0.296	0.487
1699524	Eucalypteae		Trees/shrubs	Trees/shrubs	0	0	0.632	0.079
1699517	Melaleuceae		Trees/shrubs	Trees/shrubs	0	0	1.120	0.232
164925	Melaleuca	Paperbarks	Trees/shrubs	Shrubs	0	0	1.050	0.108
1585427	Onagroideae	Willowherb subfamily	Herbs	Herbs	1	1	1.048	0.532
3939	Oenothera		Herbs	Herbs	0	0	0.938	0.455
41937	Sapindales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	0.996	0.981
4011	Anacardiaceae	Cashew family	Trees/shrubs	Trees/shrubs	0	1	0.547	0.250
23513	Rutaceae	Citrus family	Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	1.387	1.271
67937	Zanthoxylum		Trees/shrubs	Trees/shrubs	0	0	4.573	0.817
1728962	Amyridoideae	Amyris subfamily	Herbs	Herbs	0	1	2.769	0.594
1728959	Aurantioideae	Citrus subfamily	Trees/shrubs	Trees/shrubs	0	0	0.493	0.202
2706	Citrus	Citrus	Trees/shrubs	Trees/shrubs	0	0	0.440	0.171

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
23672	Sapindaceae	Soapberry family	Trees/shrubs	Trees	1	1	1.224	1.492
1977916	Hippocastanoideae	Maple subfamily	Trees/shrubs	Trees/shrubs	1	1	0.890	0.411
1977919	Acereae	Maple tribe	Trees/shrubs	Trees/shrubs	1	1	0.928	0.403
23808	Simaroubaceae		Trees/shrubs	Trees/shrubs	0	0	1.417	0.467
3602	Vitaceae	Grape family	Trees/shrubs	Shrubs	0	1	0.538	0.446
3603	Vitis	Grapevines	Trees/shrubs	Shrubs	0	1	0.489	0.094
41947	Santalales	Sandalwood order	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	17.396	1.501
25673	Balanophoraceae	Fungus root family	Herbs	Herbs	0	0	17.396	1.501
25674	Balanophora	Fungus root genus	Herbs	Herbs	0	0	17.396	1.501
1003248	Cervantesiaceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	17.396	1.501
3963	Loranthaceae	Showy mistletoe family	Trees/shrubs	Shrubs	0	0	10.710	0.344
1003268	Lorantheae	Showy mistletoes	Trees/shrubs	Shrubs	0	0	8.904	0.376
1003276	Scurrulinae		Trees/shrubs	Shrubs	0	0	8.904	0.376
1003249	Nanodeaceae		Herbs	Herbs	0	0	17.396	1.501
50152	Olax		Trees/shrubs	Trees/shrubs	0	0	17.396	1.501
1003255	Viscaceae	Mistletoe family	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	67.202	0.526
3967	Phoradendron		Trees/shrubs	Shrubs	0	0	67.202	0.526
3971	Viscum	Mistletoes	Trees/shrubs	Shrubs	1	1	81.625	0.204
1003242	Ximeniaceae		Trees/shrubs	Trees/shrubs	0	0	1.630	NA
41946	Saxifragales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	5.191	1.626
3781	Crassulaceae	Stonecrop family	Xerophytes	Herbs	1	1	1.163	1.315
202994	Rhodiola	Stonecrops	Xerophytes	Herbs	1	1	1.163	1.315
3784	Sedum	Stonecrops	Xerophytes	Herbs	1	1	1.306	1.714
51502	Cynomorium	Desert thumb	Halophytes	Herbs	0	1	5.191	1.626
3801	Ribes	Currants	Trees/shrubs	Shrubs	1	1	0.885	0.163
24945	Haloragaceae	Water-milfoil family	Freshwater aquatics	Herbs	1	1	0.433	0.266
24952	Myriophyllum	Water-milfoils	Freshwater aquatics	Herbs	1	1	0.433	0.266
42216	Hamamelidaceae	Witch-hazel family	Trees/shrubs	Trees	1	1	1.300	0.616
3792	Saxifragaceae		Herbs	Herbs	1	1	1.457	0.447
3798	Saxifraga	Saxifrages	Herbs	Herbs	1	1	1.440	0.377
4447	Liliopsida	Monocots	Mixed	Mixed	1	1	9.529	1.308
4464	Acorus	Sweet flag	Freshwater aquatics	Herbs	0	0	0.470	0.211
16360	Alismatales	Alismatids	Mixed aquatics	Herbs	1	1	5.891	0.828
4449	Alismataceae	Water-plantain family	Freshwater aquatics	Herbs	1	1	14.406	0.311
4450	Sagittaria	Arrowheads	Freshwater aquatics	Herbs	1	1	16.388	0.219
51593	Aponogeton		Freshwater aquatics	Herbs	0	0	2.730	0.331
4454	Araceae	Arum family	Herbs	Herbs	1	1	5.425	0.764
284555	Aroideae	Arum subfamily	Herbs	Herbs	1	1	7.344	0.616
74663	Amorphophallus		Herbs	Herbs	0	0	8.867	0.496
284551	Lemnoideae	Duckweed family	Freshwater aquatics	Herbs	1	1	0.820	0.544
161109	Wolffia	Least duckweed	Freshwater aquatics	Herbs	1	1	1.135	0.373
284550	Orontioideae		Freshwater aquatics	Herbs	0	0	8.700	1.024
15095	Symplocarpus	Skunk cabbages	Freshwater aquatics	Herbs	0	0	2.400	NA
421921	Philodendroideae		Herbs	Herbs	0	0	5.105	0.720
293489	Philodendreae		Herbs	Herbs	0	0	4.251	0.631
71613	Philodendron		Herbs	Herbs	0	0	3.714	0.484
284552	Pothoideae		Herbs	Herbs	0	0	5.919	0.404
293478	Monstereae		Trees/shrubs and herbs	Shrubs/herbs	0	0	7.359	0.410
25926	Cymodoceaceae	Manatee-grass family	Salt/brackish aquatics	Herbs	1	1	1.435	0.882
16365	Ruppia	Tasselweeds	Salt/brackish aquatics	Herbs	1	1	2.330	NA
26319	Hydrocharitaceae	Elodea family	Freshwater aquatics	Herbs	1	1	5.059	0.707
26324	Elodea	Waterweeds	Freshwater aquatics	Herbs	0	0	4.200	NA
13165	Najas	Water-nymphs	Freshwater aquatics	Herbs	1	1	11.630	NA

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
55487	Posidonia	Neptune-grasses	Salt/brackish aquatics	Herbs	0	1	3.130	NA
16362	Potamogetonaceae	Pondweed family	Mixed aquatics	Herbs	1	1	2.017	0.686
13228	Potamogeton	Pondweeds	Freshwater aquatics	Herbs	1	1	1.855	1.033
246706	Stuckenia	Pondweeds	Mixed aquatics	Herbs	1	1	2.017	0.686
397089	Stuckenia filiformis	Slenderleaf-pondweed	Freshwater aquatics	Herbs	1	1	2.017	0.686
85243	Tofieldiaceae	Tofieldia family	Herbs	Herbs	1	1	1.150	0.430
27254	Zosteraceae	Eelgrass family	Salt/brackish aquatics	Herbs	1	1	0.545	0.584
27257	Zostera	Eelgrasses	Salt/brackish aquatics	Herbs	1	1	0.545	0.584
1437197	Petrosaviidae		Mixed	Mixed	1	1	9.792	1.307
73496	Asparagales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	12.477	0.859
4668	Amaryllidaceae	Daffodil family	Herbs	Herbs	1	1	20.871	0.518
40553	Allioideae	Onion family	Herbs	Herbs	1	1	20.572	0.411
4678	Allium	Onions, leeks, garlics	Herbs	Herbs	1	1	20.417	0.405
703251	Amaryllidoideae	Daffodil subfamiliy	Herbs	Herbs	1	1	21.676	0.630
82236	Lycoris		Herbs	Herbs	0	0	24.450	NA
40552	Asparagaceae	Asparagus family	Herbs	Herbs	1	1	9.527	0.902
703530	Agavoideae		Herbs	Herbs	0	1	6.884	0.505
39509	Agave	Agave	Xerophytes	Herbs	0	0	6.378	0.450
4685	Asparagus	Asparagus	Halophytes	Herbs	1	1	1.825	0.549
703534	Brodiaeoideae		Herbs	Herbs	0	0	10.245	0.556
703535	Lomandroideae		Herbs	Herbs	0	0	6.534	1.322
703537	Nolinoideae		Herbs	Herbs	1	1	10.355	0.582
51383	Asphodelaceae	Asphodel family	Herbs	Herbs	1	1	14.979	0.441
703538	Asphodeloideae	Asphodel subfamily	Herbs	Herbs	1	1	16.088	0.328
25641	Aloe		Xerophytes	Herbs	0	0	16.755	0.274
44985	Hyacinthaceae	Bluebell family	Herbs	Herbs	1	1	12.228	0.674
26339	Iridaceae	Iris family	Herbs	Herbs	1	1	5.656	0.945
2508080	Crocoideae	Crocus subfamily	Herbs	Herbs	1	1	3.569	0.759
2507920	Iridoideae	Iris subfamily	Herbs	Herbs	1	1	6.284	0.889
26378	Iris	Irises	Herbs	Herbs	1	1	9.050	0.535
4747	Orchidaceae	Orchids	Herbs	Herbs	1	1	5.878	1.348
158329	Apostasioideae		Herbs	Herbs	0	0	1.256	0.817
49693	Apostasia	Grass orchids	Herbs	Herbs	0	0	0.603	0.547
158330	Cypripedioideae	Lady's slipper orchids	Herbs	Herbs	1	1	24.779	0.364
53064	Paphiopedilum	Venus slipper	Herbs	Herbs	0	0	25.352	0.187
158332	Epidendroideae		Herbs	Herbs	1	1	3.137	1.083
38206	Calanthe	Christmas orchids	Herbs	Herbs	0	0	13.200	NA
158391	Cymbidieae		Herbs	Herbs	0	0	3.093	0.481
14366	Cymbidium	Boat orchids	Herbs	Herbs	0	0	4.100	0.158
158418	Oncidiinae		Herbs	Herbs	0	0	2.992	0.501
158389	Epidendreae		Herbs	Herbs	1	1	3.813	1.301
1769708	Calypsoinae		Herbs	Herbs	1	1	8.800	0.770
48525	Corallorhiza	Coralroot orchids	Herbs	Herbs	1	1	12.817	0.437
158409	Pleurothallidinae		Herbs	Herbs	0	0	1.398	0.466
78804	Masdevallia		Herbs	Herbs	0	0	1.565	0.113
158393	Malaxideae		Herbs	Herbs	1	1	1.903	0.588
38208	Bulbophyllum		Herbs	Herbs	0	0	2.533	0.471
1759432	Malaxidinae		Herbs	Herbs	1	1	1.903	0.588
158338	Neottieae		Herbs	Herbs	0	0	16.565	0.113
48533	Neottia		Herbs	Herbs	1	1	17.310	NA
158424	Aeridinae		Herbs	Herbs	0	0	3.686	0.630
225552	Gastrochilus		Herbs	Herbs	0	0	4.170	0.485
158331	Orchidoideae		Herbs	Herbs	1	1	7.307	0.454

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
158360	Goodyerinae		Herbs	Herbs	0	0	4.167	0.673
158349	Diurideae		Herbs	Herbs	0	0	7.307	0.454
78738	Cryptostylis		Herbs	Herbs	0	0	7.307	0.454
158345	Orchideae		Herbs	Herbs	1	1	7.783	0.411
158366	Orchidinae		Herbs	Herbs	1	1	7.783	0.411
59329	Ophrys	Bee orchid genus	Herbs	Herbs	1	1	8.942	0.249
158333	Vanilloideae	Vanilla subfamily	Herbs	Herbs	0	0	4.549	0.481
78725	Cleistes		Herbs	Herbs	0	0	4.549	0.481
1729789	Vanilleae	Vanilla tribe	Herbs	Herbs	0	0	4.549	0.481
51238	Vanilla	Vanilla	Herbs	Herbs	0	0	4.520	0.424
4734	commelinids		Mixed	Mixed	1	1	3.809	1.295
40551	Arecales	Palm order	Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	3.502	1.029
4710	Arecaceae	Palms	Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	3.465	0.939
169745	Chamaedoreeae		Trees/shrubs	Trees/shrubs	0	0	3.900	0.235
169698	Calamoideae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	2.793	0.614
169733	Calameae		Herbs	Herbs	0	0	2.334	0.493
169734	Lepidocaryeae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	6.000	0.306
169700	Coryphoideae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	3.319	0.737
417014	Chuniophoeniceae		Trees/shrubs	Trees/shrubs	0	0	1.500	NA
417010	Cryosophileae		Trees/shrubs	Trees/shrubs	0	0	7.400	NA
77335	Dasypogonaceae		Trees/shrubs	Trees/shrubs	0	0	0.990	0.936
4739	Commelinales		Mixed	Mixed	0	0	13.166	0.797
4740	Commelinaceae	Dayflower family	Herbs	Herbs	0	0	15.162	0.646
54924	Haemodoraceae	Bloodwort family	Herbs	Herbs	0	0	0.933	0.446
44974	Philydraceae		Herbs	Herbs	0	0	0.590	NA
38820	Poales	Grass order	Grasses and relatives	Grasses and relatives	1	1	3.416	1.133
4613	Bromeliaceae		Grasses and relatives	Bromeliads	0	0	0.615	0.339
1909378	Bromelioideae		Grasses and relatives	Bromeliads	0	0	0.549	0.268
4615	Ananas comosus	Pineapple	Grasses and relatives	Bromeliads	0	0	0.550	NA
4609	Cyperaceae	Sedges	Grasses and relatives	Sedges	1	1	0.648	1.491
986140	Cyperoideae		Grasses and relatives	Sedges	1	1	0.593	1.379
986144	Abildgaardieae		Grasses and relatives	Sedges	1	1	0.593	1.379
13398	Carex	True sedges	Grasses and relatives	Sedges	1	1	0.439	0.367
986145	Cypereae		Grasses and relatives	Sedges	1	1	0.644	0.519
1982037	C3 Cyperus		Freshwater aquatics	Sedges	1	1	0.900	0.786
4610	Cyperus	Galingales	Grasses and relatives	Sedges	1	1	0.740	0.537
2034351	C4 Cyperus incertae sedis	Galingales	Grasses and relatives	Sedges	1	1	0.633	0.241
46324	Eleocharis	Spike sedges	Freshwater aquatics	Sedges	1	1	3.026	0.779
986142	Fuireneae		Mixed aquatics	Sedges	1	1	0.560	NA
76416	Bolboschoenus	Club-rushes	Mixed aquatics	Sedges	1	1	0.560	NA
76500	Schoenoplectus	Club-rushes	Freshwater aquatics	Sedges	1	1	0.560	NA
986147	Schoeneae	Bogrush tribe	Freshwater aquatics	Sedges	1	1	2.304	1.142
58221	Cladium	Fen sedges	Freshwater aquatics	Sedges	1	1	2.304	1.142
986141	Scirpeae		Grasses and relatives	Sedges	1	1	0.533	0.286
46326	Eriophorum	Cottonsedge	Freshwater aquatics	Sedges	1	1	0.550	0.386
46334	Scirpus		Freshwater aquatics	Sedges	1	1	0.467	0.124
986138	Hypolytreae		Grasses and relatives	Sedges	0	0	0.648	1.491
26019	Eriocaulaceae	Pipewort family	Grasses and relatives	Grasses	1	1	1.655	1.036
26021	Eriocaulon	Pipeworts	Freshwater aquatics	Grasses	1	1	2.655	0.818
14101	Juncaceae	Rushes	Grasses and relatives	Rushes	1	1	0.958	0.938
13578	Juncus	Rushes	Grasses and relatives	Rushes	1	1	0.660	0.750
4479	Poaceae	Grass family	Grasses and relatives	Grasses	1	1	5.138	0.795
147364	Anomochlooideae		Grasses and relatives	Grasses	0	0	5.138	0.795

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
359160	BOP clade		Grasses and relatives	Grasses	1	1	5.967	0.682
147366	Bambusoideae	Bamboos	Grasses and relatives	Bamboos	0	0	1.690	0.315
4590	Arundinaria	Canes	Grasses and relatives	Bamboos	0	0	2.738	0.220
1648003	Arundinariinae		Grasses and relatives	Bamboos	0	0	2.738	0.220
111430	Fargesia		Grasses and relatives	Bamboos	0	0	2.738	0.220
1648035	Bambusodae		Grasses and relatives	Bamboos	0	0	1.508	0.150
147376	Bambuseae		Grasses and relatives	Bamboos	0	0	1.547	0.085
35709	Chusquea	South American mountain bamboos	Grasses and relatives	Bamboos	0	0	1.547	0.085
4581	Bambusa		Grasses and relatives	Bamboos	0	0	1.534	0.043
147377	Olyreae		Grasses and relatives	Bamboos	0	0	0.650	NA
1648020	Olyrinae		Grasses and relatives	Bamboos	0	0	0.650	NA
1648023	Parianinae		Grasses and relatives	Bamboos	0	0	0.650	NA
147367	Oryzoideae	Rice subfamily	Grasses and relatives	Grasses	1	1	1.045	0.450
147379	Ehrharteae		Grasses and relatives	Grasses	0	0	1.281	0.290
147380	Oryzeae	Rice tribe	Freshwater aquatics	Grasses	1	1	0.962	0.499
4527	Oryza	Rice genus	Freshwater aquatics	Grasses	0	0	0.838	0.465
4530	Oryza sativa	Cultivated rice	Freshwater aquatics	Grasses	0	0	0.838	0.465
1648021	Oryzinae	Rice subtribe	Freshwater aquatics	Grasses	0	0	0.919	0.426
1648031	Zizaniinae	Wild rice subtribe	Freshwater aquatics	Grasses	0	0	1.350	0.890
1648045	Phyllorachideae		Grasses and relatives	Grasses	0	0	1.045	0.450
147368	Pooideae	Cool-season grasses	Grasses and relatives	Grasses	1	1	6.519	0.607
15367	Brachypodium	False brome grasses	Grasses and relatives	Grasses	1	1	0.551	0.537
4501	Bromus	Brome grasses	Grasses and relatives	Grasses	1	1	8.138	0.400
13648	Puccinellia	Saltmarsh grass	Halophytes	Grasses	1	1	3.557	0.445
147387	Poeae		Grasses and relatives	Grasses	1	1	5.420	0.604
1652081	Poeae Chloroplast Group 2 (Poeae type)		Grasses and relatives	Grasses	1	1	4.683	0.625
4605	Festuca	Fescues	Grasses and relatives	Grasses	1	1	5.595	0.467
1652080	Poeae Chloroplast Group 1 (Aveneae type)		Grasses and relatives	Grasses	1	1	6.049	0.581
640630	Loliinae	Ryegrass subtribe	Grasses and relatives	Grasses	1	1	5.506	0.495
640628	Poinae	Meadow grass subtribe	Grasses and relatives	Grasses	1	1	4.163	0.852
4544	Poa	Meadow grasses	Grasses and relatives	Grasses	1	1	3.648	0.736
640623	Aveninae		Grasses and relatives	Grasses	1	1	6.935	0.482
640621	Agrostidinae		Grasses and relatives	Grasses	1	1	6.209	0.474
4496	Avena	Oats	Grasses and relatives	Grasses	1	1	7.655	0.458
4520	Lolium	Ryegrasses	Grasses and relatives	Grasses	1	1	3.211	0.373
1648010	Coleanthinae		Grasses and relatives	Grasses	1	1	2.408	0.564
15452	Deschampsia	Tussock grass	Grasses and relatives	Grasses	1	1	5.882	0.262
15303	Alopecurus	Foxtail grasses	Grasses and relatives	Grasses	1	1	8.900	0.716
15376	Calamagrostis	Reed grasses	Grasses and relatives	Grasses	1	1	4.203	0.723
49767	Helictotrichon	Alpine oatgrasses	Grasses and relatives	Grasses	0	1	6.300	NA
147383	Stipeae		Grasses and relatives	Grasses	1	1	1.583	0.461
15869	Stipa	Feather grasses	Grasses and relatives	Grasses	0	1	1.140	NA
1648038	Triticodae		Grasses and relatives	Grasses	1	1	8.911	0.438
147389	Triticeae		Grasses and relatives	Grasses	1	1	8.911	0.438
1648017	Hordeinae	Barley subtribe	Grasses and relatives	Grasses	1	1	8.199	0.412
4512	Hordeum	Barlies	Grasses and relatives	Grasses	1	1	7.810	0.458
1648030	Triticinae	Wheat subtribe	Grasses and relatives	Grasses	1	1	10.320	0.435
4480	Aegilops	Goat grasses	Grasses and relatives	Grasses	0	1	10.349	0.461
15492	Elymus	Couch grasses	Grasses and relatives	Grasses	1	1	11.364	0.302
4564	Triticum	Wheats	Grasses and relatives	Grasses	0	1	11.245	0.366
4587	Thinopyrum	Wheat grasses	Grasses and relatives	Grasses	1	1	12.400	0.462
4549	Secale	Ryes	Grasses and relatives	Grasses	0	1	7.960	0.082
4513	Hordeum vulgare	Cultivated barley	Grasses and relatives	Grasses	0	0	6.720	0.314

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
4568	Triticum monococcum	Einkorn wheat	Grasses and relatives	Grasses	0	1	6.550	0.076
4571	Triticum turgidum	Durum wheat	Grasses and relatives	Grasses	0	0	12.480	0.020
4550	Secale cereale	Cultivated rye	Grasses and relatives	Grasses	0	0	8.500	0.050
147370	PACMAD clade		Grasses and relatives	Grasses	1	1	2.016	0.951
147430	Aristideae	Three-awns wire grass tribe	Grasses and relatives	Grasses	0	1	1.310	0.011
156631	Arundinoideae	Reeds	Mixed aquatics	Reeds	1	1	2.828	0.302
147431	Arundineae		Freshwater aquatics	Reeds	0	1	2.800	NA
57048	Amphipogon		Grasses and relatives	Grasses	0	0	2.800	NA
15745	Phragmites	Common reeds	Mixed aquatics	Reeds	1	1	2.380	0.621
1648043	Molinieae	Common reed tribe	Mixed aquatics	Reeds	1	1	2.837	0.369
147371	Chloridoideae		Grasses and relatives	Grasses	1	1	2.081	1.767
147435	Cynodonteae		Grasses and relatives	Grasses	0	0	2.257	1.801
66044	Triodia	Spinifex	Grasses and relatives	Grasses	0	0	2.257	1.801
751762	Eleusininae	Goose grass subtribe	Grasses and relatives	Grasses	0	0	1.412	0.623
48731	Bouteloua	Grama grasses	Grasses and relatives	Grasses	0	0	6.860	1.621
147436	Eragrostideae	Love grass subtribe	Grasses and relatives	Grasses	0	1	2.093	1.182
38413	Eragrostis	Love grasses	Grasses and relatives	Grasses	0	1	0.860	0.293
751754	Zoysieae	Zoysia tribe	Grasses and relatives	Grasses	1	1	0.660	0.514
38730	Sporobolus	Dropseeds	Grasses and relatives	Grasses	1	1	1.050	NA
751756	Sporobolinae	Dropseed subtribe	Grasses and relatives	Grasses	1	1	1.050	NA
219384	Danthonioideae		Grasses and relatives	Grasses	1	1	2.408	0.295
147432	Danthonieae		Grasses and relatives	Grasses	0	0	2.408	0.295
318921	Micrairoideae		Grasses and relatives	Grasses	0	0	1.820	NA
66023	Eriachne	Wanderrie grasses	Grasses and relatives	Grasses	0	0	1.820	NA
153999	Eriachneae		Grasses and relatives	Grasses	0	0	1.820	NA
219398	Isachneae		Grasses and relatives	Grasses	0	0	1.820	NA
147369	Panicoideae		Grasses and relatives	Grasses	0	1	1.853	0.626
147429	Andropogoneae	Sorghum tribe	Grasses and relatives	Grasses	0	1	2.505	0.578
62336	Miscanthus		Grasses and relatives	Grasses	0	0	3.024	0.221
15314	Andropogon	Broomsedges	Grasses and relatives	Grasses	0	1	3.862	0.417
4575	Zea	Maize and teosintes	Grasses and relatives	Grasses	0	0	3.800	0.358
4577	Zea mays	Maize	Grasses and relatives	Grasses	0	0	3.800	0.358
1648026	Saccharinae	Sugarcane subtribe	Grasses and relatives	Grasses	0	0	2.655	0.051
128810	Saccharum hybrid cultivar	Sugarcane	Grasses and relatives	Grasses	0	0	2.655	0.051
286192	Saccharum officinarum complex	Sugarcane	Grasses and relatives	Grasses	0	0	2.655	0.051
1648029	Tripsacinae	Maize subtribe	Grasses and relatives	Grasses	0	0	3.923	0.344
147428	Paniceae		Grasses and relatives	Grasses	0	1	1.492	0.443
4554	Setaria	Bristle grasses	Grasses and relatives	Grasses	0	1	1.459	0.446
1293361	Cenchrinae		Grasses and relatives	Grasses	0	1	1.738	0.424
4539	Panicum	Panicgrass	Grasses and relatives	Grasses	0	1	1.553	0.417
45618	Echinochloa	Barnyard grasses	Grasses and relatives	Grasses	0	1	1.328	0.069
1293362	Melinidinae		Grasses and relatives	Grasses	0	1	1.150	NA
4523	Neurachne		Grasses and relatives	Grasses	0	0	0.710	NA
1293364	Neurachninae	Mulga grass subtribe	Grasses and relatives	Grasses	0	0	0.710	NA
1293358	Otachyriinae		Grasses and relatives	Grasses	0	0	1.151	0.301
1293356	Paspaleae		Grasses and relatives	Grasses	0	0	1.151	0.301
1293359	Paspalinae		Grasses and relatives	Grasses	0	0	1.151	0.301
318915	Tristachyideae		Grasses and relatives	Grasses	0	0	1.853	0.626
147374	Phareae	Stalk grass tribe	Grasses and relatives	Grasses	0	0	1.260	NA
4731	Typhaceae	Bulrush family	Freshwater aquatics	Bulrushes	1	1	0.343	0.408
4729	Sparganium	Bur-reeds	Freshwater aquatics	Bulrushes	1	1	0.500	NA
4732	Typha	Bulrushes	Freshwater aquatics	Bulrushes	1	1	0.265	0.187
4618	Zingiberales	Banana order	Herbs	Herbs	0	0	1.327	2.085

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
4630	Costaceae	Banana family	Herbs	Herbs	0	0	1.640	0.419
4653	Heliconia	Lobster-claws	Herbs	Herbs	0	0	0.470	0.037
4657	Orchidantha	Orchid-flowers	Herbs	Herbs	0	0	3.600	NA
4619	Marantaceae	Arrowroot family	Herbs	Herbs	0	0	0.445	0.214
4637	Musaceae	Banana family	Herbs	Herbs	0	0	0.670	0.121
4638	Ensete	False banana	Herbs	Herbs	0	0	0.570	0.099
4640	Musa	Bananas and plantains	Herbs	Herbs	0	0	0.678	0.115
4641	Musa acuminata	Bananas	Herbs	Herbs	0	0	0.850	NA
4642	Zingiberaceae	Ginger family	Herbs	Herbs	0	0	1.842	1.943
94326	Alpinia	Shell gingers	Herbs	Herbs	0	0	2.800	NA
40548	Dioscoreales		Herbs	Herbs	1	1	1.070	1.238
48531	Burmannia	Bluethreads	Herbs	Herbs	0	0	1.740	NA
4671	Dioscoreaceae		Herbs	Herbs	1	1	1.097	1.164
4672	Dioscorea		Herbs	Herbs	1	1	1.107	1.178
167601	Trichopus		Herbs	Herbs	0	0	1.097	1.164
114201	Nartheciaceae		Herbs	Herbs	1	1	0.390	0.073
119987	Aletris		Herbs	Herbs	0	0	0.390	0.073
44981	Tacca	Arrowroot genus	Herbs	Herbs	0	0	0.620	0.349
4667	Liliales		Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	29.080	0.783
56740	Alstroemeriaceae		Herbs	Herbs	0	0	24.089	0.343
41218	Colchicaceae	Autumn crocus family	Herbs	Herbs	1	1	7.014	0.864
4699	Uvularia	Bellworts	Herbs	Herbs	0	0	12.240	0.344
4677	Liliaceae	Lily family	Herbs	Herbs	1	1	33.862	0.516
59070	Fritillaria	Fritillaries	Herbs	Herbs	1	1	56.580	0.283
4688	Lilium	Lillies	Herbs	Herbs	0	1	37.995	0.196
50362	Melanthiaceae	Bunchflower family	Herbs	Herbs	1	1	30.543	1.065
49674	Trillium		Herbs	Herbs	0	0	58.540	0.341
50241	Veratrum	False hellebores	Herbs	Herbs	0	1	2.789	0.217
40550	Pandanales	Pandan order	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	0.771	0.418
49662	Stemonaceae	Baibu family	Herbs	Herbs	0	0	0.730	NA
85281	Stemona	Baibu	Herbs	Herbs	0	0	0.730	NA
114079	Petrosaviaceae		Herbs	Herbs	0	0	1.900	NA
114080	Petrosavia		Herbs	Herbs	0	0	1.900	NA
232347	Magnoliidae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.882	0.870
3424	Canellaceae		Trees/shrubs	Trees/shrubs	0	0	5.800	NA
3432	Laurales		Trees/shrubs	Trees	0	1	1.534	0.511
3427	Calycanthaceae	Sweetshrubs	Trees/shrubs	Trees/shrubs	0	0	0.947	0.080
3433	Lauraceae	Laurel family	Trees/shrubs	Trees	0	1	1.687	0.551
3400	Magnoliales		Trees/shrubs	Trees/shrubs	0	0	1.522	0.747
22140	Annonaceae	Custard apple family	Trees/shrubs	Trees/shrubs	0	0	1.200	0.506
3401	Magnoliaceae	Magnolia family	Trees/shrubs	Trees/shrubs	0	0	3.644	0.595
3402	Magnolia	Magnolia	Trees/shrubs	Trees	0	0	4.095	0.542
16736	Piperales		Herbs	Herbs	0	1	2.226	0.929
16727	Aristolochiaceae	Birthwort family	Herbs	Herbs	0	1	3.216	0.848
12947	Aristolochia	Birthworts	Herbs	Herbs	0	1	0.370	0.264
13215	Piper	Pepper vines	Herbs	Herbs	0	0	1.250	0.474
232378	Proteales		Trees/shrubs and herbs	Trees/shrubs and herbs	0	1	3.135	2.107
4430	Nelumbo		Freshwater aquatics	Herbs	0	0	0.605	0.853
4402	Platanus	Plane trees	Trees/shrubs	Trees	0	1	1.610	0.272
4328	Proteaceae		Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	2.844	2.166
22910	Sabiaceae		Trees/shrubs	Trees/shrubs	0	0	3.135	2.107
41768	Ranunculales	Buttercup order	Trees/shrubs and herbs	Trees/shrubs and herbs	1	1	6.431	0.905
41773	Berberidaceae	Barberry family	Trees/shrubs	Shrubs	1	1	4.343	1.453

Continued on next page

Table D.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
1461175	Berberidoideae	Barberry subfamily	Trees/shrubs	Shrubs	1	1	3.974	1.269
22774	Berberis	Barberry	Trees/shrubs	Shrubs	1	1	1.435	0.503
168832	Leontice		Herbs	Herbs	0	1	3.974	1.269
3455	Menispermaceae	Moonseeds	Trees/shrubs and herbs	Trees/shrubs and herbs	0	0	1.100	NA
1462614	Papaveroideae	Poppy subfamily	Herbs	Herbs	1	1	2.924	0.675
3468	Papaver	Poppies	Herbs	Herbs	1	1	3.332	0.459
3440	Ranunculaceae	Buttercup family	Trees/shrubs and herbs	Shrubs/herbs	1	1	8.705	0.781
261450	Coptis chinensis	Chinese goldthread	Herbs	Herbs	0	0	0.400	NA
1463138	Ranunculoideae	Buttercup subfamily	Herbs	Herbs	1	1	9.272	0.564
1463147	Anemoneae	Anemone tribe	Herbs	Herbs	1	1	12.661	0.478
1463144	Caltheae		Herbs	Herbs	1	1	16.500	NA
3448	Caltha		Herbs	Herbs	1	1	16.500	NA
3449	Caltha palustris	Marsh marigold	Herbs	Herbs	1	1	16.500	NA
1463143	Cimicifugeae	Baneberry tribe	Herbs	Herbs	1	1	10.000	0.070
46988	Actaea	Baneberries	Herbs	Herbs	1	1	10.350	0.048
1463140	Delphinieae	Delphinium tribe	Herbs	Herbs	0	1	2.280	NA
1463148	Ranunculeae	Buttercup tribe	Herbs	Herbs	1	1	6.367	0.565
3445	Ranunculus	Buttercups	Herbs	Herbs	1	1	6.714	0.548
1463137	Thalictroideae		Herbs	Herbs	1	1	1.037	0.915
46968	Thalictrum	Meadow-rues	Herbs	Herbs	1	1	1.091	0.889
261007	Nymphaeales	Waterlily order	Freshwater aquatics	Herbs	1	1	1.769	0.626
4422	Cabombaceae	Fanwort family	Freshwater aquatics	Herbs	1	1	2.385	0.691
4425	Cabomba		Freshwater aquatics	Herbs	0	0	3.550	NA
4410	Nymphaeaceae	Waterlily family	Freshwater aquatics	Herbs	1	1	1.884	0.629
4411	Barclaya		Freshwater aquatics	Herbs	0	0	1.884	0.629
4415	Nuphar	Waterlilies	Freshwater aquatics	Herbs	1	1	2.817	0.070
4418	Nymphaea	Waterlilies	Freshwater aquatics	Herbs	1	1	1.293	0.510
1521260	Lycopodiopsida	Lycopods	Herbs	Herbs	1	1	0.971	2.128
13838	Isoetes	Quillworts	Freshwater aquatics	Herbs	1	1	6.860	1.053
3250	Lycopodiaceae	Core clubmoss family	Herbs	Herbs	1	1	3.769	0.263
1965347	Lycopodioideae		Herbs	Herbs	1	1	3.365	0.187
3246	Selaginella	Spikemosses	Herbs	Herbs	1	1	0.124	0.391
131209	Zygnemophyceae		Algae	Algae	NA	1	6.760	1.125
131210	Desmidiales	Desmids	Algae	Algae	1	1	7.769	1.016
31314	Desmidiaceae		Algae	Algae	NA	NA	7.344	1.050
33100	Cosmarium		Algae	Algae	1	1	1.803	0.647
130992	Euastrum		Algae	Algae	NA	NA	15.410	0.733
3176	Zygnematales		Algae	Algae	1	1	1.421	0.674
31319	Mesotaeniaceae		Algae	Algae	1	1	0.675	0.262
37306	Cylindrocystis		Freshwater aquatics	Algae	1	1	0.675	0.262
3178	Zygnemataceae		Algae	Algae	1	1	1.783	0.504
2041619	unclassified Mougeotia		Algae	Algae	NA	NA	3.200	NA
3179	Spirogyra	Blanketweeds	Freshwater aquatics	Algae	1	1	2.050	NA
2041618	unclassified Spirogyra	Blanketweeds	Freshwater aquatics	Algae	1	1	2.050	NA
43951	Zygnema		Mixed	Algae	1	1	1.400	0.141
151659	environmental samples		Mixed	Mixed	NA	NA	NA	NA

Appendix E

Taxa found in the Metazoa (no Primates) data

Table E.1: List of taxa found in the Metazoa (excluding Primates) data, sorted by full taxonomy. Includes NCBI taxonomic ID, full taxon name, common name, ecological category, informal taxon group (e.g. crustaceans; bivalves), native status in Great Britain and Europe, mean estimated C-value (haploid genome size), and the coefficient of variance (CV) for that C-value. Native status was not necessary for taxa outside Metazoa. Mean C-value was calculated for taxa in Metazoa only; due to the imprecise nature of initial taxonomic assignment, this dataset includes some taxa outside Metazoa and also some inside Primates. Non-native status is highlighted in gold. Mean C-values are shaded by size: values <1 are in blue and generate an increase from raw read count to biogenomic mass, and values >1 are in red and generate a decrease. Magnitude is shown by saturation. CV values ≥ 1 suggest imprecision and are highlighted in gold.

Table E.1

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
2157	Archaea	Archaea	Mixed	Archaea	1	1	NA	NA
2	Bacteria	Bacteria	Mixed	Bacteria	1	1	NA	NA
48479	environmental samples		Mixed	Bacteria	NA	NA	NA	NA
1196022	unclassified Rhodothermaceae		Mixed	Bacteria	1	1	NA	NA
49546	Flavobacteriaceae		Mixed	Bacteria	1	1	NA	NA
1224	Proteobacteria		Mixed	Bacteria	1	1	NA	NA
28211	Alphaproteobacteria		Mixed	Bacteria	1	1	NA	NA
356	Rhizobiales		Mixed	Bacteria	1	1	NA	NA
41294	Bradyrhizobiaceae		Mixed	Bacteria	NA	NA	NA	NA
407	Methylobacterium		Mixed	Bacteria	NA	NA	NA	NA
68287	Mesorhizobium		Terrestrial	Bacteria	NA	NA	NA	NA
80840	Burkholderiales		Mixed	Bacteria	NA	NA	NA	NA
1236	Gammaproteobacteria		Mixed	Bacteria	1	1	NA	NA
2742	Marinobacter		Salt/brackish aquatics	Bacteria	NA	NA	NA	NA
1177	Nostoc		Mixed	Bacteria	NA	NA	NA	NA
33154	Opisthokonta	Opisthokonts	Mixed	Eukaryotes	1	1	NA	NA
33208	Metazoa	Animals	Mixed	Animals	1	1	4.031	2.337
6072	Eumetazoa		Mixed	Animals	1	1	4.080	2.321
33511	Deuterostomia	Deuterostomes	Mixed	Animals	1	1	5.170	2.152
7711	Chordata	Chordates	Mixed	Animals	1	1	5.208	2.146
7737	Branchiostoma		Salt/brackish aquatics	Invertebrates	1	1	0.590	NA
7742	Vertebrata	Vertebrates	Mixed	Vertebrates	1	1	5.213	2.145
7776	Gnathostomata	Jawed vertebrates	Mixed	Vertebrates	1	1	5.089	2.243
7898	Actinopterygii	Ray-finned fishes	Mixed aquatics	Fishes	1	1	1.289	0.609
186623	Actinopteri		Mixed aquatics	Fishes	1	1	1.258	0.575
41665	Neopterygii		Mixed aquatics	Fishes	1	1	1.235	0.526
1489341	Osteoglossocephalai		Mixed aquatics	Fishes	1	1	1.210	0.520
8045	Gadidae	Cod family	Salt/brackish aquatics	Fishes	1	1	0.813	0.250
28738	Cyprinodontiformes	Toothcarps	Mixed aquatics	Fishes	0	1	1.281	0.619
1489913	Atherinomorphae		Mixed aquatics	Fishes	1	1	1.246	0.605
1489388	Euteleosteomorpha		Mixed aquatics	Fishes	1	1	1.115	0.592
1489908	Ovalentaria		Mixed aquatics	Fishes	1	1	1.070	0.520
1489872	Percomorphaceae		Mixed aquatics	Fishes	1	1	0.947	0.433
123369	Euacanthomorphacea		Mixed aquatics	Fishes	1	1	0.947	0.431
123368	Acanthomorphata		Mixed aquatics	Fishes	1	1	0.945	0.428
1489922	Eupercaria	Perch series	Mixed aquatics	Fishes	1	1	0.921	0.341
41705	Protacanthopterygii		Mixed aquatics	Fishes	1	1	2.471	0.315
8087	Cyprinodontoidei		Mixed aquatics	Fishes	0	1	0.925	0.280
1489904	Carangaria		Mixed aquatics	Fishes	1	1	0.755	0.246
8015	Salmonidae	Salmon family	Mixed aquatics	Fishes	1	1	2.844	0.155
504568	Salmoninae	Salmon subfamily	Mixed aquatics	Fishes	1	1	2.829	0.128
318546	Pseudocrenilabrinae		Mixed aquatics	Fishes	0	0	1.055	0.123
8016	Oncorhynchus	Pacific salmon and trout	Mixed aquatics	Fishes	0	0	2.682	0.111
8028	Salmo	European salmon and trout	Mixed aquatics	Fishes	1	1	3.005	0.066
75365	Sinocyclocheilus	Golden-line fish	Freshwater aquatics	Fishes	0	0	2.300	NA
186628	Characiphysae		Freshwater aquatics	Fishes	0	1	1.493	0.400
7952	Cypriniformes	Carp order	Freshwater aquatics	Fishes	1	1	1.407	0.361
564289	Cyprinidae intergeneric hybrids	Carp family hybrids	Freshwater aquatics	Fishes	1	1	1.397	0.306
7956	Carassius	Crucian carps	Freshwater aquatics	Fishes	1	1	1.973	0.190
186634	Otomorpha		Mixed aquatics	Fishes	1	1	1.413	0.363
32519	Ostariophysi	Carp subcohort	Mixed aquatics	Fishes	1	1	1.426	0.362

Continued on next page

Table E.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
186626	Otophysi		Mixed aquatics	Fishes	1	1	1.426	0.362
8287	Sarcopterygii	Lobe-finned fish	Mixed	Vertebrates	1	1	7.962	1.864
1338369	Dipnotetrapodomorpha	Tetropods and lungfish	Mixed	Vertebrates	1	1	7.957	1.865
40674	Mammalia	Mammals	Mixed	Mammals	1	1	3.205	0.263
9721	Cetacea	Whales	Mixed aquatics	Mammals	1	1	3.336	0.097
9722	Odontoceti	Toothed whales	Mixed aquatics	Mammals	1	1	3.403	0.086
9612	Canis lupus	Grey wolf	Terrestrial	Mammals	1	1	2.810	NA
33550	Hystricomorpha	Guinea pig suborder	Terrestrial	Mammals	0	1	4.049	0.249
9989	Rodentia	Rodents	Terrestrial	Mammals	1	1	3.615	0.235
9935	Ovis	Sheep genus	Terrestrial	Mammals	0	1	3.008	0.172
9443	Primates	Primates	Terrestrial	Mammals	1	1	3.584	0.152
9528	Cercopithecinae	Cercopithecine monkeys	Terrestrial	Mammals	1	1	3.868	0.149
9527	Cercopithecidae	Old World monkeys	Terrestrial	Mammals	1	1	3.893	0.143
9963	Caprinae	Goat subfamily	Terrestrial	Mammals	0	1	3.073	0.142
9526	Catarrhini	Old World monkeys and apes	Terrestrial	Mammals	1	1	3.752	0.141
314293	Simiiformes	Simians	Terrestrial	Mammals	1	1	3.706	0.135
376913	Haplorrhini	Dry-nosed primates	Terrestrial	Mammals	1	1	3.706	0.135
9431	Vespertilionidae	Vespertilionid bats	Terrestrial	Mammals	1	1	2.371	0.121
30560	Microchiroptera	Microbats	Terrestrial	Mammals	1	1	2.488	0.120
9397	Chiroptera	Bats	Terrestrial	Mammals	1	1	2.437	0.118
9789	Equus	Horses, zebras and asses	Terrestrial	Mammals	1	1	3.349	0.114
379583	Feliformia	Cat-like carnivorans	Terrestrial	Mammals	1	1	2.965	0.113
27592	Bovinae	Cattle subfamily	Terrestrial	Mammals	1	1	3.727	0.112
39107	Murinae	Old World rats and mice	Terrestrial	Mammals	1	1	3.238	0.095
10167	Bathyergidae	African mole-rats	Terrestrial	Mammals	0	0	3.175	0.089
9479	Platyrrhini	New World monkeys	Terrestrial	Mammals	0	0	3.522	0.087
9821	Suidae	Pig family	Terrestrial	Mammals	1	1	3.128	0.075
9822	Sus	Pigs	Terrestrial	Mammals	1	1	3.128	0.075
9823	Sus scrofa	Wild boar and domestic pig	Terrestrial	Mammals	1	1	3.128	0.075
9903	Bos	Wild and domestic cattle	Terrestrial	Mammals	1	1	3.601	0.063
77225	Pteropodinae	Flying fox subfamily	Terrestrial	Mammals	0	0	2.211	0.054
9596	Pan	Chimpanzees	Terrestrial	Mammals	0	0	3.675	0.046
9401	Pteropus	Flying foxes	Terrestrial	Mammals	0	0	2.303	0.042
314147	Glires	Rodents and lagomorphs	Mixed	Mammals	1	1	3.559	0.226
337677	Cricetidae	Hamster family	Mixed	Mammals	1	1	3.214	0.153
9787	Perissodactyla	Odd-toed ungulates	Mixed	Mammals	1	1	3.284	0.132
33554	Carnivora	Carnivorans	Mixed	Mammals	1	1	3.007	0.090
9655	Mustelidae	Weasel family	Mixed	Mammals	1	1	2.945	0.065
9263	Metatheria	Marsupials	Terrestrial	Mammals	0	0	4.005	0.167
38609	Diprotodontia		Terrestrial	Mammals	0	0	3.901	0.197
32561	Sauria	Diapsids	Mixed	Vertebrates	1	1	1.656	0.362
8476	Emydidae	Terrapins	Freshwater aquatics	Testudines	0	1	2.683	0.246
1526411	Mareca	Wigeon genus	Freshwater aquatics	Birds	1	1	1.386	0.160
8495	Alligator	Alligators	Freshwater aquatics	Crocodilians	0	0	2.575	0.047
8830	Anatidae	Waterfowl	Mixed aquatics	Birds	1	1	1.388	0.130
8783	Palaeognathae	Ratites and tinamous	Terrestrial	Birds	0	0	1.586	0.225
9030	Gallus	Junglefowl and chicken	Terrestrial	Birds	0	0	1.256	0.121
9072	Phasianinae	Pheasant subfamily	Terrestrial	Birds	1	1	1.256	0.121
9223	Psittaciformes	Parrot order	Terrestrial	Birds	0	0	1.439	0.107
8976	Galliformes	Landfowl	Terrestrial	Birds	1	1	1.303	0.104
40155	Estrildinae		Terrestrial	Birds	0	0	1.335	0.090
1329799	Archelosauria		Mixed	Vertebrates	1	1	1.467	0.320
1579337	Durocryptodira		Mixed	Testudines	1	1	2.846	0.254

Continued on next page

Table E.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
8492	Archosauria	Archosaurs	Mixed	Vertebrates	1	1	1.375	0.183
8825	Neognathae		Mixed	Birds	1	1	1.354	0.125
9126	Passeriformes	Perching birds	Mixed	Birds	1	1	1.324	0.090
42167	Acanthophiinae	Taipan subfamily	Terrestrial	Squamates	0	0	2.222	0.285
34989	Colubroidea	Colubrid snake superfamily	Mixed	Squamates	1	1	2.188	0.242
1329911	Toxicofera	Snakes and related lizards	Mixed	Squamates	1	1	2.146	0.240
7712	Tunicata	Tunicates	Salt/brackish aquatics	Tunicates	1	1	0.246	1.148
7713	Ascidiacea	Sea squirts	Salt/brackish aquatics	Tunicates	1	1	0.290	1.054
7716	Phlebobranchia		Salt/brackish aquatics	Tunicates	1	1	0.140	0.515
30274	Ascidiidae		Salt/brackish aquatics	Tunicates	1	1	0.110	0.643
7718	Ciona	Vase tunicates	Salt/brackish aquatics	Tunicates	1	1	0.200	NA
7720	Stolidobranchia	Lightbulb sea squirts	Salt/brackish aquatics	Tunicates	1	1	0.740	NA
27574	Molgula	Sea grapes	Salt/brackish aquatics	Tunicates	1	1	0.740	NA
7721	Styelidae		Salt/brackish aquatics	Tunicates	1	1	0.740	NA
7724	Styela		Salt/brackish aquatics	Tunicates	1	1	0.740	NA
7674	Echinacea		Salt/brackish aquatics	Echinoderms	1	1	0.844	0.151
6157	Platyhelminthes	Flatworms	Mixed	Flatworms	1	1	2.062	1.669
52056	Microstomum		Mixed aquatics	Flatworms	1	1	0.480	NA
166333	Acotylea		Mixed aquatics	Flatworms	1	1	2.211	1.629
39220	Dalyellioida		Mixed aquatics	Flatworms	1	1	3.413	1.449
66749	Provorticidae		Salt/brackish aquatics	Flatworms	1	1	3.413	1.449
27902	Typhloplanidae		Mixed aquatics	Flatworms	1	1	4.072	1.373
27903	Mesostominae		Freshwater aquatics	Flatworms	1	1	3.770	NA
1292248	Planarioidea		Freshwater aquatics	Flatworms	1	1	1.965	0.550
33317	Protostomia		Mixed	Invertebrates	1	1	1.926	2.015
1206794	Ecdysozoa		Mixed	Invertebrates	1	1	2.002	2.135
6231	Nematoda	Nematode worms	Mixed	Nematodes	1	1	0.169	2.033
54125	Pristionchus		Terrestrial	Nematodes	1	1	0.170	NA
6237	Caenorhabditis		Terrestrial	Nematodes	1	1	0.116	0.267
48796	Setaria		Terrestrial	Nematodes	1	1	0.130	0.218
6247	Strongyloides		Terrestrial	Nematodes	1	1	0.230	NA
33284	Tylenchoidea		Mixed	Nematodes	1	1	0.051	0.436
46005	Pratylenchinae		Mixed	Nematodes	1	1	0.020	NA
46009	Hirschmanniella		Mixed	Nematodes	0	1	0.020	NA
119088	Enoplea		Mixed aquatics	Nematodes	1	1	0.143	0.726
88770	Panarthropoda		Mixed	Invertebrates	1	1	2.072	2.094
6656	Arthropoda	Arthropods	Mixed	Arthropods	1	1	2.050	2.112
6946	Acariformes		Mixed	Mites	1	1	0.128	0.407
83137	Sarcoptiformes	Biting mites	Mixed	Mites	1	1	0.150	NA
6951	Astigmata		Mixed	Mites	1	1	0.150	NA
223454	Hemisarcoptoidea		Terrestrial	Mites	1	1	0.150	NA
223472	Psoroptidia		Terrestrial	Mites	1	1	0.150	NA
83163	Analgoidea		Terrestrial	Mites	1	1	0.150	NA
289414	Proctophyllodinae	Group of feather mites	Terrestrial	Mites	1	1	0.150	NA
289415	Proctophyllodes	Group of feather mites	Terrestrial	Mites	1	1	0.150	NA
1111262	Amerodectes	Group of feather mites	Terrestrial	Mites	0	0	0.150	NA
474258	Picalgoides	Group of feather mites	Terrestrial	Mites	1	1	0.150	NA
6947	Prostigmata	Sucking mites	Mixed	Mites	1	1	0.085	0.083
32262	Tetranychidae	Spider mites	Terrestrial	Mites	1	1	0.085	0.083
83139	Eupodina		Mixed	Mites	1	1	0.085	0.083
768130	Rhombognathus		Salt/brackish aquatics	Mites	1	1	0.085	0.083
62624	Varroa	Honey bee mites	Terrestrial	Mites	0	0	0.580	NA
6893	Araneae	Spiders	Mixed	Spiders	1	1	2.399	0.463

Continued on next page

Table E.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
6905	Araneomorphae	Cross-fanged spiders	Mixed	Spiders	1	1	2.405	0.463
81835	Linyphiidae	Money spiders	Terrestrial	Spiders	1	1	1.371	0.279
94020	RTA clade		Mixed	Spiders	1	1	2.821	0.403
94015	Dionycha		Mixed	Spiders	1	1	2.822	0.409
74980	Pardosa	Thin-legged wolf spiders	Terrestrial	Spiders	1	1	2.050	0.042
546906	Trechaleidae		Mixed	Spiders	0	0	2.405	0.463
101159	Phalangioidea		Terrestrial	Harvestmen	1	1	2.356	0.507
197563	Mandibulata		Mixed	Arthropods	1	1	2.136	2.263
71419	Polydesmida	Flat-backed millipedes	Terrestrial	Arthropods	1	1	0.410	0.224
197562	Pancrustacea		Mixed	Arthropods	1	1	2.150	2.258
6657	Crustacea	Crustaceans	Mixed	Crustaceans	1	1	5.141	1.528
2172821	Multicrustacea		Mixed	Crustaceans	1	1	6.539	1.343
116569	Neocopepoda		Mixed aquatics	Copepods	1	1	3.205	1.038
6833	Calanoida		Mixed aquatics	Copepods	1	1	4.471	0.795
6835	Calanidae		Salt/brackish aquatics	Copepods	1	1	7.719	0.461
88013	Temoridae		Mixed aquatics	Copepods	1	1	2.045	1.119
116571	Podoplea		Mixed aquatics	Copepods	1	1	0.952	0.821
84308	Cyclopoida		Mixed aquatics	Copepods	1	1	0.997	0.799
263416	Cyclops	Water fleas	Mixed aquatics	Copepods	1	1	1.201	0.875
136190	Oithona		Mixed aquatics	Copepods	1	1	0.250	NA
41212	Harpacticoida		Mixed aquatics	Copepods	1	1	0.250	NA
329895	Nannopus		Salt/brackish aquatics	Copepods	1	1	0.250	NA
6675	Cirripedia	Barnacles	Salt/brackish aquatics	Barnacles	1	1	1.296	0.441
6676	Thoracica		Salt/brackish aquatics	Barnacles	1	1	1.374	0.406
116167	Sessilia	Acorn barnacles	Salt/brackish aquatics	Barnacles	1	1	1.438	0.430
72041	Eumalacostraca		Mixed	Crustaceans	1	1	7.859	1.244
6692	Pleocyemata		Mixed	Crustaceans	1	1	6.652	0.995
516941	Inachidae		Salt/brackish aquatics	Crabs	1	1	3.368	0.312
116704	Eubrachyura		Mixed aquatics	Crabs	1	1	3.041	0.772
6694	Caridea	Caridean shrimps	Mixed aquatics	Crustaceans	1	1	12.335	0.703
98081	Crangonidae		Salt/brackish aquatics	Crustaceans	1	1	18.116	0.721
6701	Pandalidae	Pandalid shrimps	Salt/brackish aquatics	Crustaceans	1	1	8.530	NA
6821	Amphipoda	Amphipods	Mixed	Crustaceans	1	1	6.804	1.469
44327	Gammaridea		Mixed aquatics	Crustaceans	1	1	28.830	0.826
111560	Niphargus		Freshwater aquatics	Crustaceans	1	1	5.690	0.521
44329	Gammaroidea		Mixed aquatics	Crustaceans	1	1	5.690	0.521
96850	Jaera		Salt/brackish aquatics	Crustaceans	1	1	1.900	NA
6670	Ostracoda	Ostracods	Mixed	Ostracods	1	1	1.104	0.580
182511	Cytheroidea		Salt/brackish aquatics	Ostracods	1	1	0.700	NA
6960	Hexapoda	Insects and relatives	Mixed	Arthropods	1	1	1.102	1.857
30001	Collembola	Springtails	Mixed	Springtails	1	1	0.230	NA
7496	Pterygota	Winged insects	Mixed	Insects	1	1	1.089	1.889
33340	Neoptera		Mixed	Insects	1	1	1.044	1.986
33392	Holometabola	Metamorphosing insects	Mixed	Insects	1	1	0.506	0.799
85604	Amphiesmenoptera		Mixed	Insects	1	1	0.632	0.498
7088	Lepidoptera	Butterflies and moths	Mixed	Lepidopterans	1	1	0.621	0.465
889681	Erikssonia	Eriksson's coppers	Terrestrial	Butterflies	0	0	0.230	NA
344716	Dircennina		Terrestrial	Butterflies	0	0	0.290	NA
42282	Satyrinae	Browns	Terrestrial	Butterflies	1	1	0.490	NA
40083	Phycitinae		Terrestrial	Moths	1	1	0.730	NA
42265	Riodininae		Terrestrial	Butterflies	0	0	0.358	0.266
37572	Papilionoidea	Butterflies	Terrestrial	Butterflies	1	1	0.358	0.266
33415	Nymphalidae	Brushfoots	Terrestrial	Butterflies	1	1	0.388	0.206

Continued on next page

Table E.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
134396	Sesamia		Terrestrial	Moths	0	1	0.975	0.007
104431	Obtectomera		Mixed	Lepidopterans	1	1	0.639	0.499
41197	Heteroneura		Mixed	Lepidopterans	1	1	0.632	0.498
37567	Ditrysia		Mixed	Lepidopterans	1	1	0.640	0.495
30225	Arctiinae		Mixed	Moths	1	1	0.697	0.349
30263	Trichoptera	Caddisflies	Freshwater aquatics	Trichopterans	1	1	0.632	0.498
600663	Polyplectropus		Freshwater aquatics	Trichopterans	0	0	0.632	0.498
93873	Annulipalpia	Fixed-retreat makers	Freshwater aquatics	Trichopterans	1	1	0.632	0.498
177955	Ecnomidae		Freshwater aquatics	Trichopterans	1	1	0.632	0.498
1683728	Plenitentoria		Freshwater aquatics	Trichopterans	1	1	0.632	0.498
41033	Limnephiloidea		Freshwater aquatics	Trichopterans	1	1	0.632	0.498
177653	Apataniidae		Freshwater aquatics	Trichopterans	1	1	0.632	0.498
7041	Coleoptera	Beetles	Mixed	Coleopterans	1	1	0.756	0.772
41071	Adephaga		Mixed	Coleopterans	1	1	0.832	0.790
535382	Caraboidea	Ground beetle superfamily	Terrestrial	Coleopterans	1	1	0.545	0.552
41073	Carabidae	Ground beetles	Terrestrial	Coleopterans	1	1	0.574	0.513
795056	Elaphropus		Terrestrial	Coleopterans	1	1	0.574	0.513
932593	Exocelina		Freshwater aquatics	Coleopterans	0	0	1.318	0.575
41084	Polyphaga		Mixed	Coleopterans	1	1	0.764	0.697
63712	Doryphorini		Terrestrial	Coleopterans	1	1	0.460	NA
41088	Cucujiformia		Terrestrial	Coleopterans	1	1	0.710	0.722
484807	Aphanocephalus		Terrestrial	Coleopterans	0	0	0.606	0.614
41102	Hydrophilidae	Water scavenger beetle family	Freshwater aquatics	Coleopterans	1	1	1.294	0.302
7147	Diptera	True flies	Mixed	Dipterans	1	1	0.375	0.861
7203	Brachycera		Mixed	Dipterans	1	1	0.312	0.708
43733	Muscomorpha		Mixed	Dipterans	1	1	0.309	0.695
43738	Schizophora		Terrestrial	Dipterans	1	1	0.302	0.701
1713957	Macrostomus		Terrestrial	Dipterans	0	0	0.309	0.695
92557	Empididae	Dagger flies	Terrestrial	Dipterans	1	1	0.309	0.695
43741	Acalyptratae		Terrestrial	Dipterans	1	1	0.230	0.362
7215	Drosophila	Fruit flies	Terrestrial	Dipterans	1	1	0.216	0.259
32281	Drosophila		Terrestrial	Dipterans	1	1	0.216	0.259
32357	obscura subgroup		Terrestrial	Dipterans	1	1	0.194	0.221
32346	melanogaster group		Terrestrial	Dipterans	1	1	0.202	0.214
32321	repleta group		Terrestrial	Dipterans	1	1	0.194	0.206
32351	melanogaster subgroup		Terrestrial	Dipterans	1	1	0.169	0.190
198037	mojavensis species complex		Terrestrial	Dipterans	0	0	0.183	0.175
32352	montium subgroup		Terrestrial	Dipterans	0	0	0.228	0.129
59854	Scioniini	Austral horse flies	Mixed	Dipterans	0	0	0.312	0.708
7148	Nematocera	Midge suborder	Mixed	Dipterans	1	1	0.435	0.920
43793	Cecidomyiinae		Freshwater aquatics	Dipterans	1	1	0.123	0.285
7190	Simuliidae	Blackflies	Freshwater aquatics	Dipterans	1	1	0.190	NA
7149	Chironomidae	Non-biting midges	Freshwater aquatics	Dipterans	1	1	0.141	0.321
7150	Chironomus		Freshwater aquatics	Dipterans	1	1	0.181	0.210
41042	Tipulidae	Crane flies and relatives	Mixed	Dipterans	1	1	0.435	0.920
7399	Hymenoptera	Sawflies, wasps, ants and bees	Mixed	Hymenopterans	1	1	0.428	0.527
7400	Apocrita	Wasps, ants and bees	Mixed	Hymenopterans	1	1	0.413	0.495
7434	Aculeata	Stinging wasps, ants and bees	Mixed	Hymenopterans	1	1	0.429	0.465
156305	Chelostoma		Terrestrial	Hymenopterans	1	1	0.565	0.663
7458	Apidae	Honey bee family	Terrestrial	Hymenopterans	1	1	0.549	0.491
7459	Apis	Honey bees	Terrestrial	Hymenopterans	1	1	0.229	0.279
34695	Myrmicinae		Terrestrial	Hymenopterans	1	1	0.359	0.369
36668	Formicidae	Ants	Terrestrial	Hymenopterans	1	1	0.361	0.353

Continued on next page

Table E.1 continued

ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
143999	Attini	Fungus-growing ants	Terrestrial	Hymenopterans	0	0	0.363	0.339
7479	Formicinae		Terrestrial	Hymenopterans	1	1	0.292	0.192
1955251	Parasitoida		Terrestrial	Hymenopterans	1	1	0.342	0.639
7422	Chalcidoidea	Chalcid wasps	Terrestrial	Hymenopterans	1	1	0.409	0.407
33373	Sternorrhyncha	Aphids and relatives	Terrestrial	Hemipterans	1	1	0.518	0.293
58612	Grylloblatta	Ice crawlers	Terrestrial	Insects	0	0	5.490	0.787
327033	Echinoderidae		Salt/brackish aquatics	Kinorhynchs	1	1	0.560	NA
1206795	Lophotrochozoa	Lophotrochozoans	Mixed	Invertebrates	1	1	1.611	0.817
6340	Annelida	Annelid worms	Mixed	Annelids	1	1	1.298	0.967
2109251	Naididae	Detritus worms	Mixed aquatics	Annelids	1	1	2.072	0.864
6341	Polychaeta	Bristle worms	Mixed aquatics	Annelids	1	1	1.360	0.913
6362	Sabellida	Tube worms and relatives	Salt/brackish aquatics	Annelids	1	1	1.112	0.636
1002757	Fabriciidae		Salt/brackish aquatics	Annelids	1	1	1.112	0.636
1002765	Fabriciola		Salt/brackish aquatics	Annelids	1	1	1.112	0.636
184866	Manayunkia		Salt/brackish aquatics	Annelids	1	1	1.112	0.636
36125	Sabellariidae	Honeycomb worm family	Salt/brackish aquatics	Annelids	1	1	1.112	0.636
205110	Protodriloides		Salt/brackish aquatics	Annelids	1	1	0.240	NA
42115	Arenicolidae	Lugworms	Salt/brackish aquatics	Annelids	1	1	0.900	NA
46603	Orbiniidae		Salt/brackish aquatics	Annelids	1	1	1.977	0.665
46604	Scoloplos		Salt/brackish aquatics	Annelids	1	1	1.977	0.665
46590	Cirratulidae		Salt/brackish aquatics	Annelids	1	1	1.300	1.115
46599	Spionidae		Salt/brackish aquatics	Annelids	1	1	0.500	NA
6373	Terebellida		Salt/brackish aquatics	Annelids	1	1	0.989	0.314
10206	Gymnolaemata		Salt/brackish aquatics	Bryozoans	1	1	0.700	0.565
10207	Cheilostomatida		Salt/brackish aquatics	Bryozoans	1	1	0.817	0.805
558764	Flustrina		Salt/brackish aquatics	Bryozoans	1	1	0.817	0.805
558780	Adeonidae		Salt/brackish aquatics	Bryozoans	1	1	0.817	0.805
51985	Ctenostomatida		Mixed aquatics	Bryozoans	1	1	0.550	NA
97264	Stenolaemata		Salt/brackish aquatics	Bryozoans	1	1	0.250	NA
41321	Paucitubulatina		Mixed aquatics	Gastrotrichs	1	1	0.230	0.565
41372	Chaetonotidae		Mixed aquatics	Gastrotrichs	1	1	0.080	NA
6447	Mollusca	Molluscs	Mixed	Molluscs	1	1	2.053	0.618
6544	Bivalvia	Bivalves	Mixed aquatics	Bivalves	1	1	1.747	0.405
735337	Euheterodonta		Mixed aquatics	Bivalves	1	1	1.719	0.268
6601	Myoidea		Salt/brackish aquatics	Bivalves	1	1	1.400	0.000
6580	Veneroida		Mixed aquatics	Bivalves	1	1	1.697	0.247
105710	Cardioidea	Cockle superfamily	Salt/brackish aquatics	Bivalves	1	1	1.335	0.037
55708	Cardiidae	Cockles	Salt/brackish aquatics	Bivalves	1	1	1.518	0.214
55709	Cerastoderma		Salt/brackish aquatics	Bivalves	1	1	1.370	NA
98297	Tellinoidea		Salt/brackish aquatics	Bivalves	1	1	1.970	0.274
121180	Abra		Salt/brackish aquatics	Bivalves	1	1	1.970	0.274
856330	Nuculanoida		Salt/brackish aquatics	Bivalves	1	1	3.533	0.479
6547	Mytilidae	Mussels	Mixed aquatics	Bivalves	1	1	1.787	0.290
301959	Mytilinae		Mixed aquatics	Bivalves	1	1	1.565	0.218
6548	Mytilus		Salt/brackish aquatics	Bivalves	1	1	1.645	0.130
38948	Perna		Salt/brackish aquatics	Bivalves	1	1	1.110	NA
98302	Ostreoidea	Oyster superfamily	Salt/brackish aquatics	Bivalves	1	1	0.944	0.304
37858	Ostrea		Salt/brackish aquatics	Bivalves	1	1	1.235	0.074
6643	Octopus		Salt/brackish aquatics	Cephalopods	1	1	4.127	0.271
6448	Gastropoda	Snails and slugs	Mixed	Gastropods	1	1	2.309	0.646
69555	Caenogastropoda		Mixed	Gastropods	1	1	2.144	0.507
216294	Littorinimorpha		Mixed	Gastropods	1	1	1.174	0.170
1345656	Truncatelloidea		Mixed	Gastropods	1	1	1.250	NA

Continued on next page

Table E.1 continued

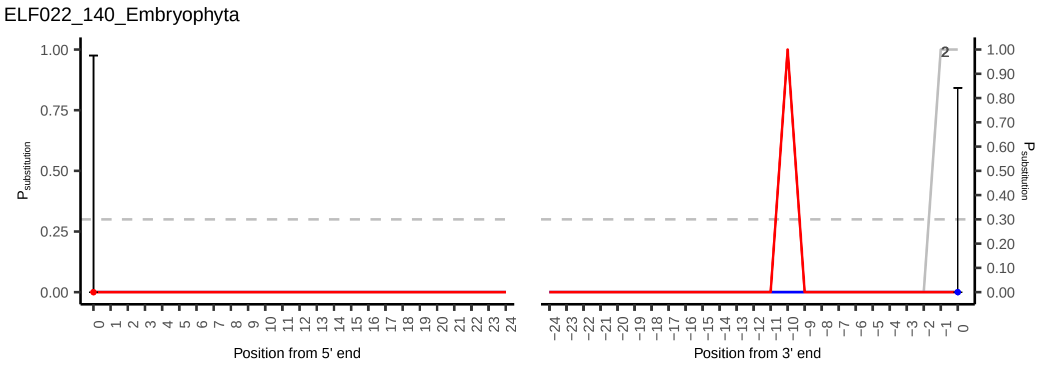
ID	Taxon	Common name	Ecological category	Group	GB native	European native	Mean C-value	CV
55731	Hydrobiidae	Mud snails	Mixed aquatics	Gastropods	1	1	0.680	NA
216305	Heterobranchia		Mixed	Gastropods	1	1	1.736	0.392
71469	Doto		Salt/brackish aquatics	Gastropods	1	1	0.955	0.213
6651	Neoloricata		Salt/brackish aquatics	Polyplacophorans	1	1	1.572	0.406
6652	Chitonida		Salt/brackish aquatics	Polyplacophorans	1	1	1.393	0.330
6222	Lineidae		Salt/brackish aquatics	Nemerteans	1	1	1.400	NA
10194	Brachionus		Mixed aquatics	Rotifers	1	1	0.141	0.589
6073	Cnidaria	Cnidarians	Mixed aquatics	Cnidarians	1	1	0.857	0.833
6102	Hexacorallia	Hexacorals	Salt/brackish aquatics	Cnidarians	1	1	0.618	0.464
6103	Actiniaria	Sea anemones	Salt/brackish aquatics	Sea anemones	1	1	0.461	0.473
42822	Actiniidae		Salt/brackish aquatics	Sea anemones	1	1	0.598	0.316
45620	Urticina		Salt/brackish aquatics	Sea anemones	1	1	0.598	0.316
6115	Metridium	Plumose anemones	Salt/brackish aquatics	Sea anemones	1	1	0.461	0.473
227527	Sagartiidae		Salt/brackish aquatics	Sea anemones	1	1	0.461	0.473
6127	Acropora		Salt/brackish aquatics	Stony corals	0	0	0.430	NA
6132	Octocorallia	Octocorals	Salt/brackish aquatics	Cnidarians	1	1	0.640	NA
6074	Hydrozoa	Hydrozoans	Mixed aquatics	Hydrozoans	1	1	1.216	0.768
37516	Hydroidolina		Mixed aquatics	Hydrozoans	1	1	1.197	0.829
406427	Anthoathecata		Mixed aquatics	Hydrozoans	1	1	0.954	0.519
6077	Corynidae		Salt/brackish aquatics	Hydrozoans	1	1	0.700	NA
6078	Sarsia		Salt/brackish aquatics	Hydrozoans	1	1	0.700	NA
406428	Filifera		Mixed aquatics	Hydrozoans	1	1	0.570	0.447
168713	Bougainvilliidae		Salt/brackish aquatics	Hydrozoans	1	1	0.750	NA
168714	Bougainvillia		Salt/brackish aquatics	Hydrozoans	1	1	0.750	NA
500008	Leptothecata	Thecate hydroids	Mixed aquatics	Hydrozoans	1	1	0.410	0.448
27800	Campanulariidae		Mixed aquatics	Hydrozoans	1	1	0.410	0.448
32569	Obelia	Sea fur	Mixed aquatics	Hydrozoans	1	1	0.410	0.448
308586	Haleciidae		Salt/brackish aquatics	Hydrozoans	1	1	0.410	0.448
36081	Mitrocomidae		Salt/brackish aquatics	Hydrozoans	1	1	0.410	0.448
308572	Plumulariidae		Salt/brackish aquatics	Hydrozoans	1	1	0.410	0.448
308618	Nemertesia		Salt/brackish aquatics	Hydrozoans	1	1	0.410	0.448
86619	Sertulariidae		Salt/brackish aquatics	Hydrozoans	1	1	0.410	0.448
6142	Scyphozoa	True jellyfish	Salt/brackish aquatics	Cnidarians	1	1	0.476	0.424
140493	Tentaculata		Salt/brackish aquatics	Ctenophores	1	1	1.735	1.162
1779146	Heteroscleromorpha		Mixed aquatics	Poriferans	1	1	0.253	1.151
1779161	Spongillida	Freshwater sponges	Freshwater aquatics	Poriferans	1	1	0.355	0.086
1779162	Suberitida		Salt/brackish aquatics	Poriferans	1	1	0.587	1.496
6061	Halichondria		Salt/brackish aquatics	Poriferans	1	1	0.120	0.118

Appendix F

MetaDamage plots per sample for Embryophyta

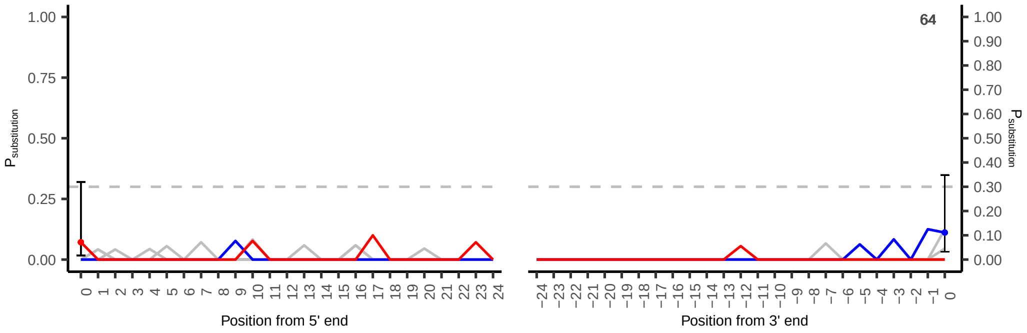
MetaDamage profiles for Embryophyta reads in each sample, grouped by core and ordered by ascending depth. $P_{substitution}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right. Plots for samples with no data are omitted.

F.0.1 ELF022

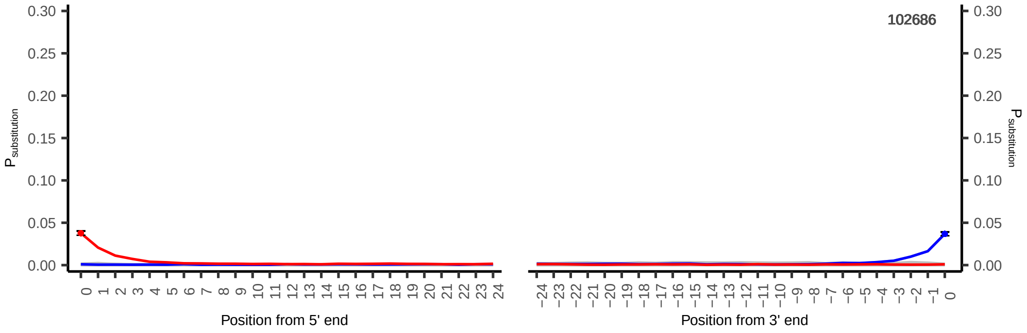


F.1 ELF027

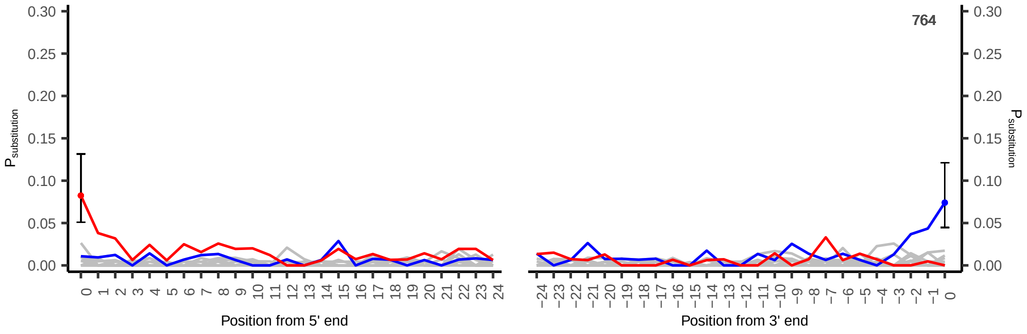
ELF027_177_Embryophyta



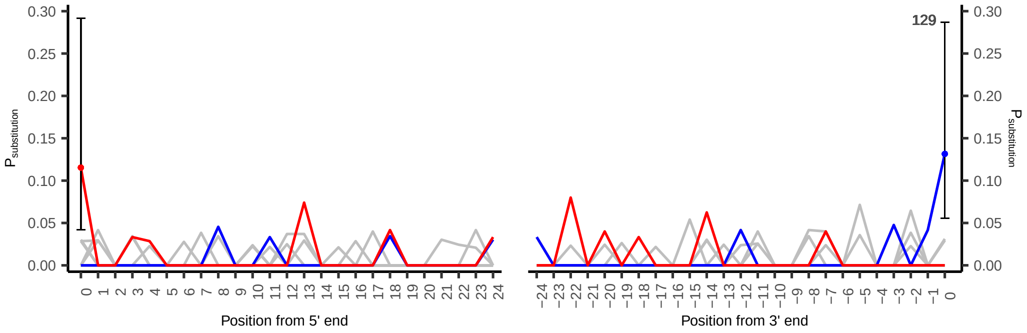
ELF027_174_Embryophyta

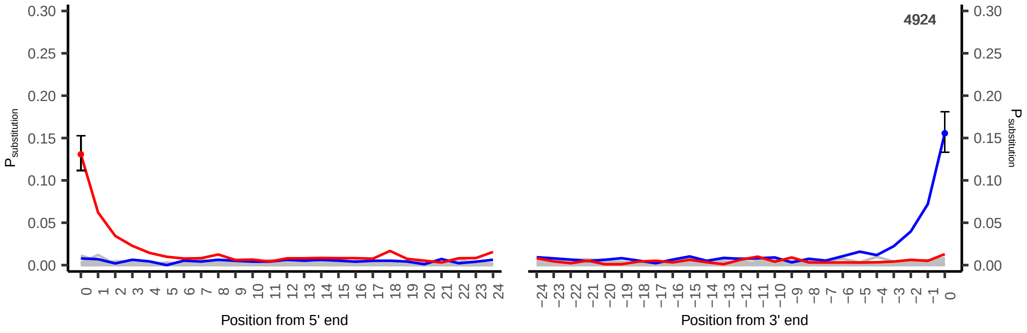


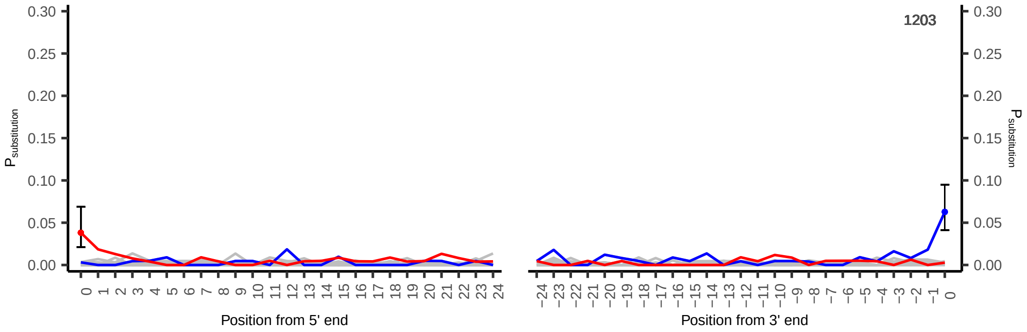
ELF027_160_Embryophyta

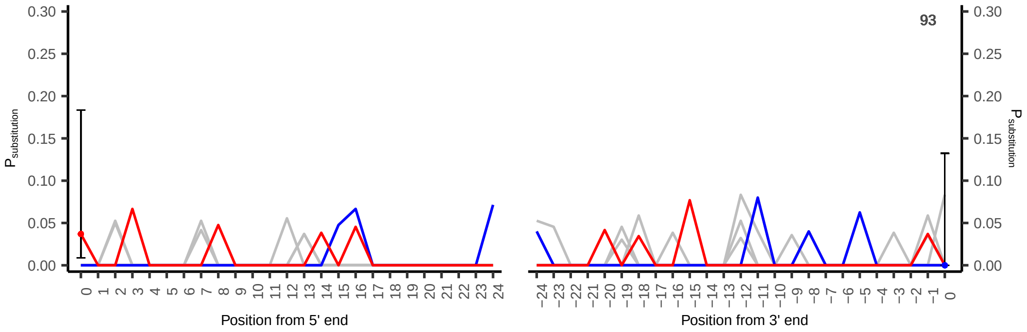


F.2 ELF059

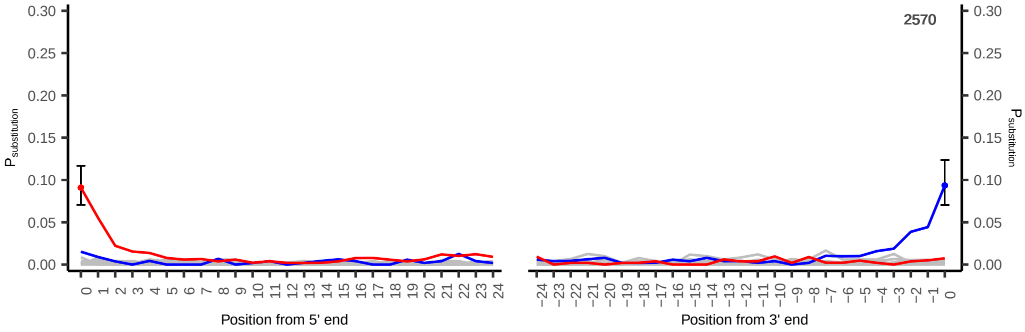

ELF059_378_Embryophyta


ELF059_359_Embryophyta

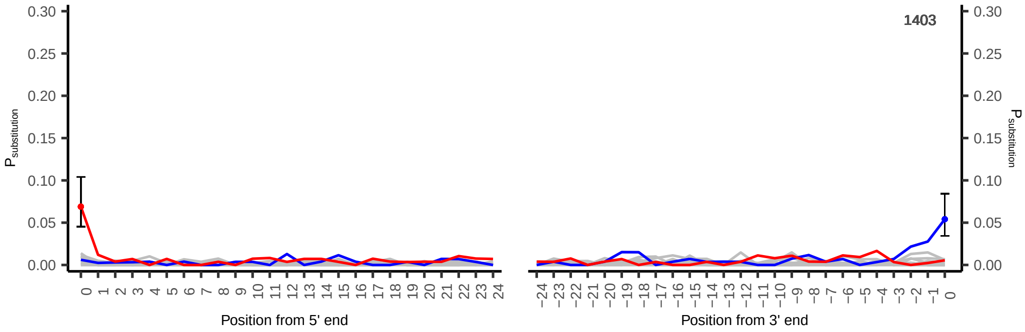

ELF059_337_Embryophyta


ELF059_280_Embryophyta

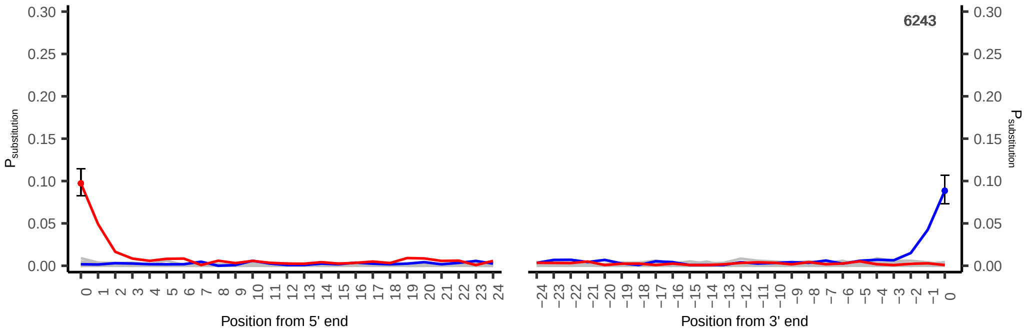
ELF059_270_Embryophyta



ELF059_230_Embryophyta

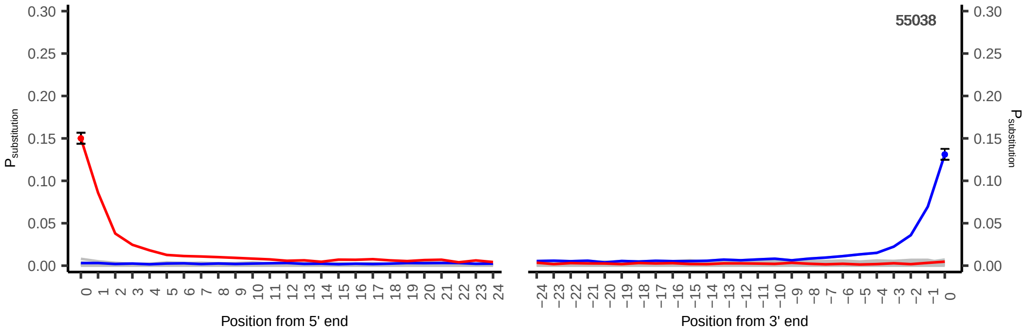


ELF059_210_Embryophyta

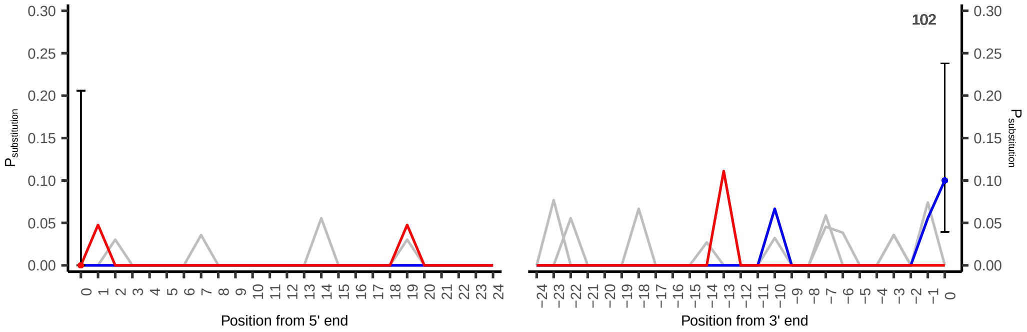


F.3 ELF059A

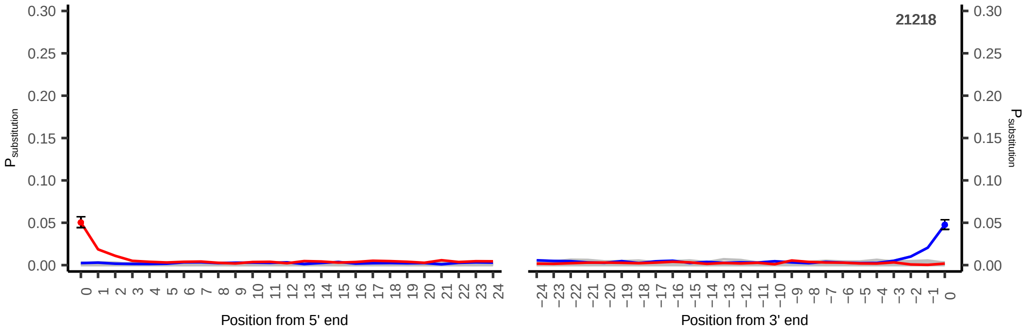
ELF059A_355_Embryophyta



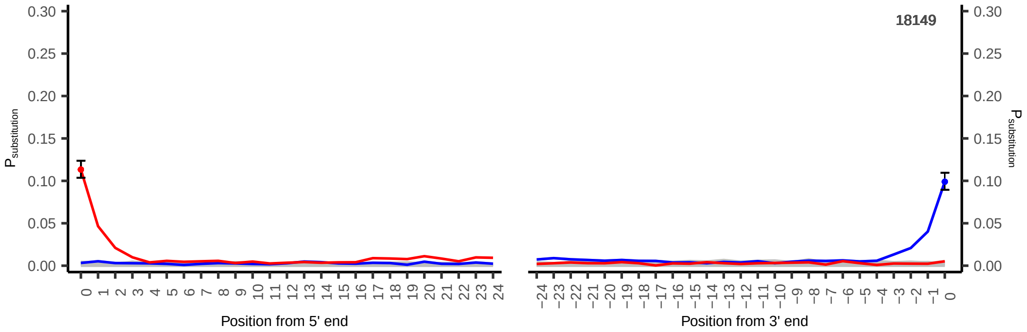
ELF059A_320_Embryophyta



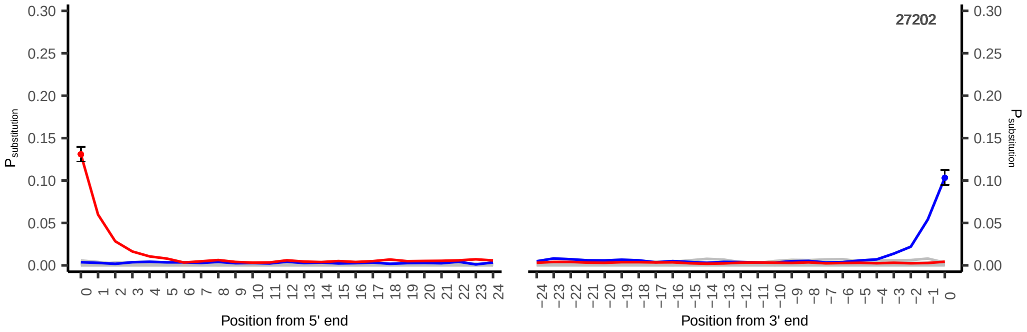
ELF059A_250_Embryophyta



ELF059A_190_Embryophyta

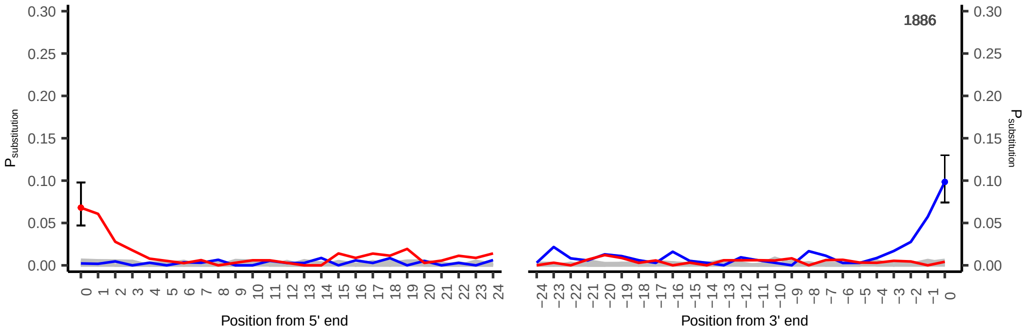


ELF059A_135_Embryophyta

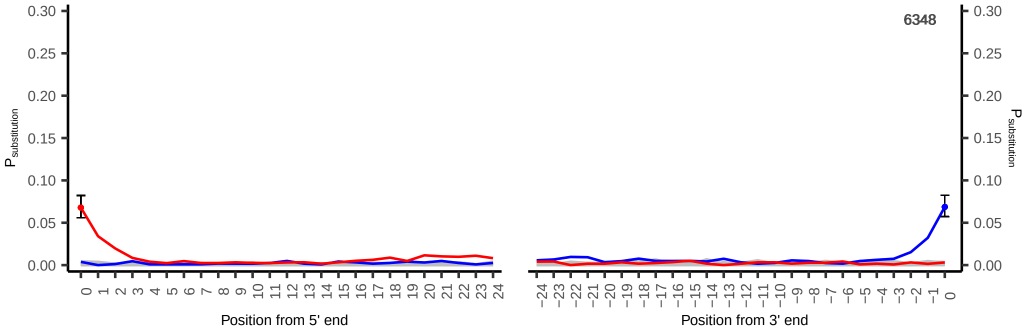


F.4 ELF060

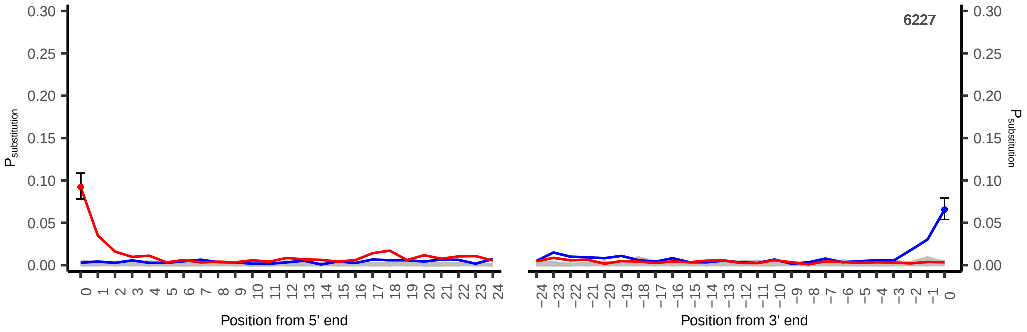
ELF060_465_Embryophyta



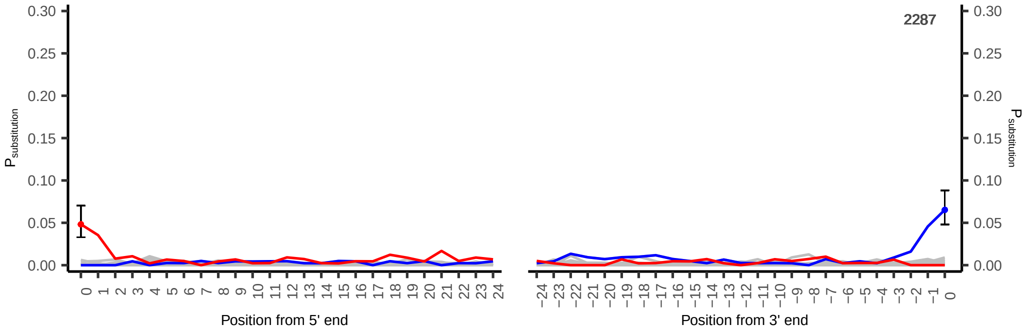
ELF060_420_Embryophyta



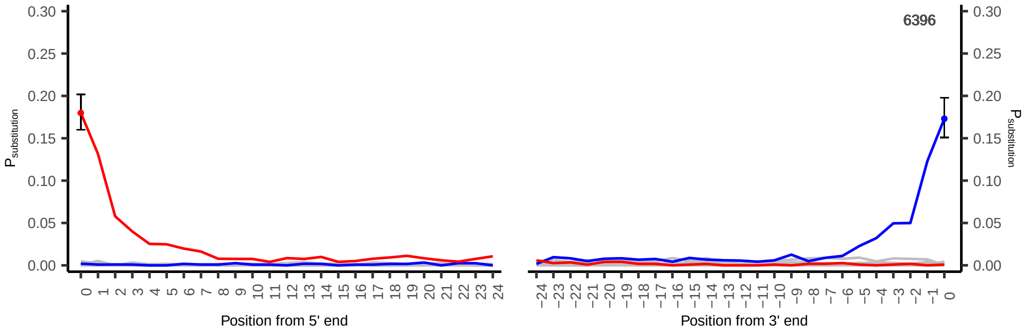
ELF060_350_Embryophyta

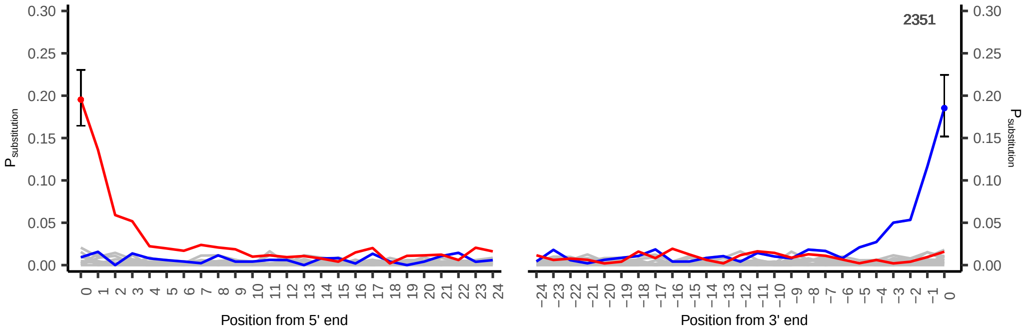


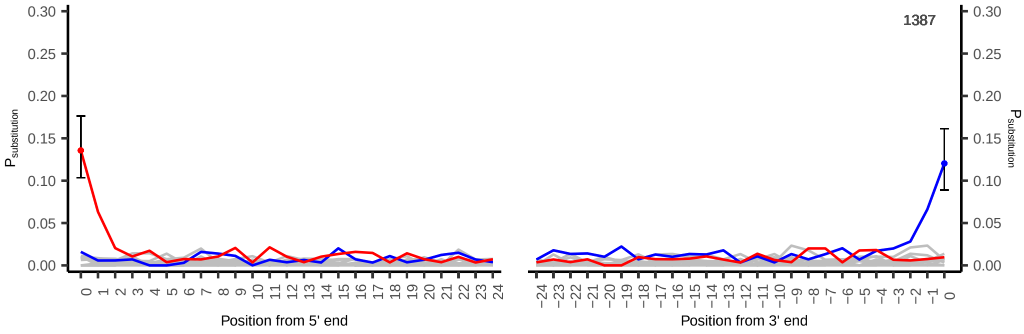
ELF060_250_Embryophyta

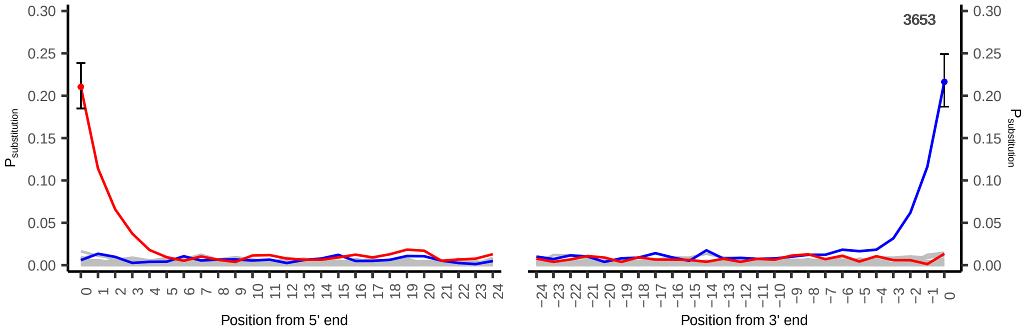


F.5 ELF032A

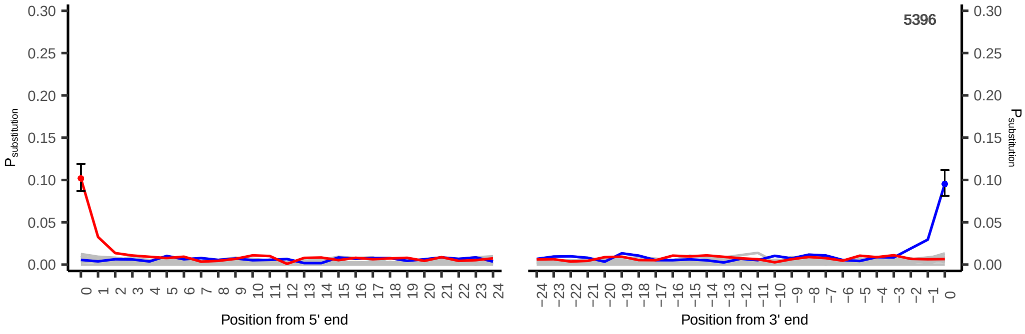

ELF032A_177_Embryophyta


ELF032A_153_Embryophyta

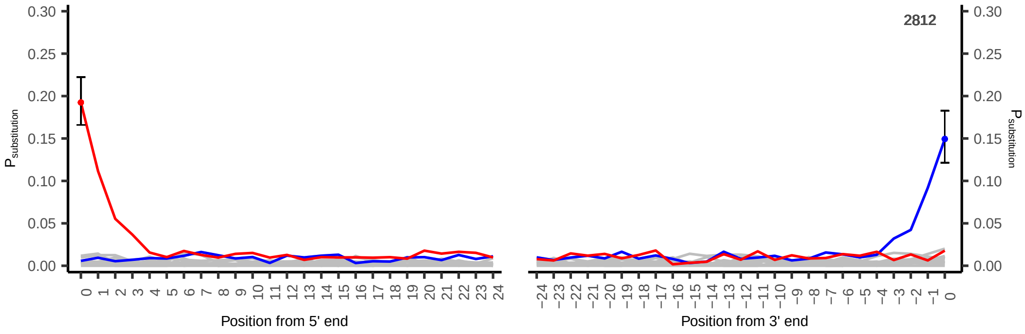

ELF032A_117_Embryophyta


ELF032A_095_Embryophyta

ELF032A_074_Embryophyta

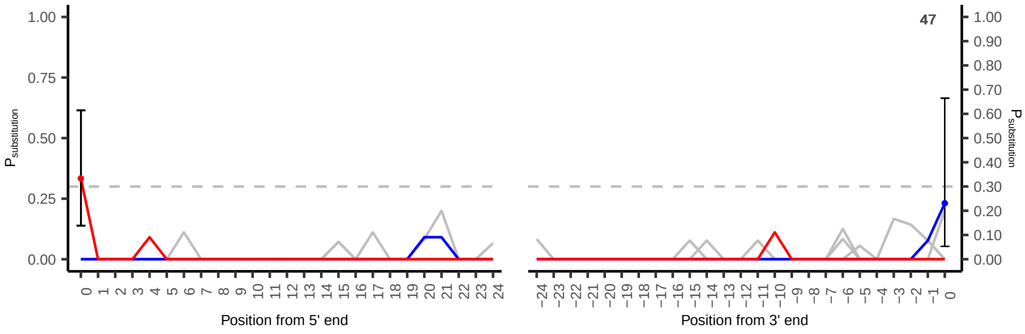


ELF032A_047_Embryophyta

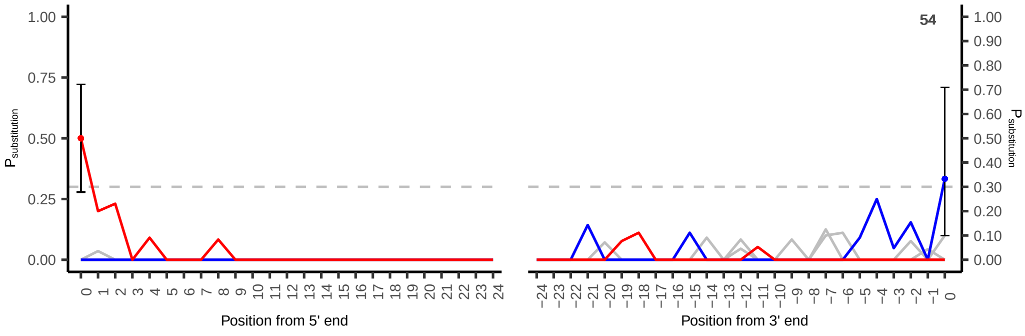


F.6 ELF033

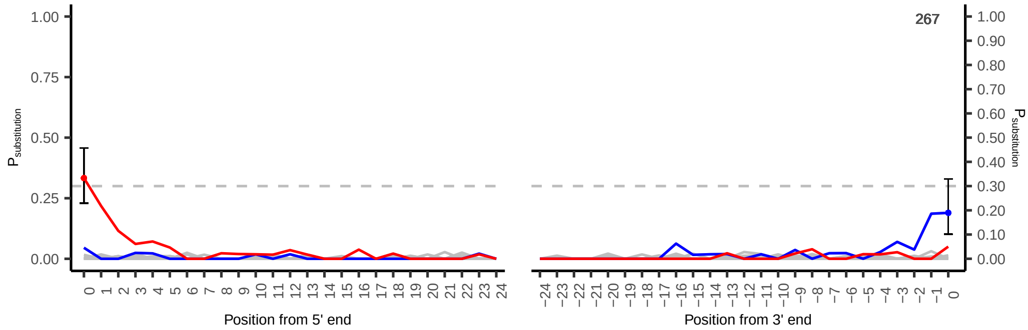
ELF033_203_Embryophyta



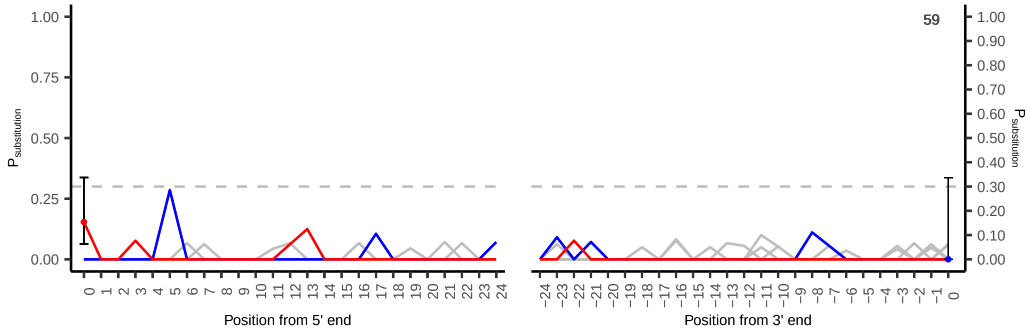
ELF033_195_Embryophyta



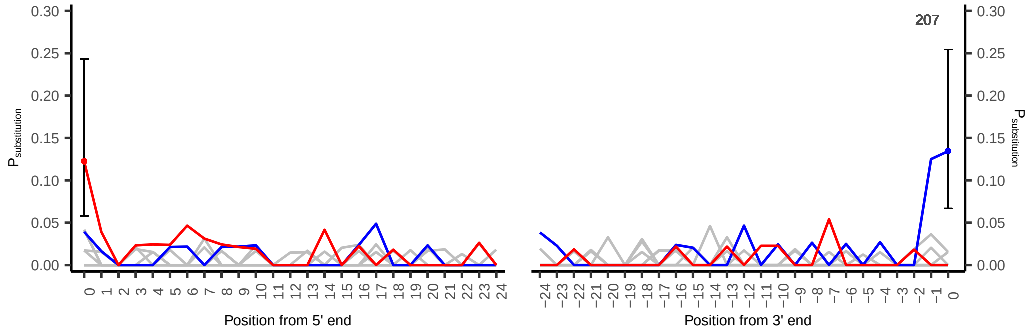
ELF033_187_Embryophyta



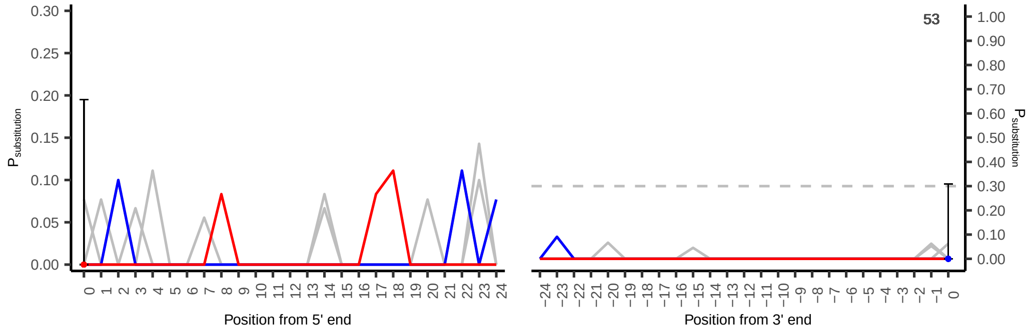
ELF033_183_Embryophyta

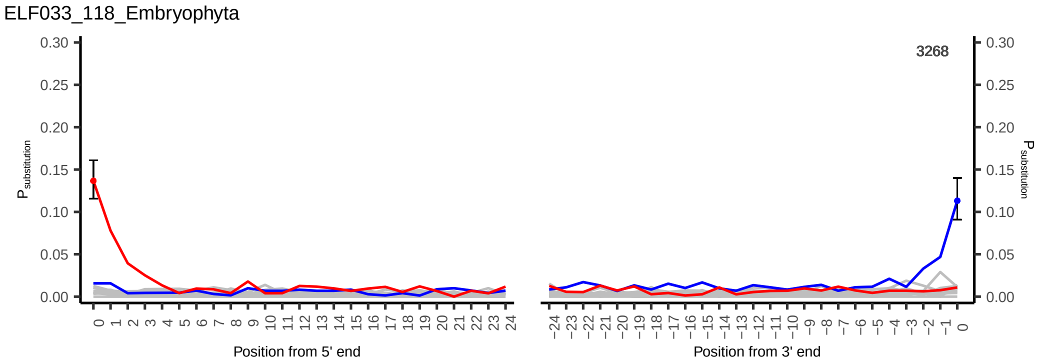


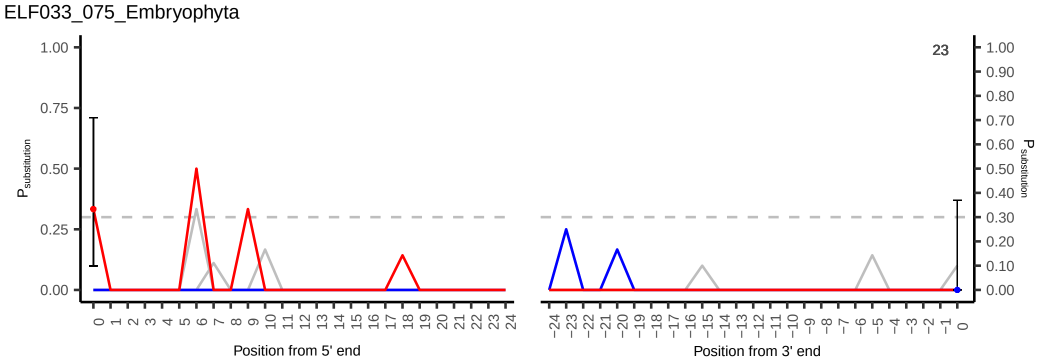
ELF033_176_Embryophyta

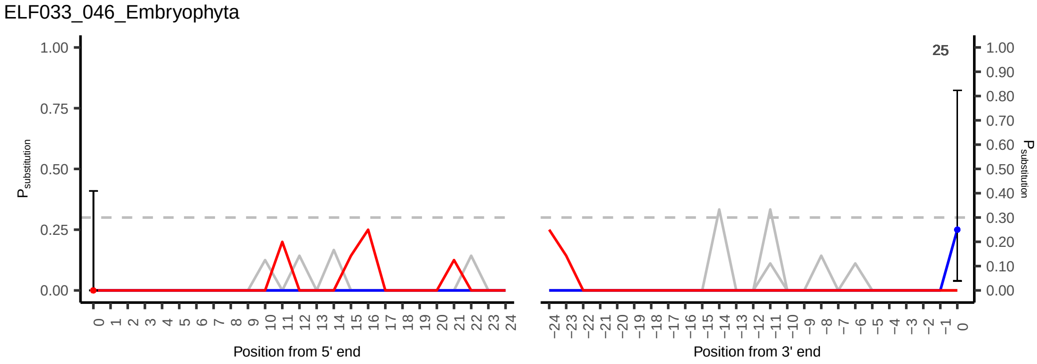


ELF033_155_Embryophyta



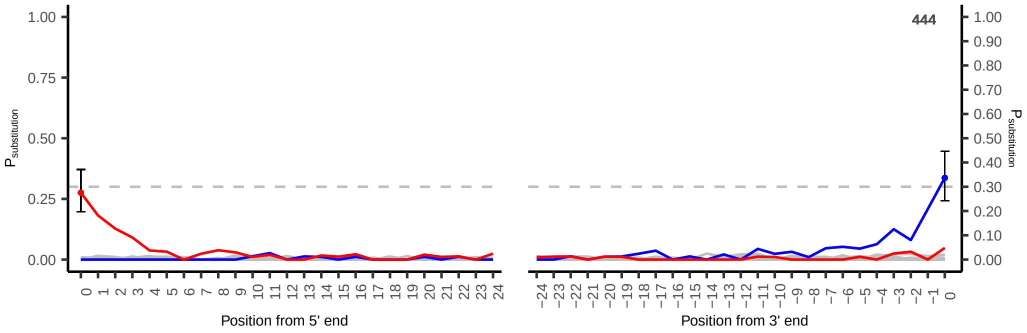

ELF033_118_Embryophyta
3268
P_substitution
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end


ELF033_075_Embryophyta
23
P_substitution
0.00
0.25
0.50
0.75
1.00
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

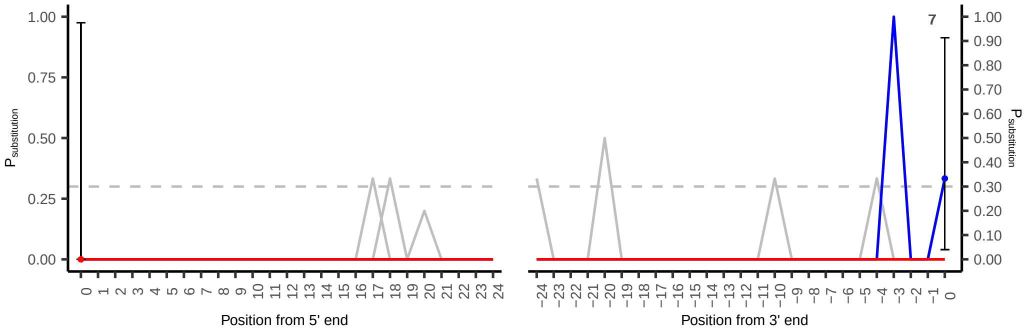

ELF033_046_Embryophyta
25
P_substitution
0.00
0.25
0.50
0.75
1.00
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

F.7 ELF033A

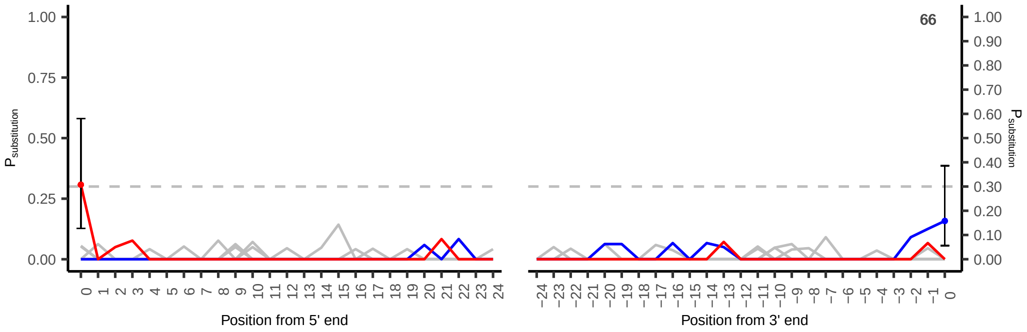
ELF033A_158_Embryophyta



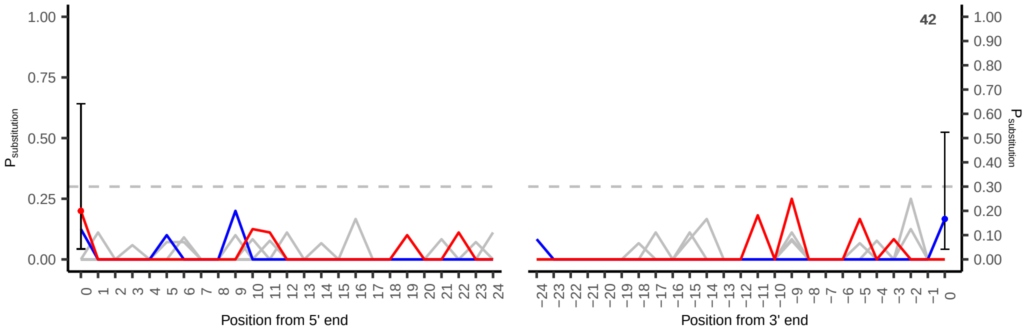
ELF033A_126_Embryophyta



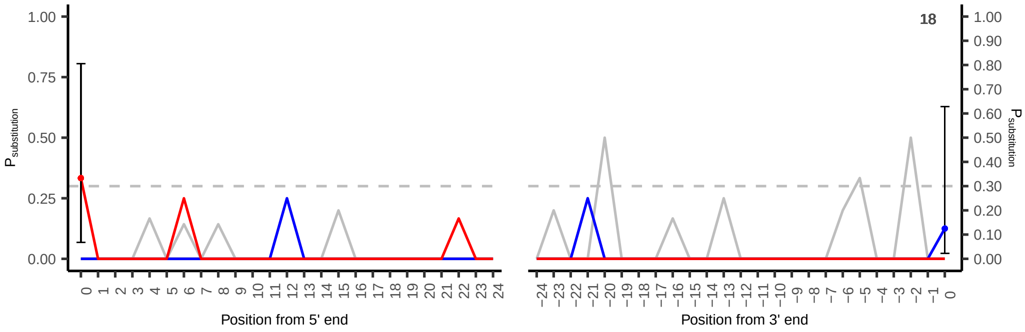
ELF033A_111_Embryophyta



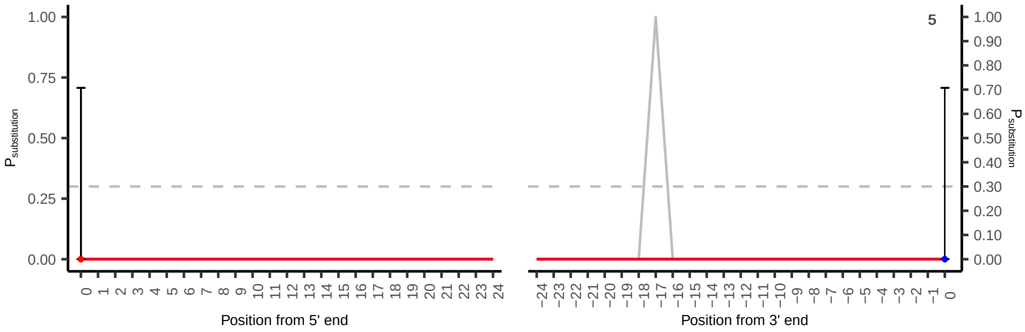
ELF033A_098_Embryophyta



ELF033A_070_Embryophyta



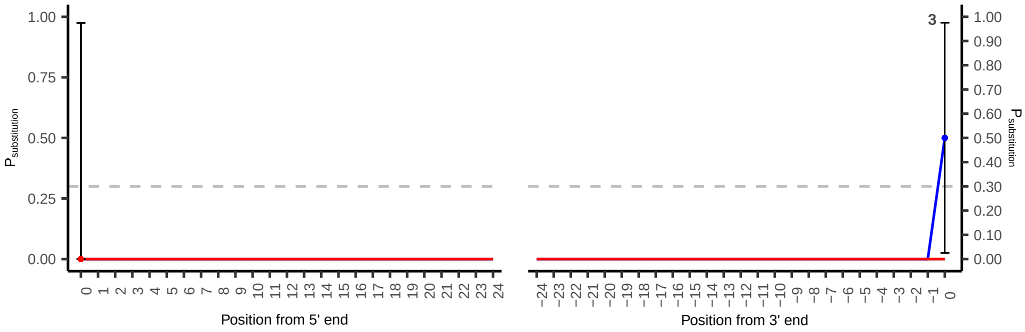
ELF033A_050_Embryophyta



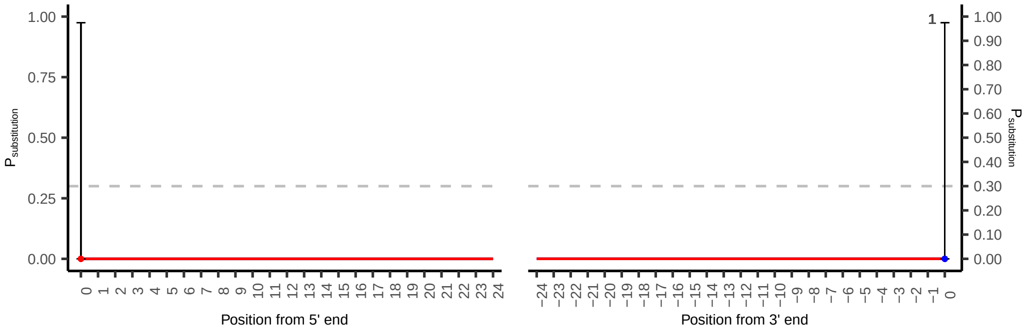
F.8 ELF034A

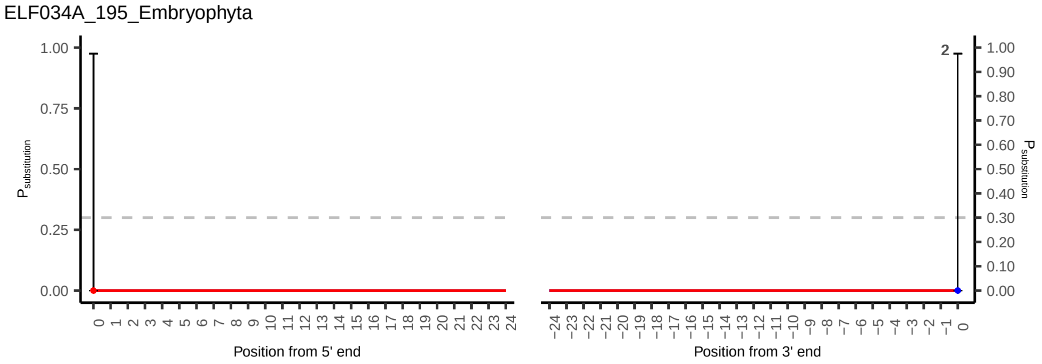
No data for ELF034A_225.

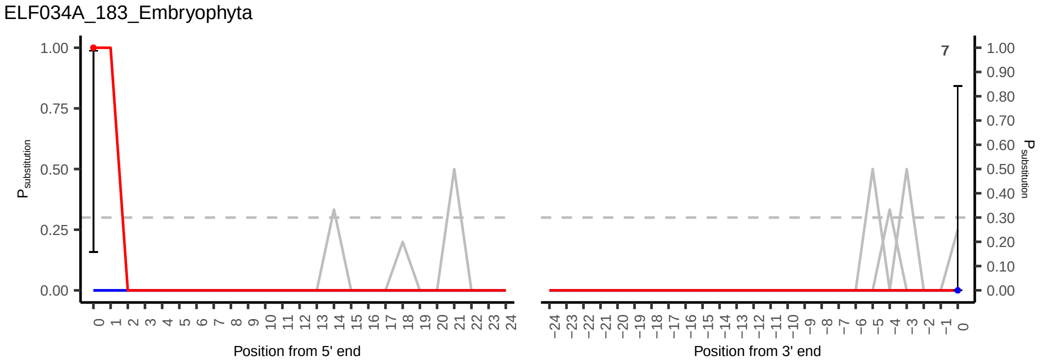
ELF034A_282_Embryophyta

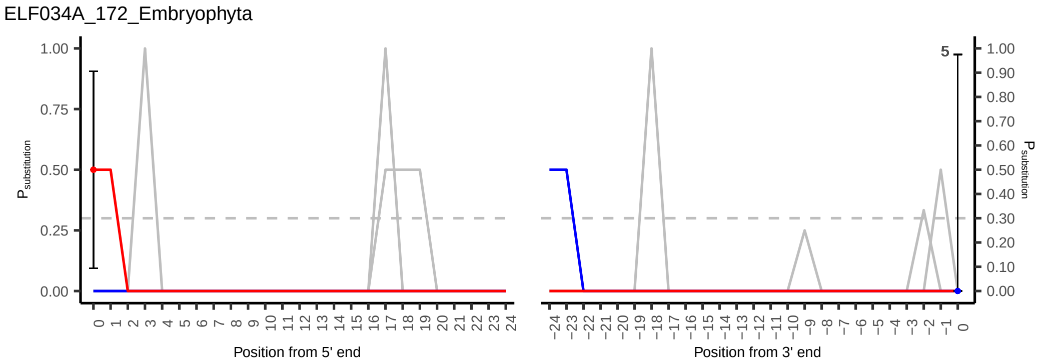


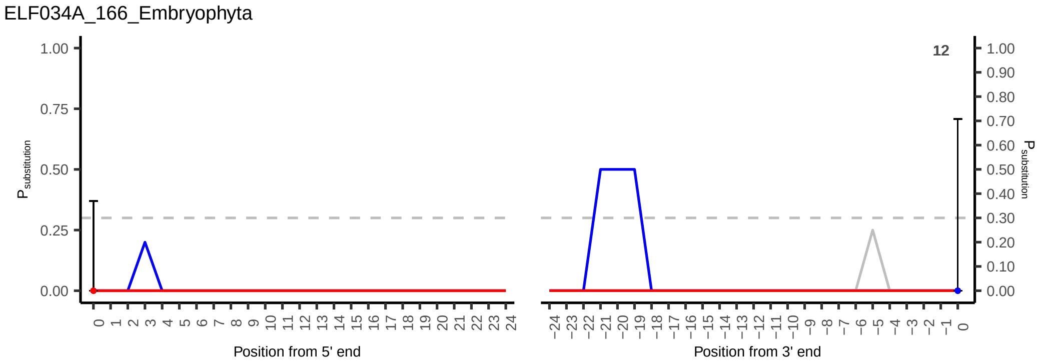
ELF034A_261_Embryophyta

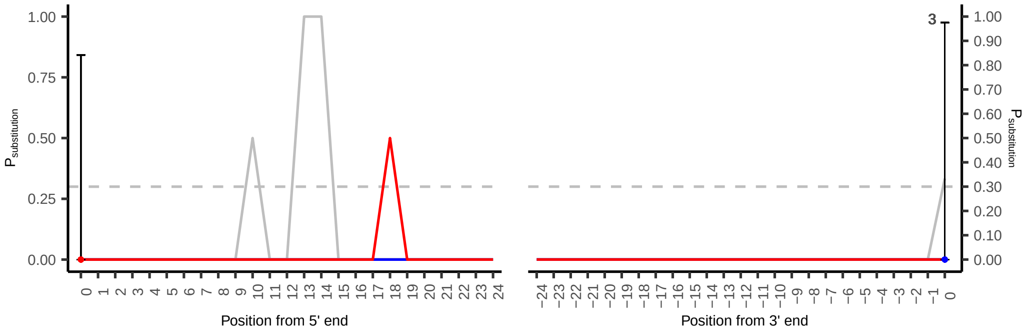


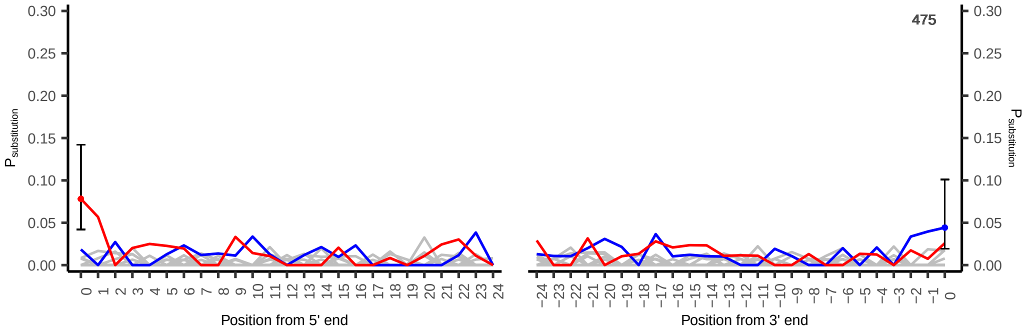

ELF034A_195_Embryophyta

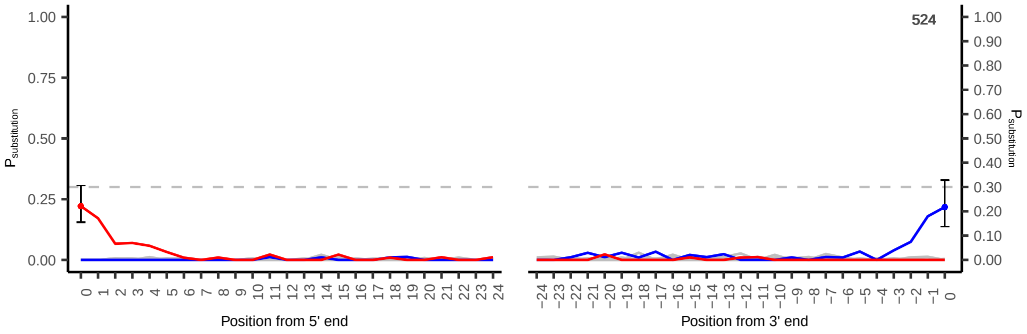

ELF034A_183_Embryophyta

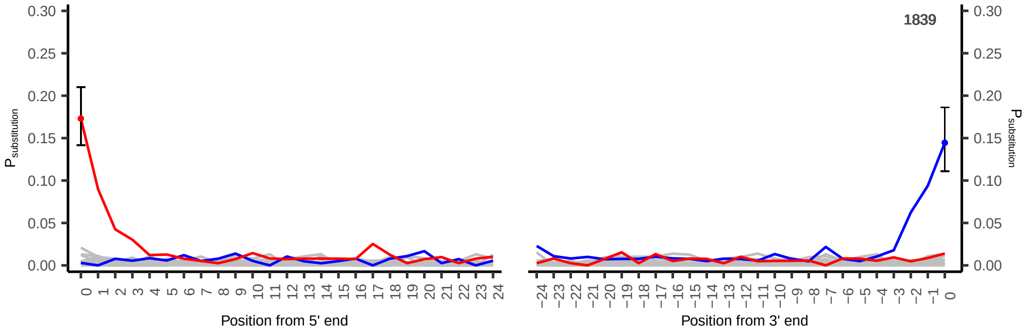

ELF034A_172_Embryophyta


ELF034A_166_Embryophyta


ELF034A_146_Embryophyta


ELF034A_126_Embryophyta

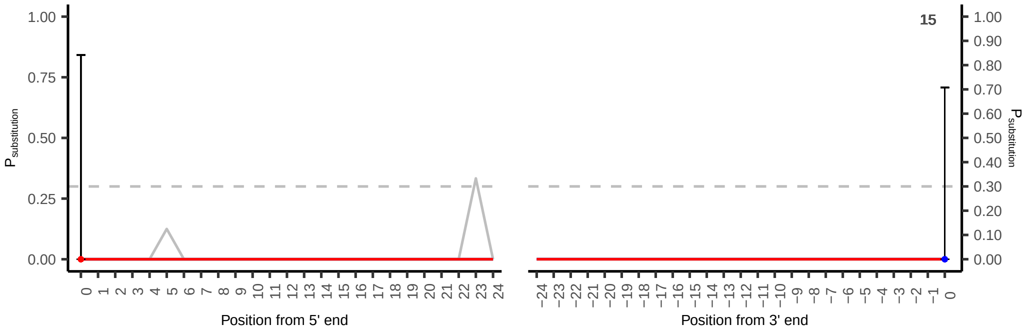

ELF034A_081_Embryophyta


ELF034A_063_Embryophyta

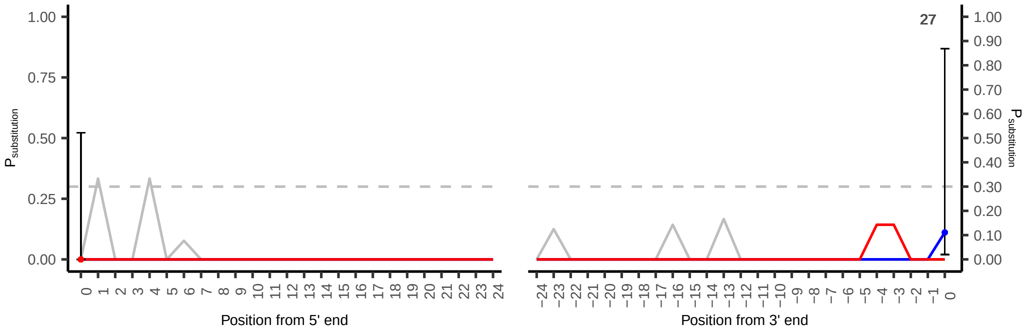
F.9 ELF034

No data for ELF034_157, 132, 094, or 079.

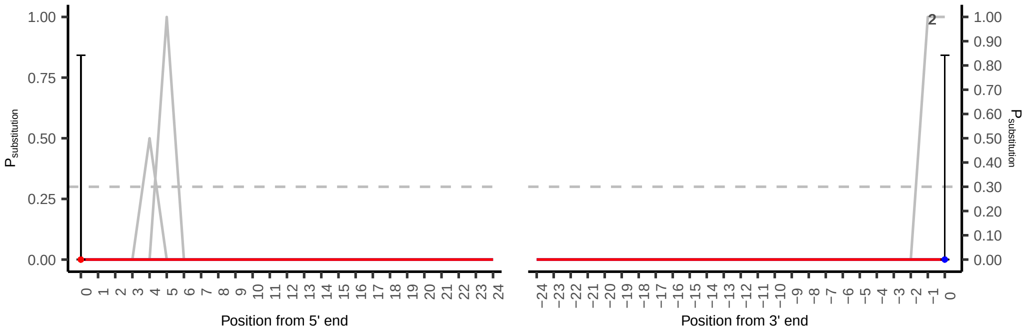
ELF034_219_Embryophyta



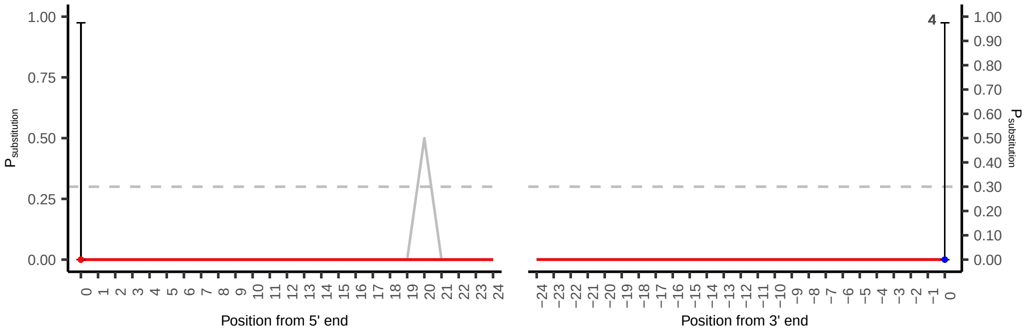
ELF034_202_Embryophyta



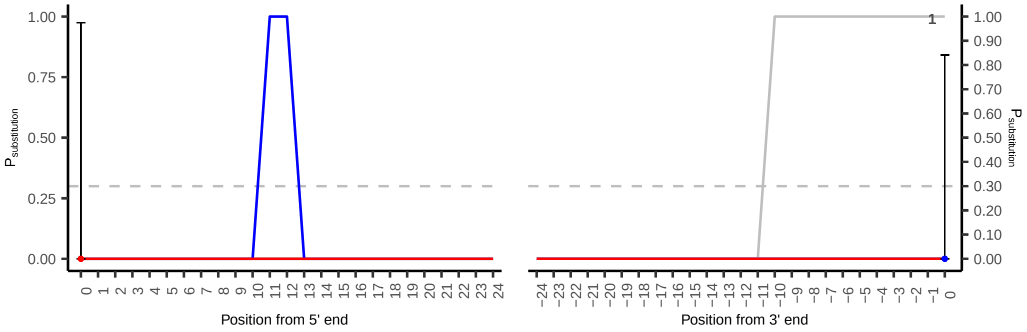
ELF034_185_Embryophyta



ELF034_177_Embryophyta

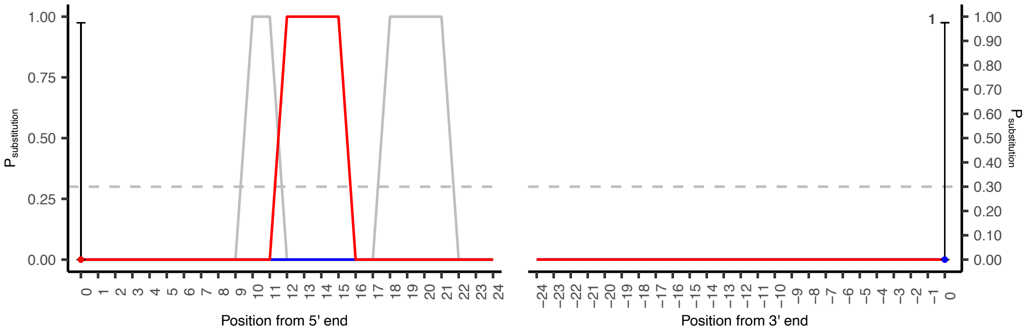


ELF034_061_Embryophyta

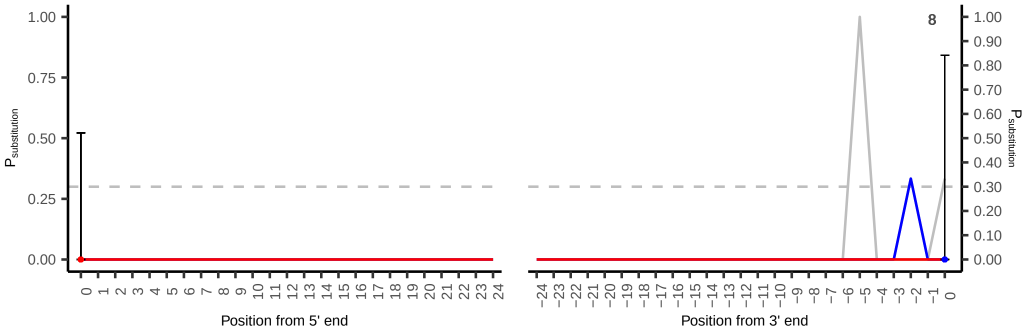


F.10 ELF054

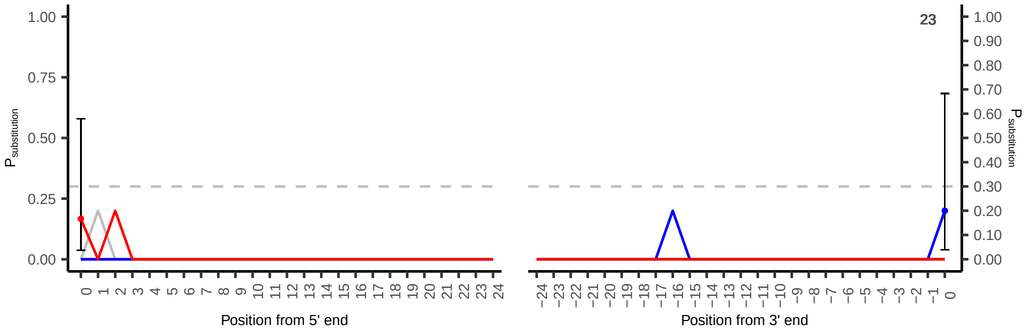
ELF054_356_Embryophyta

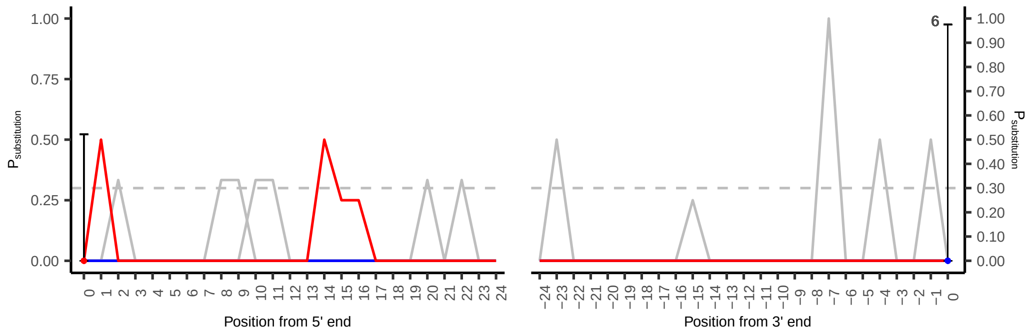


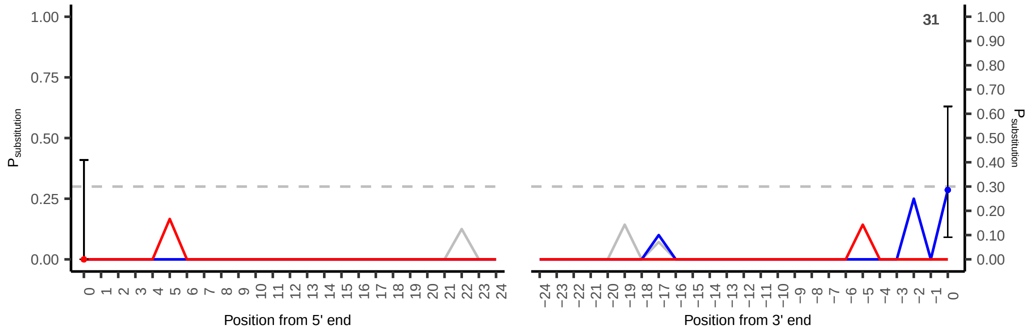
ELF054_330_Embryophyta

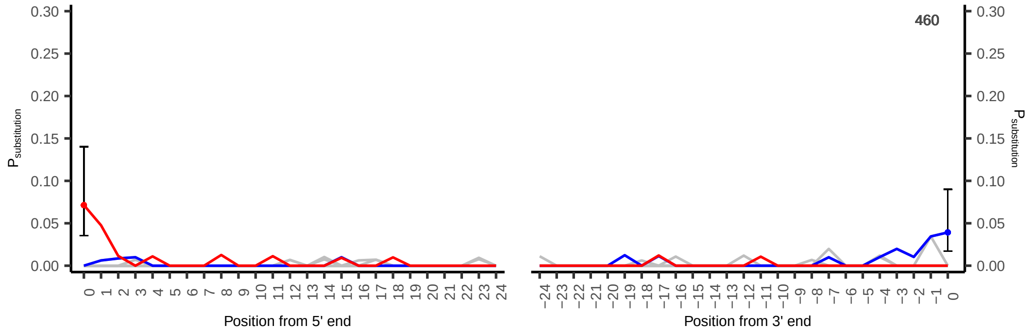


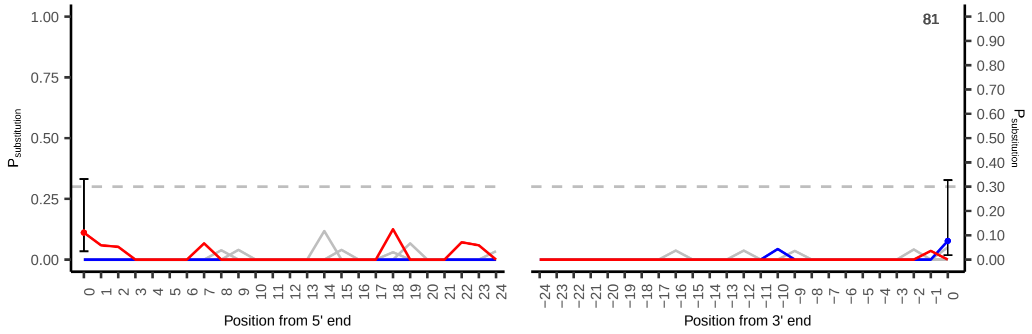
ELF054_315_Embryophyta

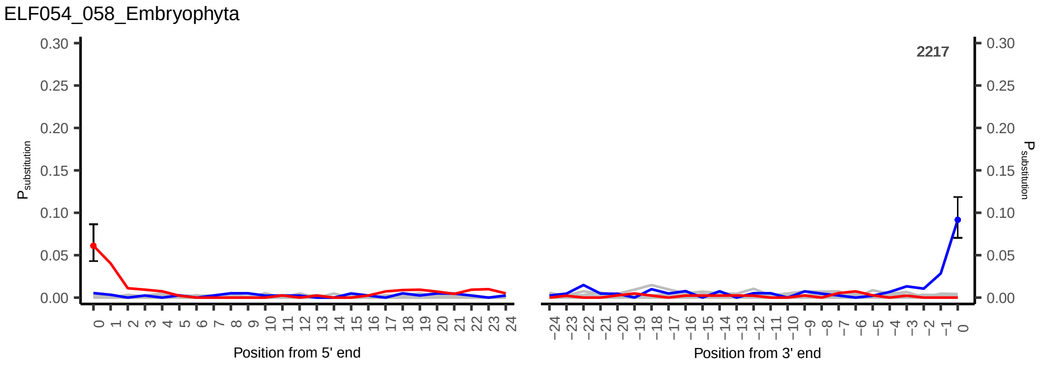



ELF054_291_Embryophyta


ELF054_268_Embryophyta


ELF054_182_Embryophyta

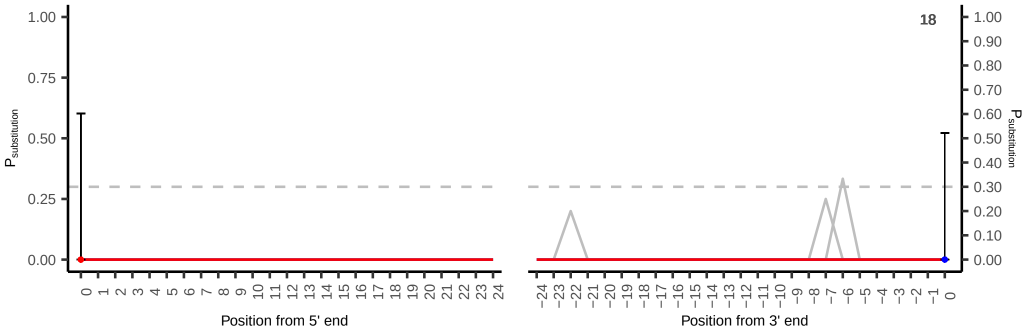

ELF054_140_Embryophyta


ELF054_058_Embryophyta
2217
0.30
0.25
0.20
0.15
0.10
0.05
0.00
$P_{substitution}$
Position from 5' end
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 3' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0

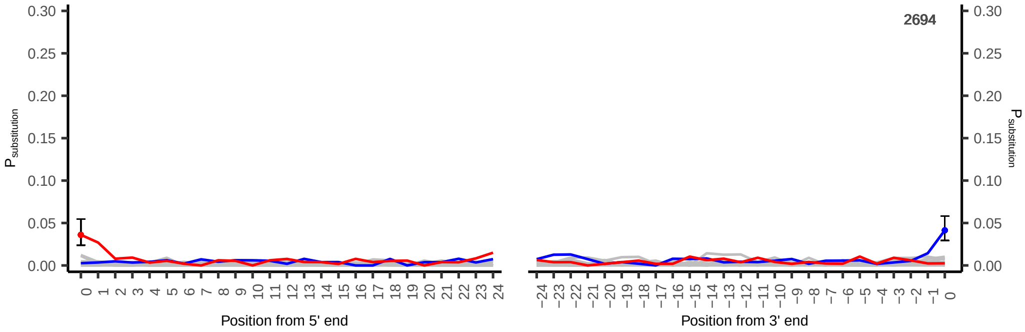
F.11 ELF039

No data for ELF039_145.

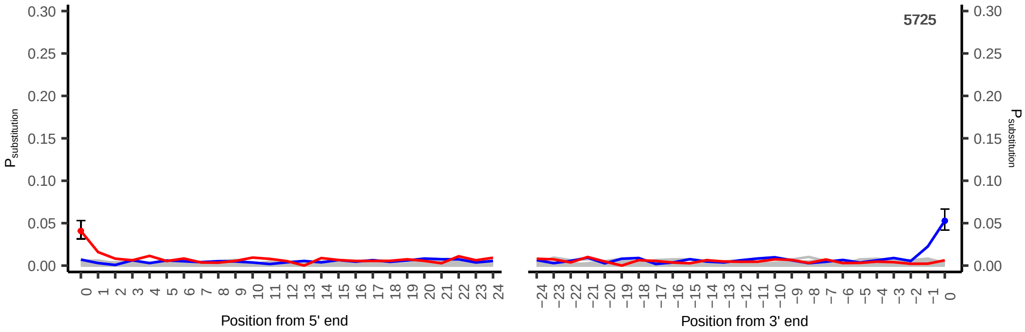
ELF039_485_Embryophyta



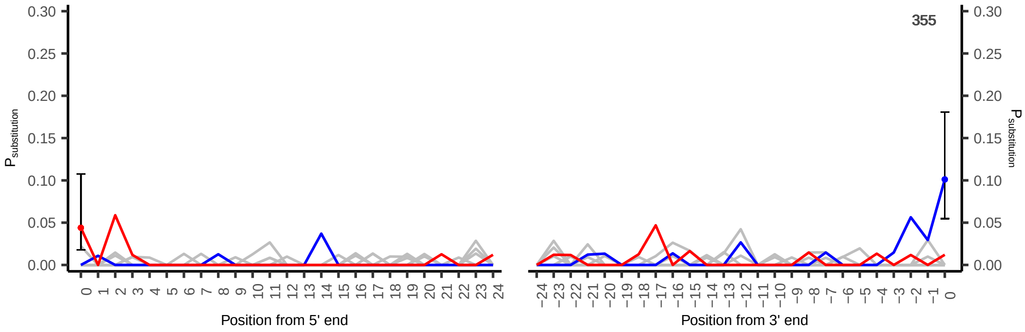
ELF039_460_Embryophyta

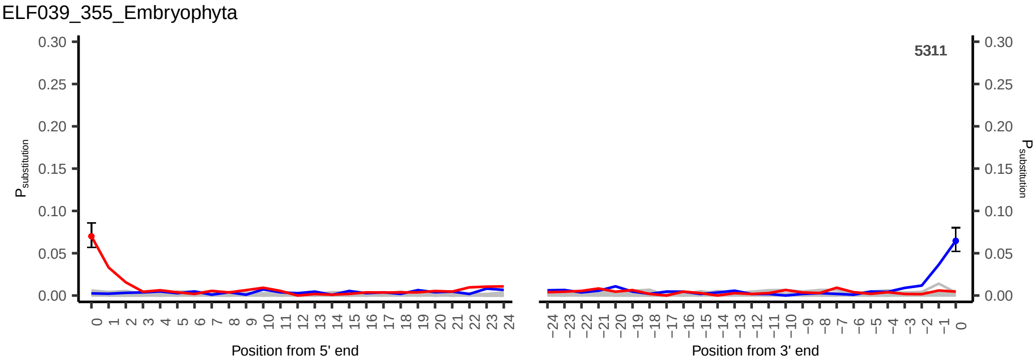


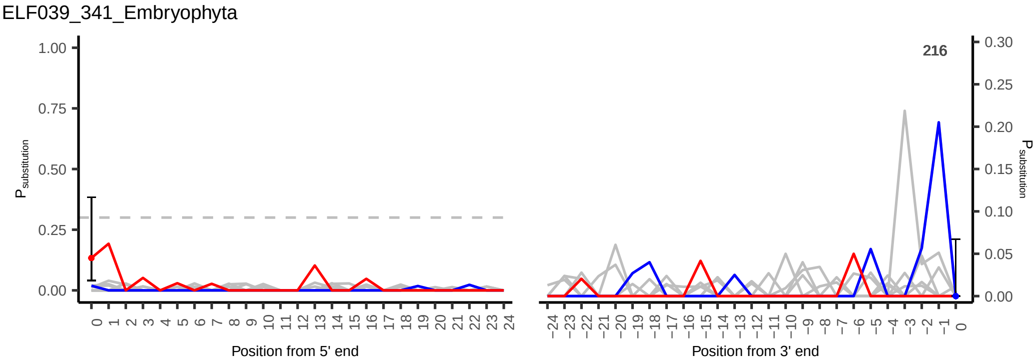
ELF039_415_Embryophyta

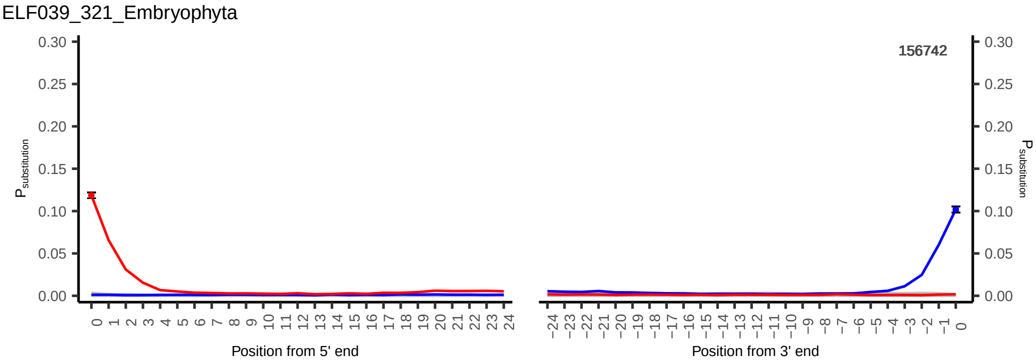


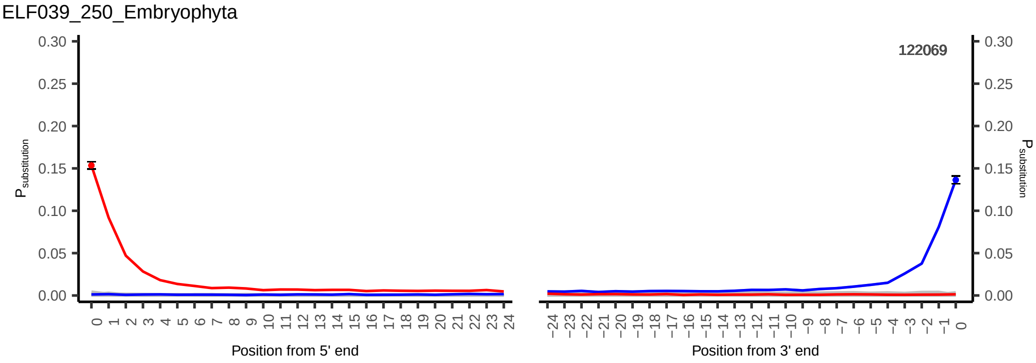
ELF039_384_Embryophyta




ELF039_355_Embryophyta
5311
$P_{substitution}$
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Position from 5' end
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 3' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0

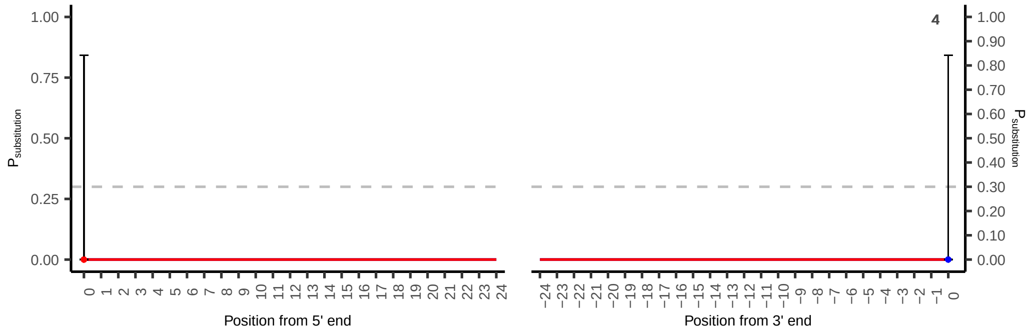

ELF039_341_Embryophyta
216
$P_{substitution}$
0.00
0.25
0.50
0.75
1.00
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Position from 5' end
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 3' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0


ELF039_321_Embryophyta
156742
$P_{substitution}$
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Position from 5' end
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 3' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0

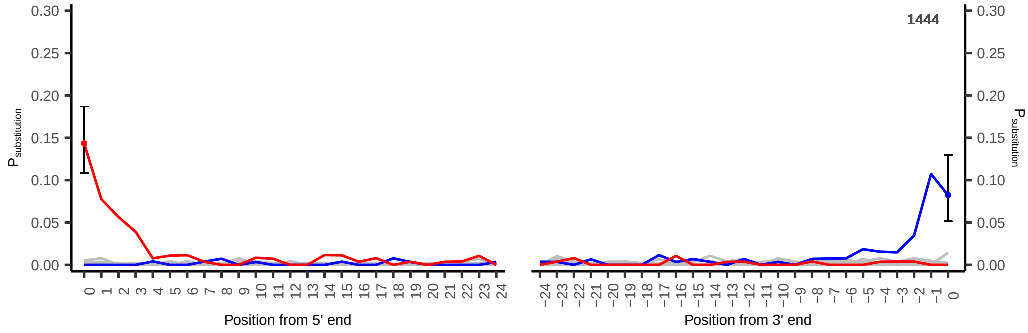

ELF039_250_Embryophyta
122069
$P_{substitution}$
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Position from 5' end
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 3' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0

F.12 ELF040A

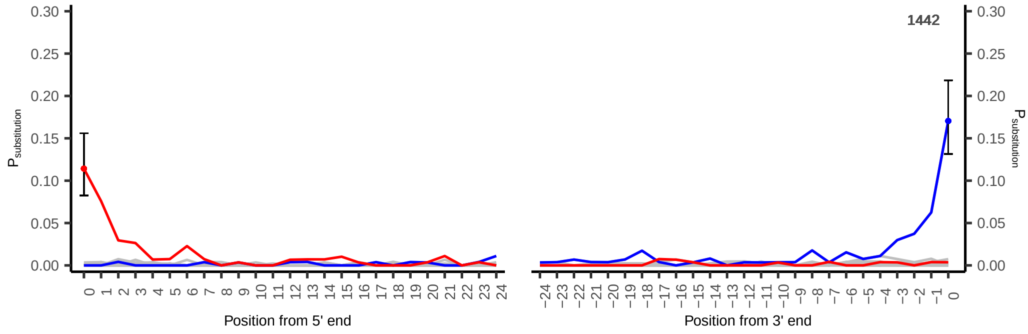
ELF040A_487_Embryophyta



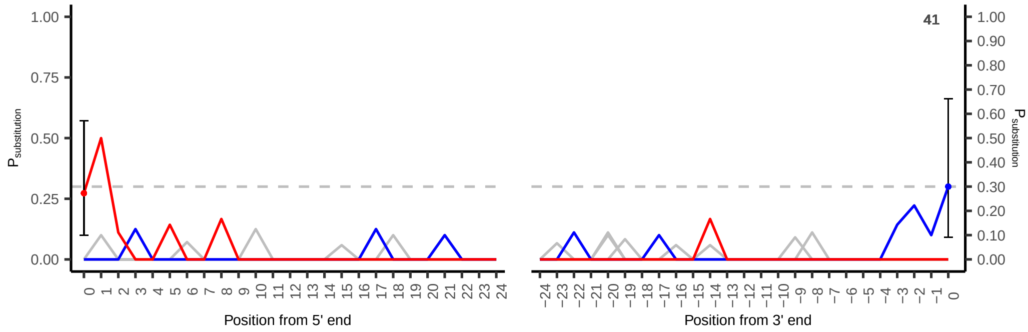
ELF040A_350_Embryophyta



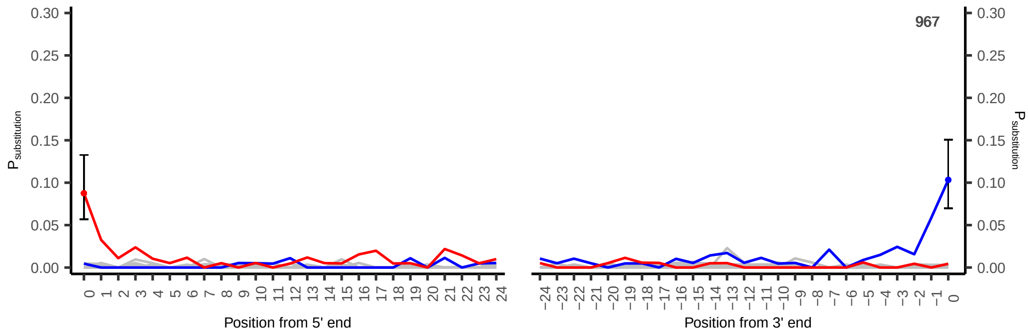
ELF040A_298_Embryophyta



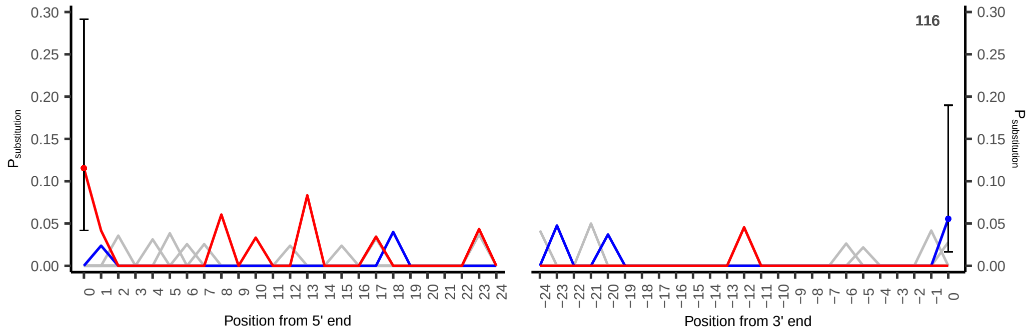
ELF040A_208_Embryophyta



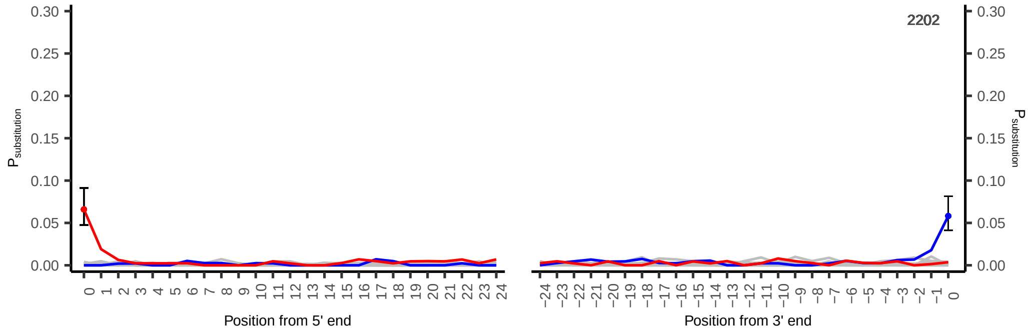
ELF040A_192_Embryophyta



ELF040A_112_Embryophyta

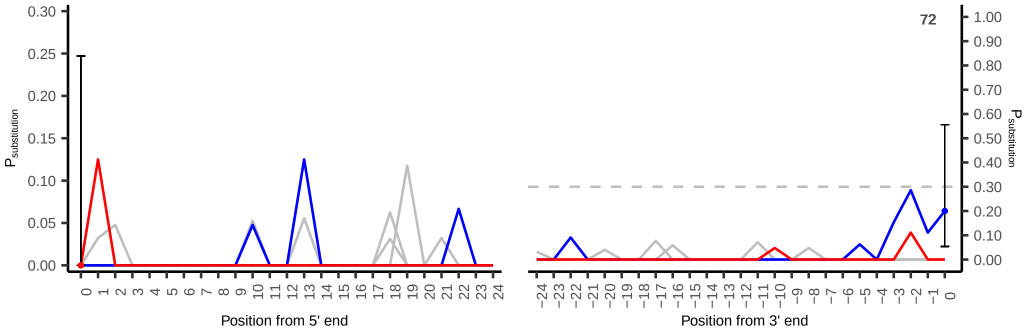


ELF040A_095_Embryophyta

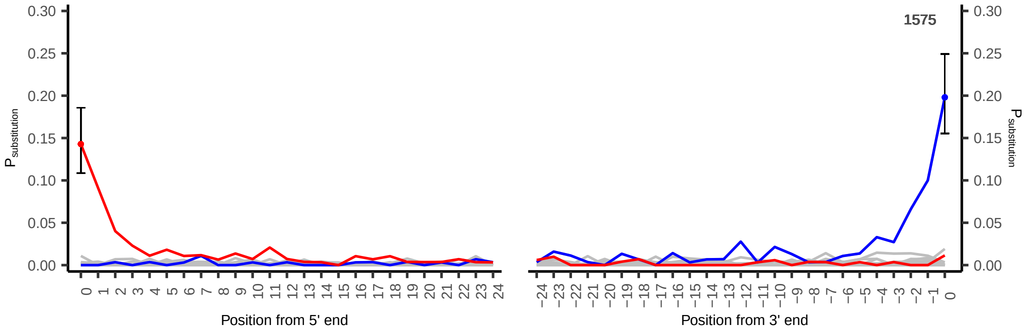


F.13 ELF041

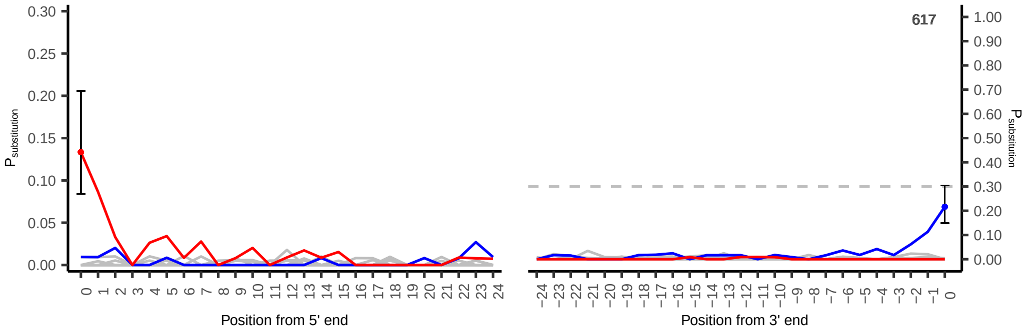
ELF041_295_Embryophyta



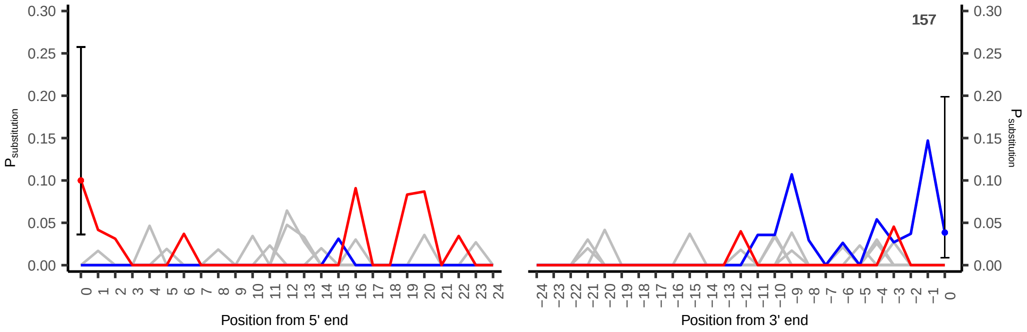
ELF041_180_Embryophyta



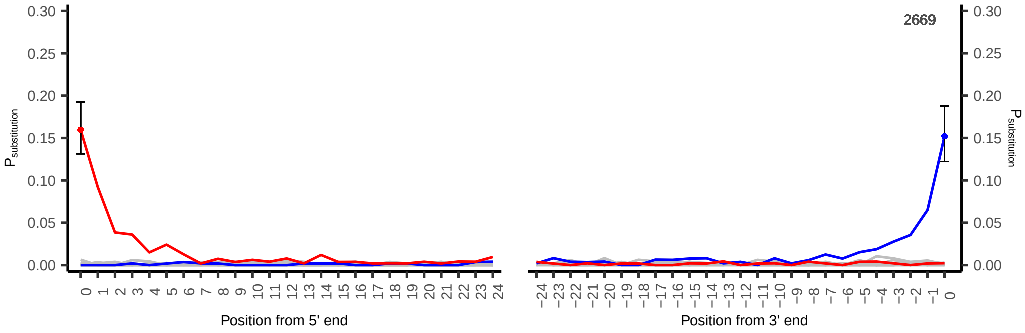
ELF041_110_Embryophyta

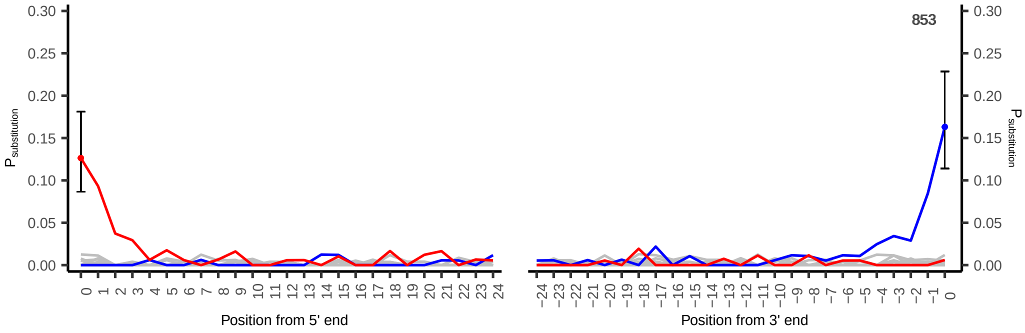


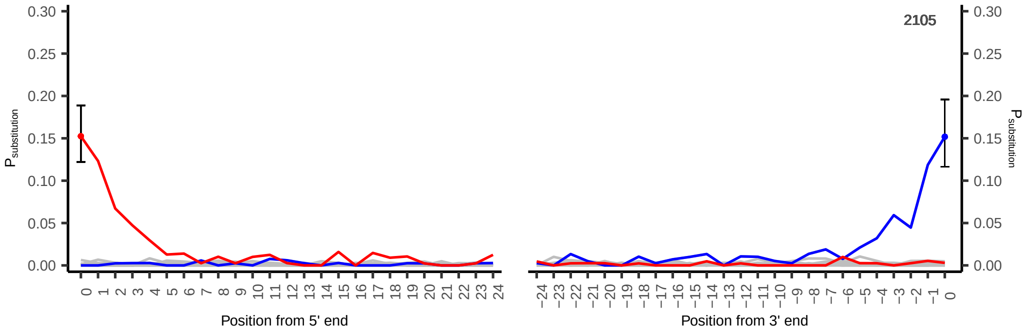
ELF041_087_Embryophyta

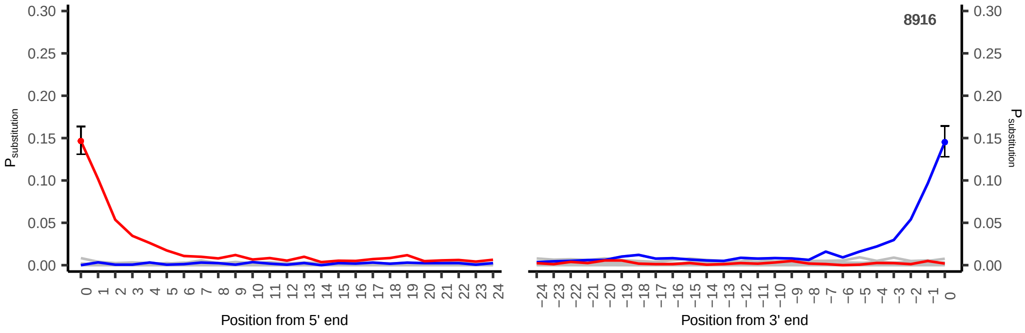


F.14 ELF042


ELF042_350_Embryophyta

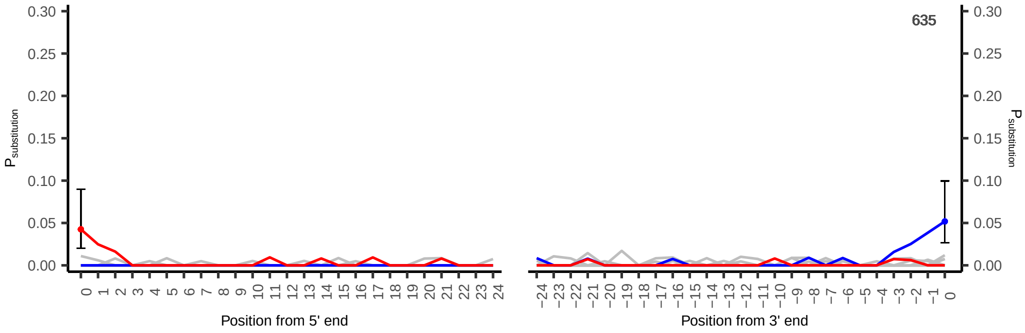

ELF042_250_Embryophyta


ELF042_151_Embryophyta

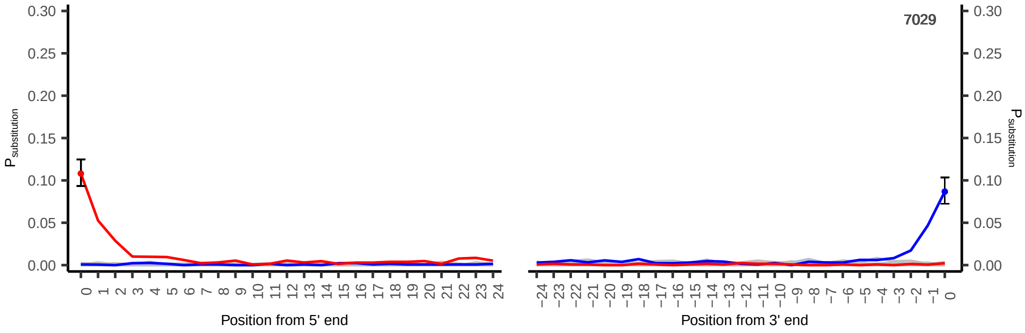

ELF042_065_Embryophyta

F.15 ELF044

ELF044_137_Embryophyta

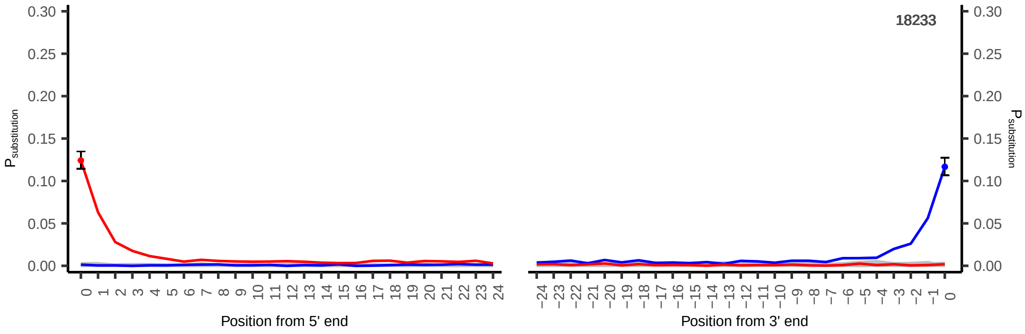


ELF044_090_Embryophyta



F.16 ELF044A

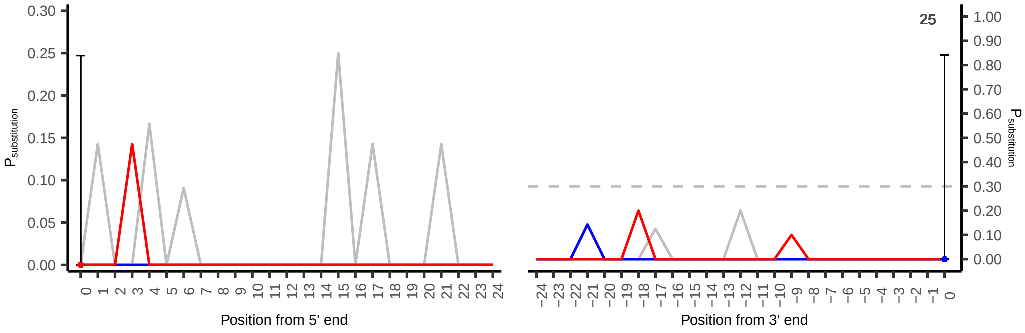
ELF044A_097_Embryophyta



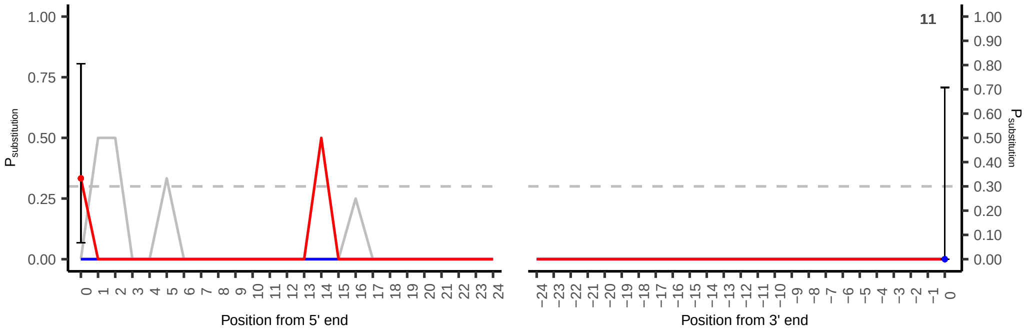
F.17 ELF031

No data for ELF031_033.

ELF031_056_Embryophyta



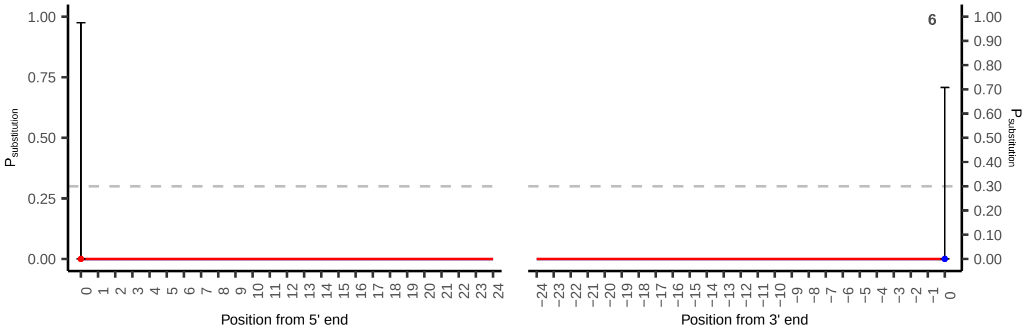
ELF031_043_Embryophyta



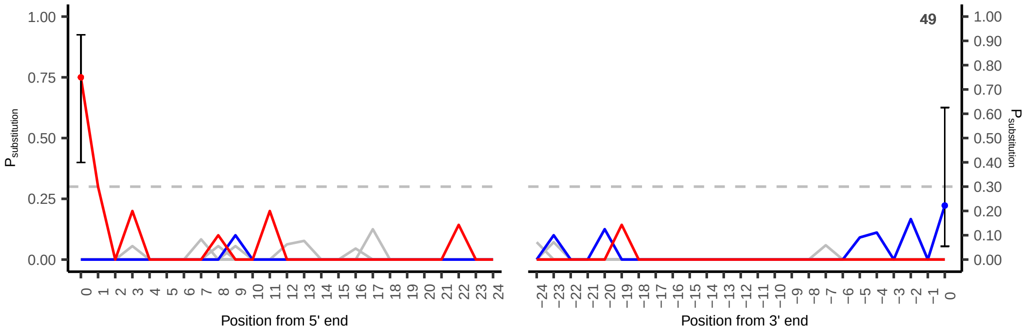
F.18 ELF031A

No data for ELF031A_310 or 219.

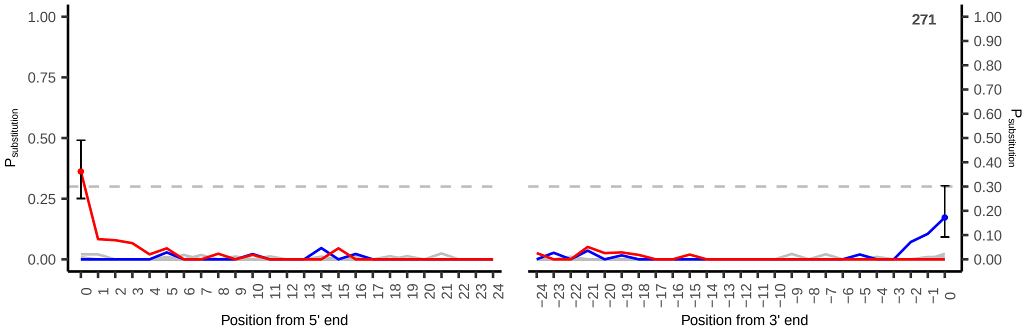
ELF031A_281_Embryophyta



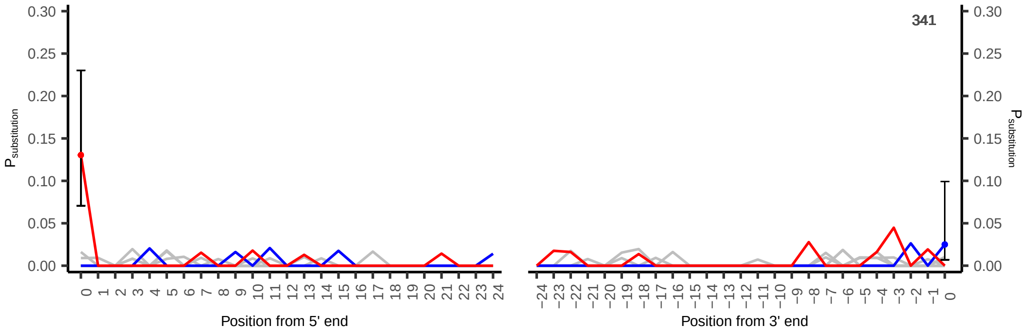
ELF031A_202_Embryophyta



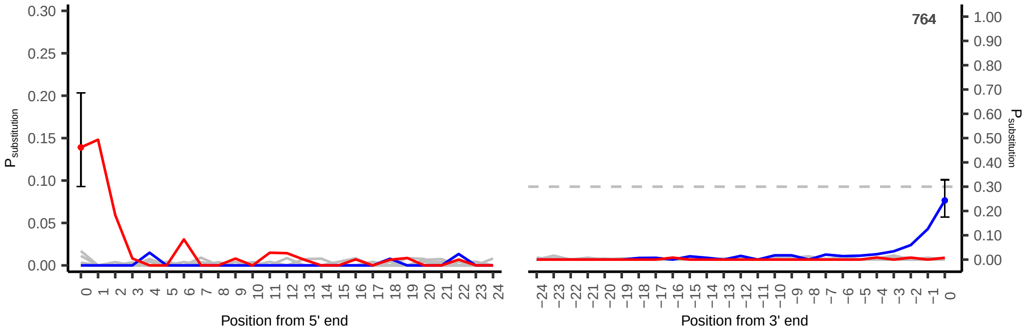
ELF031A_177_Embryophyta

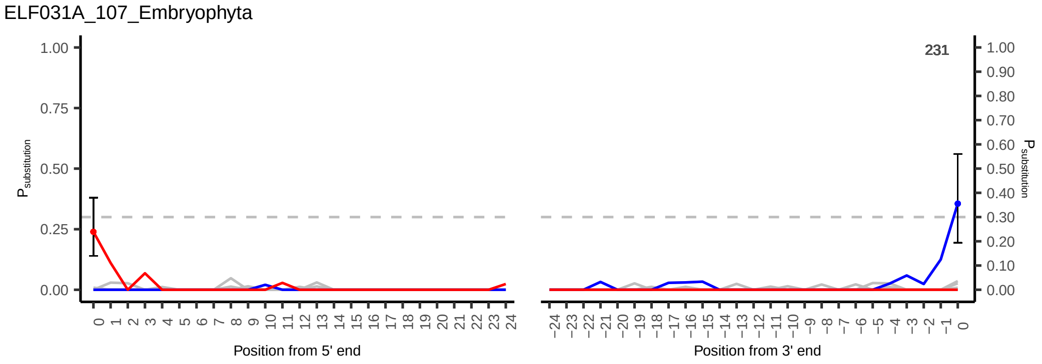


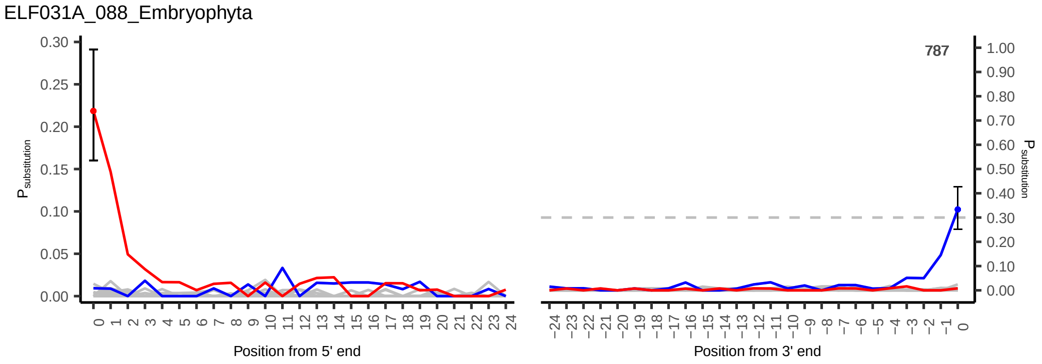
ELF031A_152_Embryophyta

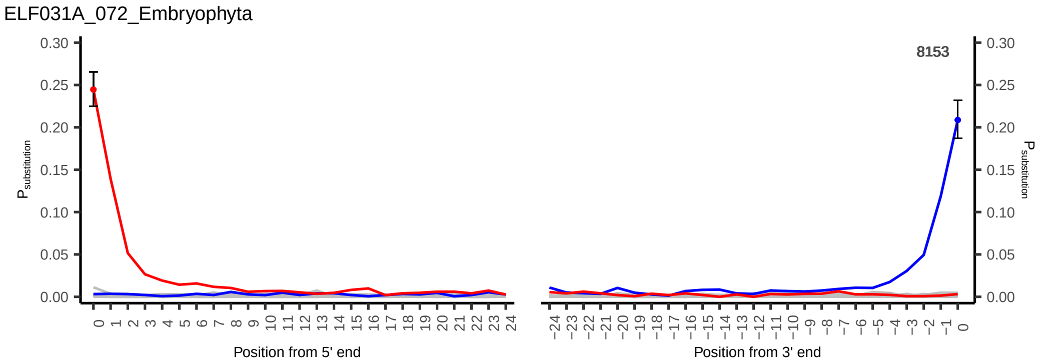


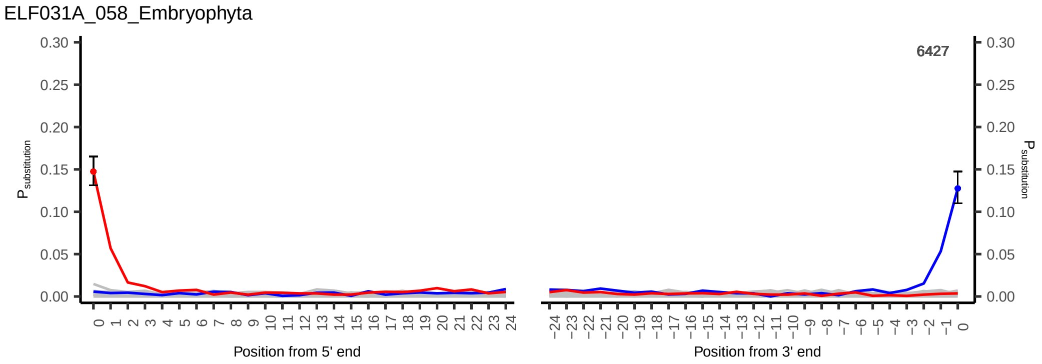
ELF031A_123_Embryophyta




ELF031A_107_Embryophyta
231
$P_{substitution}$
1.00
0.75
0.50
0.25
0.00
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Position from 5' end
Position from 3' end

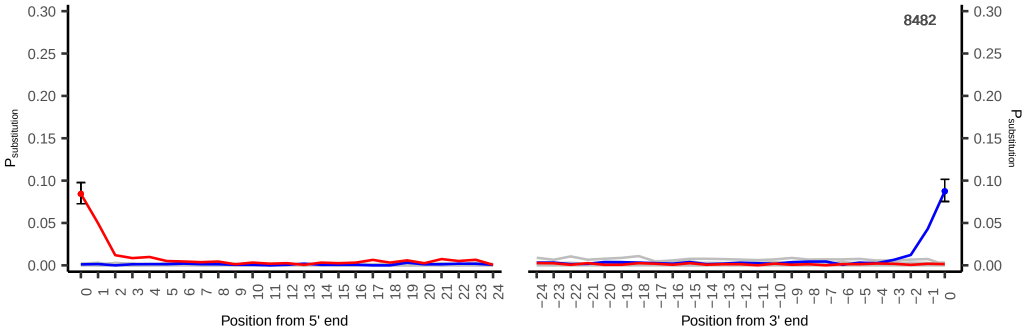

ELF031A_088_Embryophyta
787
$P_{substitution}$
0.30
0.25
0.20
0.15
0.10
0.05
0.00
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Position from 5' end
Position from 3' end


ELF031A_072_Embryophyta
8153
$P_{substitution}$
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Position from 5' end
Position from 3' end

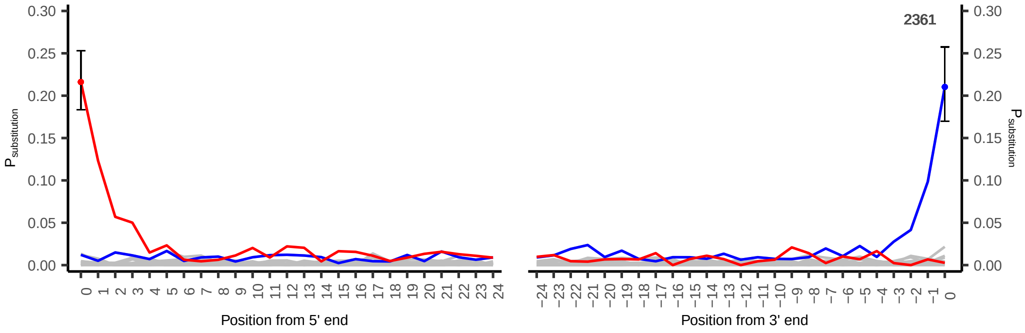

ELF031A_058_Embryophyta
6427
$P_{substitution}$
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Position from 5' end
Position from 3' end

F.19 ELF051

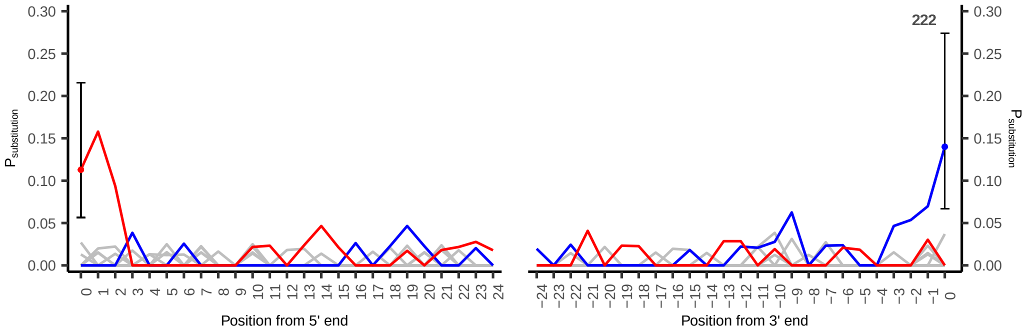
ELF051_292_Embryophyta



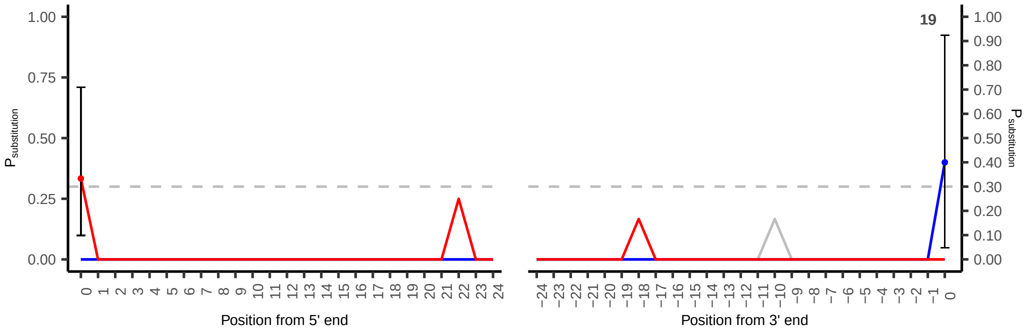
ELF051_255_Embryophyta

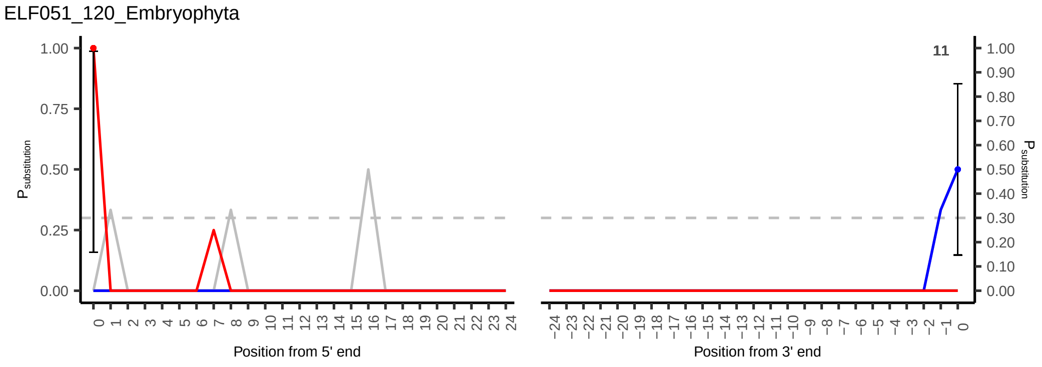


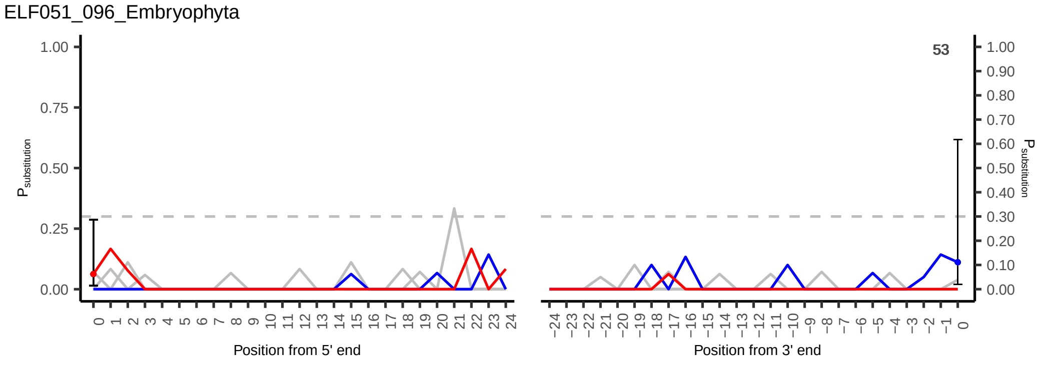
ELF051_196_Embryophyta



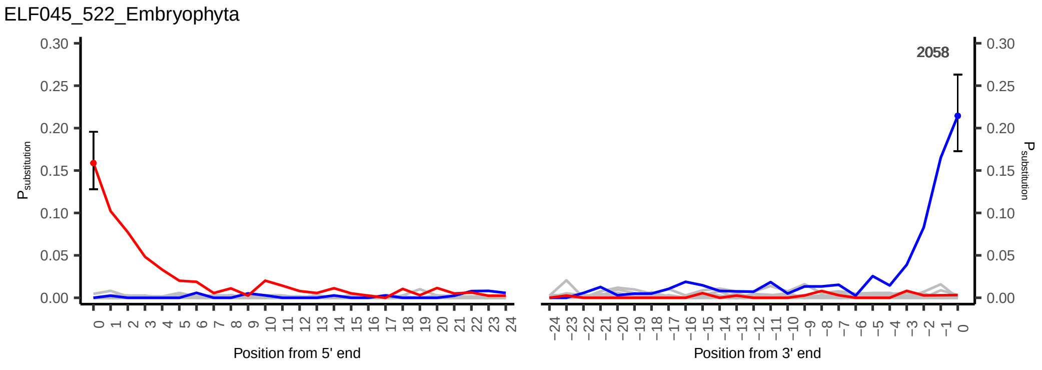
ELF051_151_Embryophyta

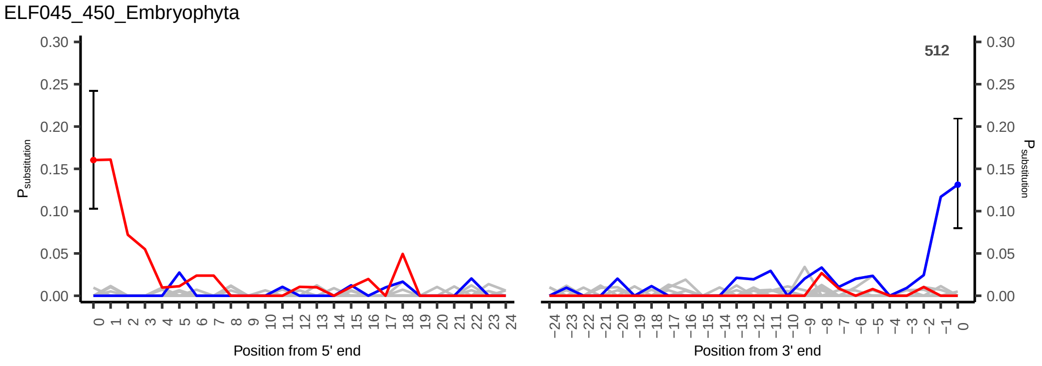



ELF051_120_Embryophyta
11
Psubstitution
1.00
0.75
0.50
0.25
0.00
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

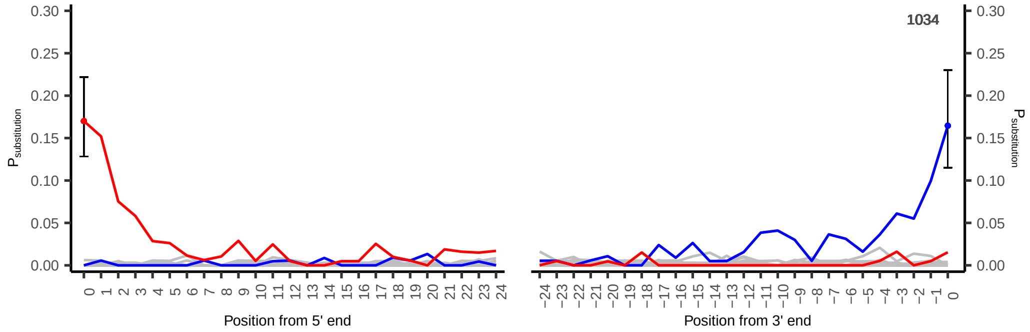

ELF051_096_Embryophyta
53
Psubstitution
1.00
0.75
0.50
0.25
0.00
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

F.20 ELF045

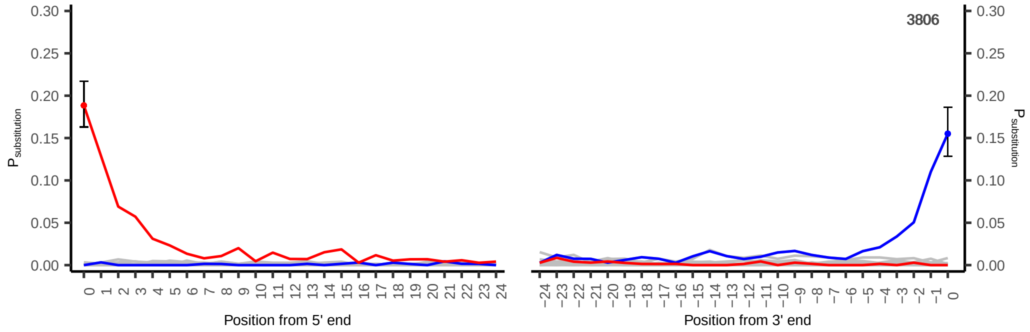

ELF045_522_Embryophyta
2058
Psubstitution
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end


ELF045_450_Embryophyta
512
Psubstitution
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

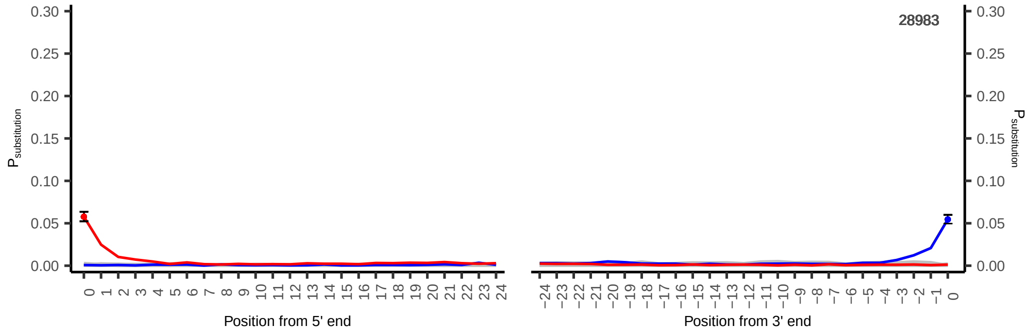
ELF045_346_Embryophyta



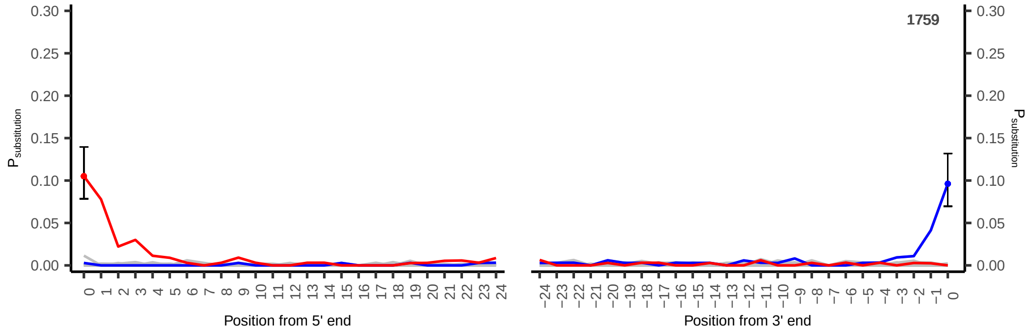
ELF045_252_Embryophyta



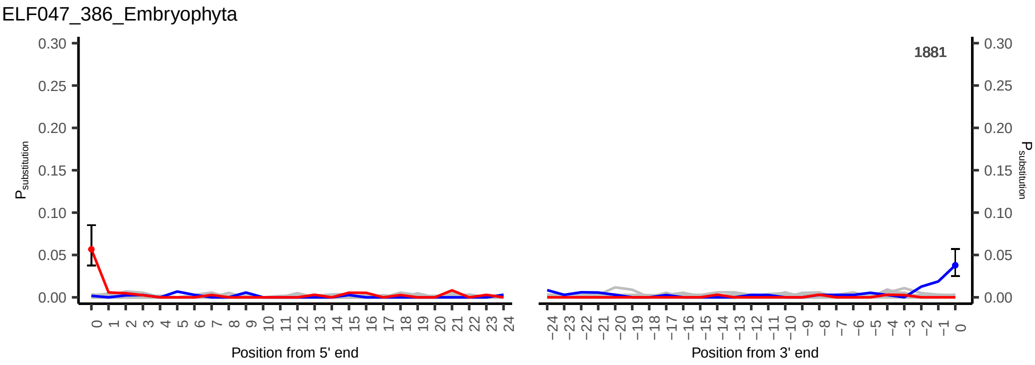
ELF045_145_Embryophyta

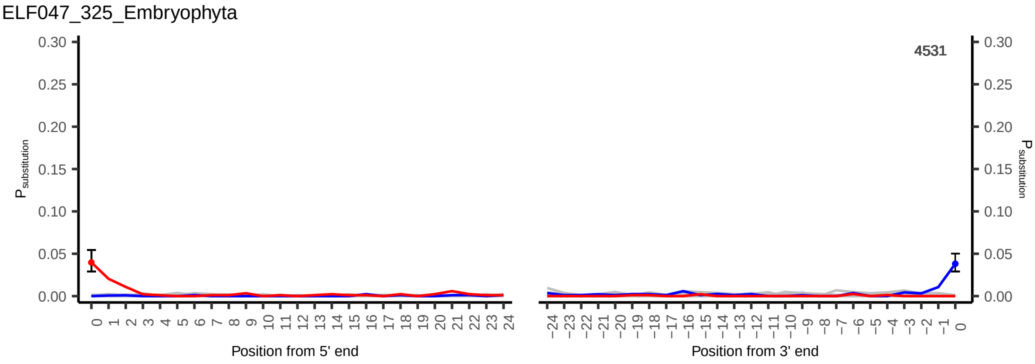


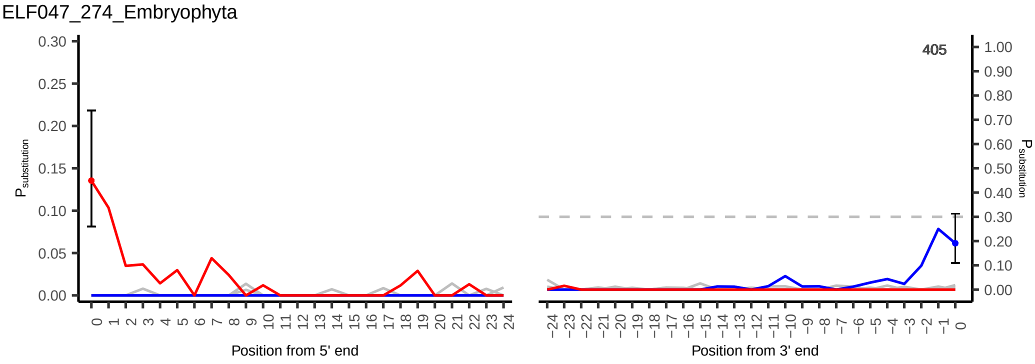
ELF045_090_Embryophyta

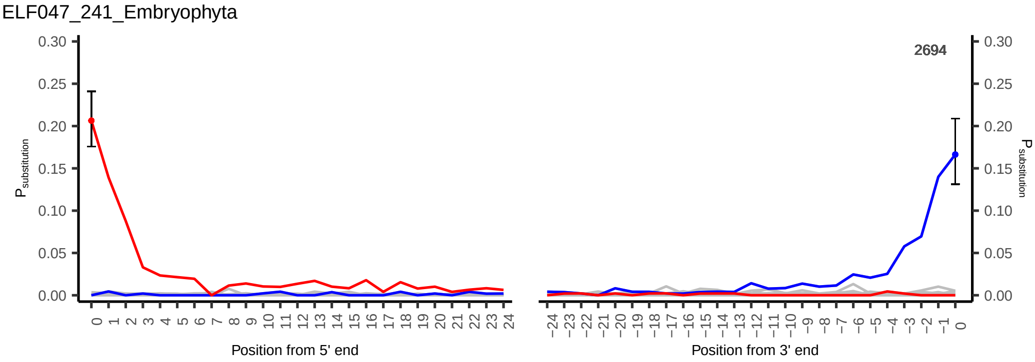


F.21 ELF047

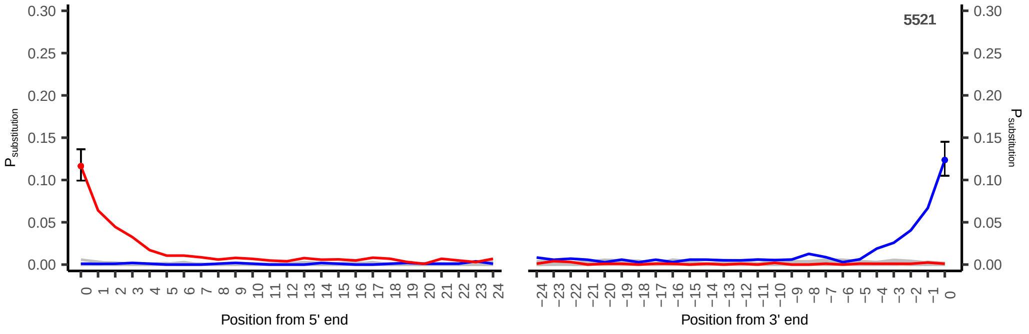

ELF047_386_Embryophyta
1881
P substitution
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end


ELF047_325_Embryophyta
4531
P substitution
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

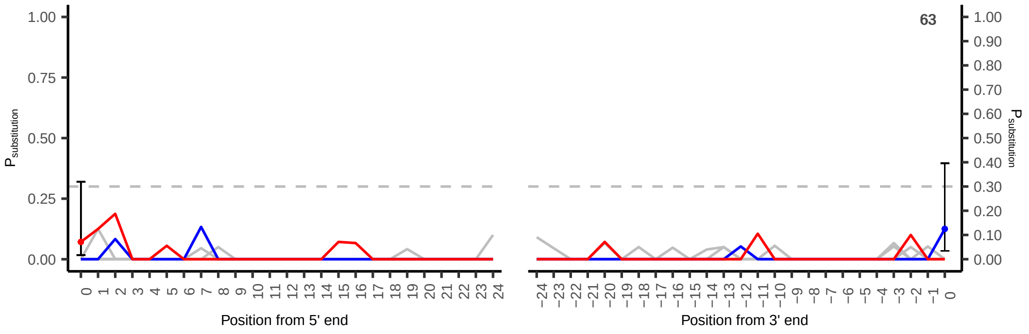

ELF047_274_Embryophyta
405
P substitution
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end


ELF047_241_Embryophyta
2694
P substitution
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

ELF047_150_Embryophyta

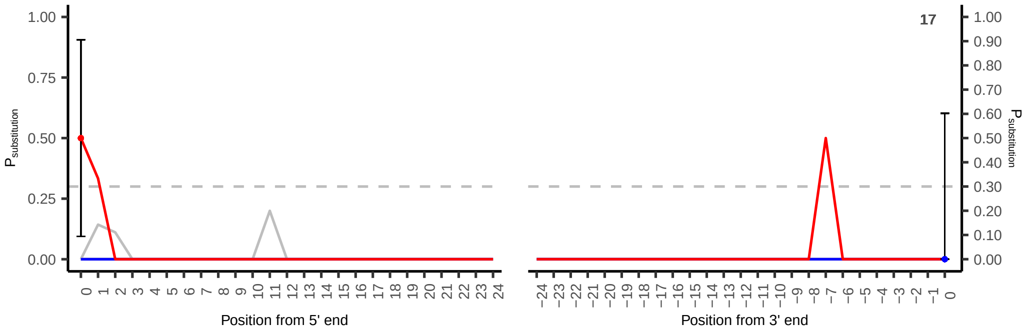


ELF047_070_Embryophyta

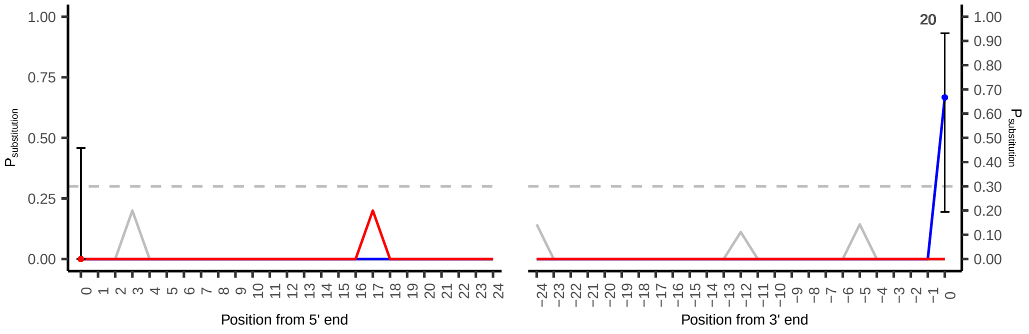


F.22 ELF047A

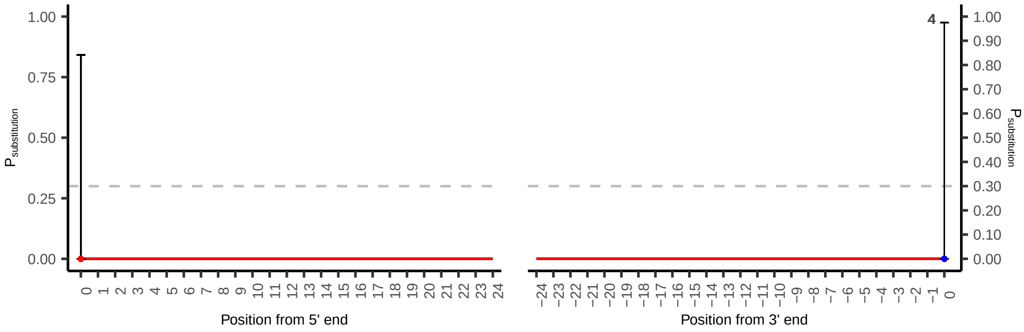
ELF047A_354_Embryophyta



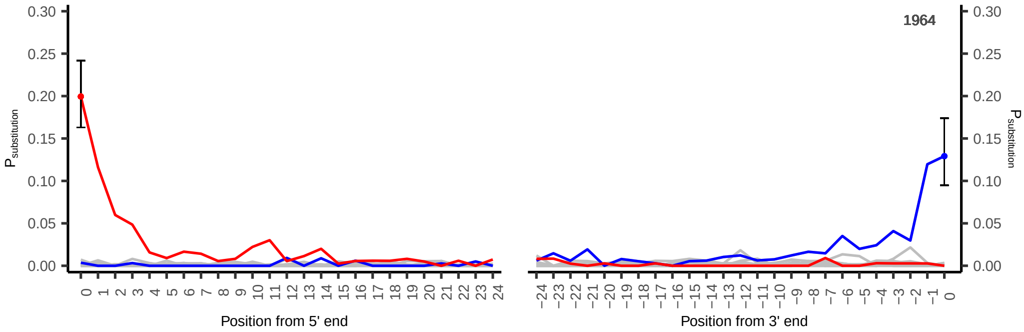
ELF047A_256_Embryophyta



ELF047A_150_Embryophyta

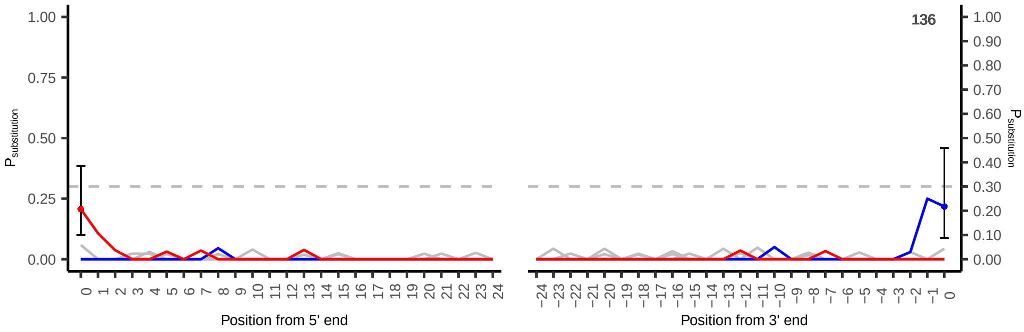


ELF047A_050_Embryophyta

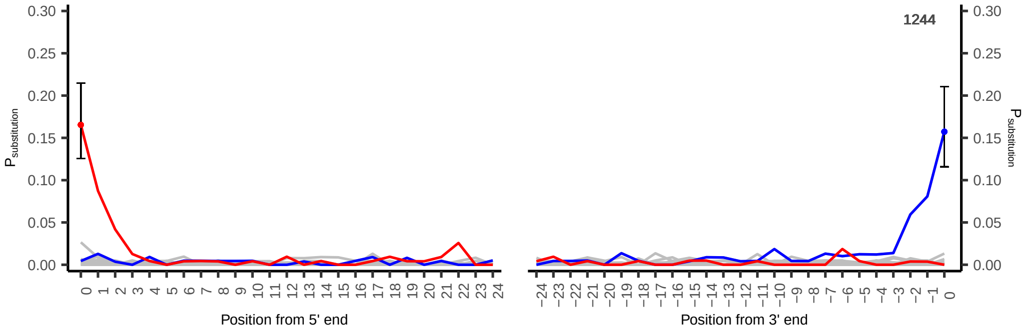


F.23 ELF049

ELF049_361_Embryophyta



ELF049_295_Embryophyta

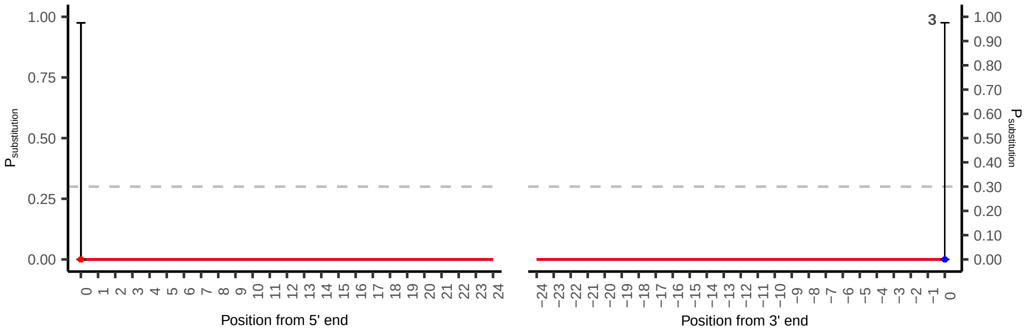


F.24 ELF046A

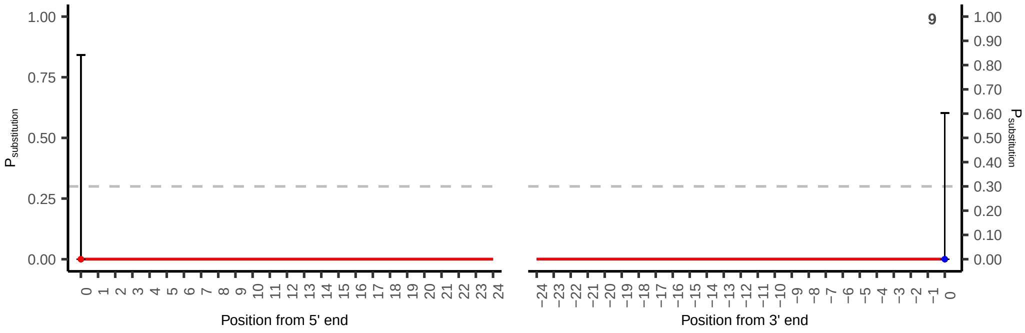
No data for ELF046A_270.

F.25 ELF050

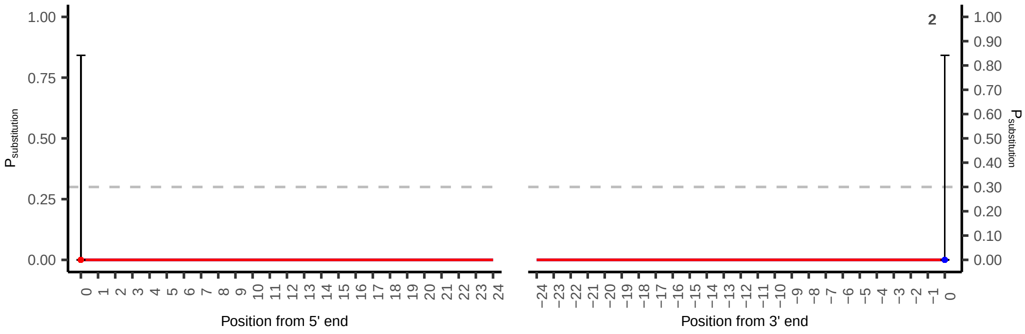
ELF050_595_Embryophyta

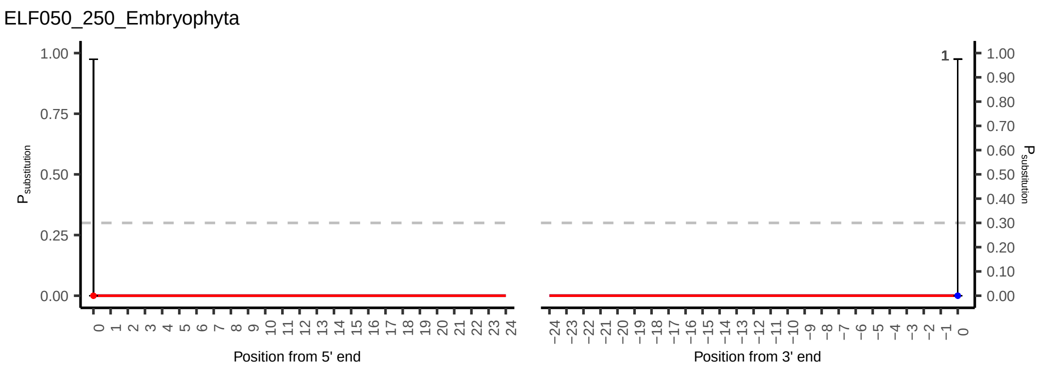


ELF050_450_Embryophyta



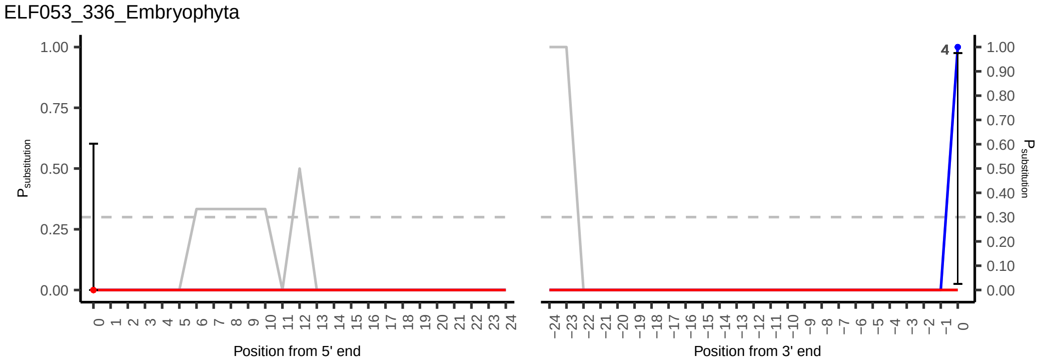
ELF050_354_Embryophyta

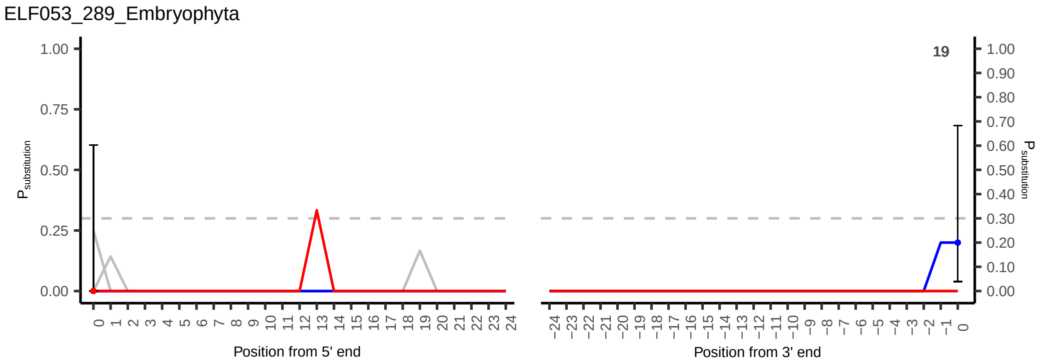


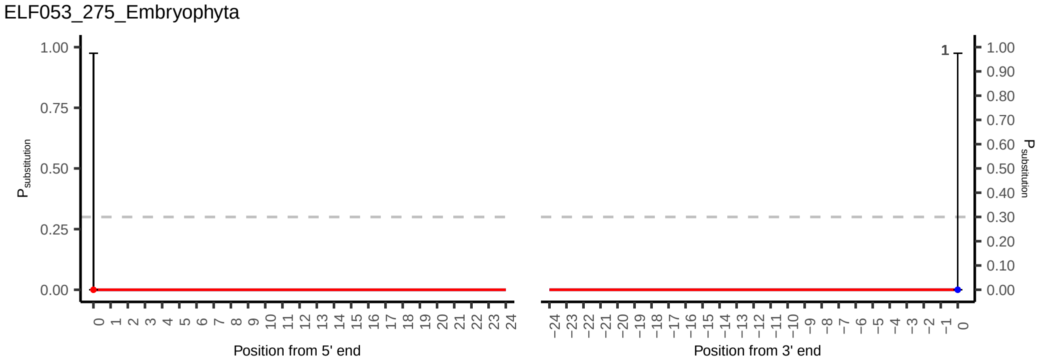

ELF050_250_Embryophyta

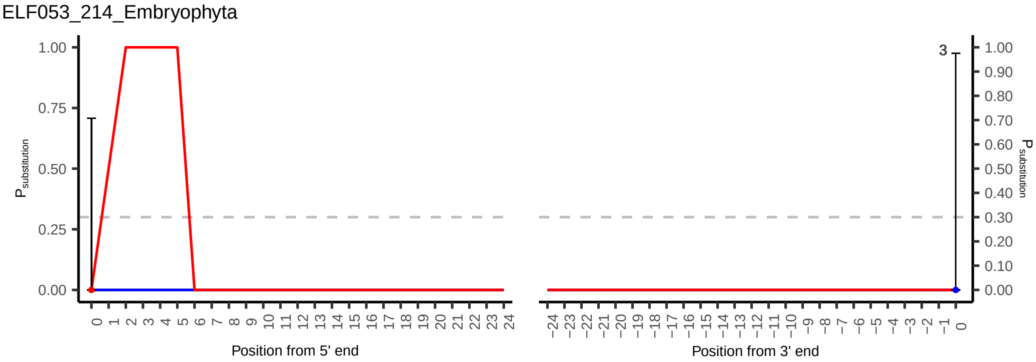
F.26 ELF053

No data in ELF053_179.


ELF053_336_Embryophyta


ELF053_289_Embryophyta


ELF053_275_Embryophyta

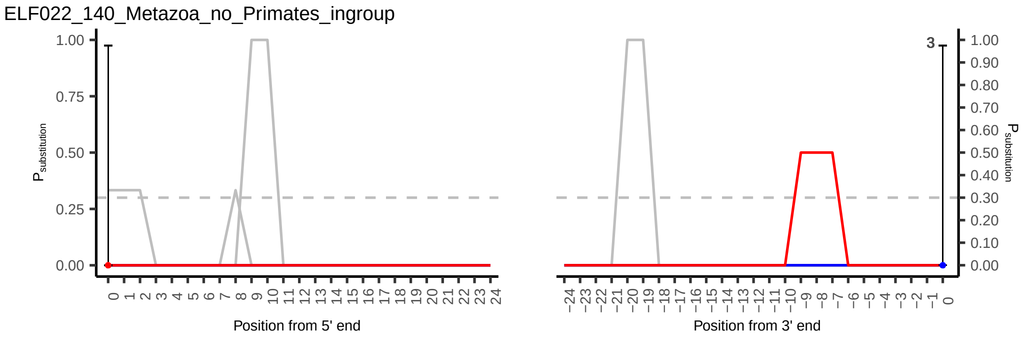

ELF053_214_Embryophyta
1.00
0.75
0.50
0.25
0.00
$P_{substitution}$
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end
3
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
$P_{substitution}$

Appendix G

MetaDamage plots per sample for Metazoa (no Primates)

MetaDamage profiles for Metazoa (excluding Primates) reads in each sample, grouped by core and ordered by ascending depth. $P_{\text{substitution}}$ is the proportion of reads that, at a particular position, present a different base to their reference sequence. C-to-T and G-to-A mismatch proportions are in red and blue respectively; others are in grey. Error bars are 95% confidence intervals. Total reads analysed are in the top right. Plots for samples with no data are omitted.

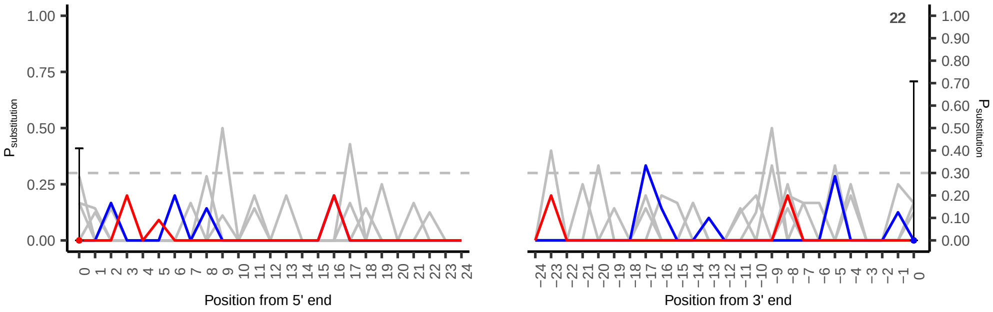
G.0.1 ELF022



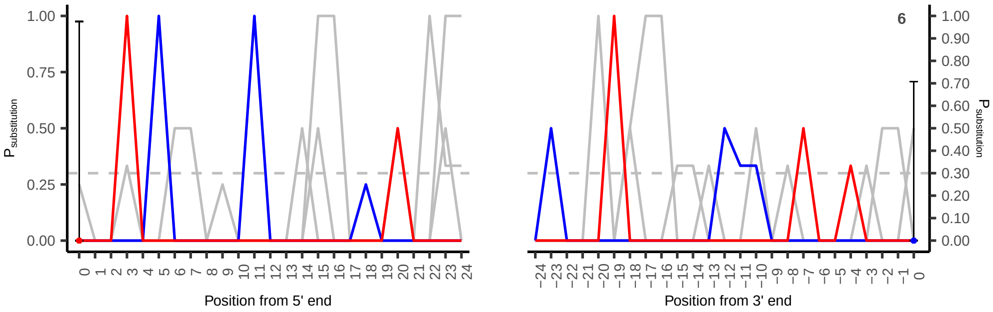
G.1 ELF027

No data for ELF027_177.

ELF027_174_Metazoa_no_Primates_ingroup

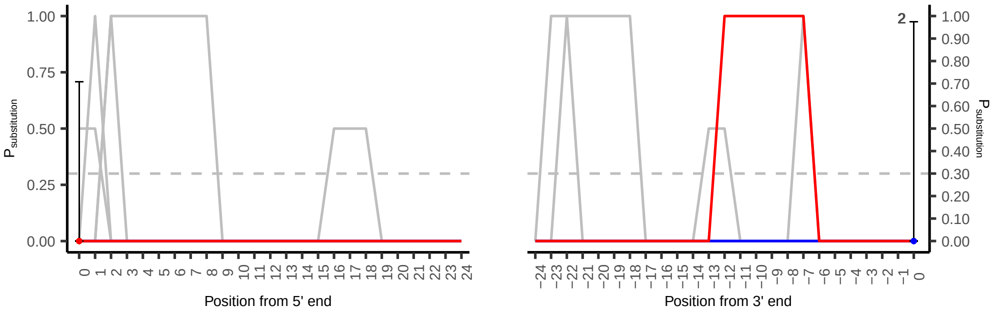


ELF027_160_Metazoa_no_Primates_ingroup

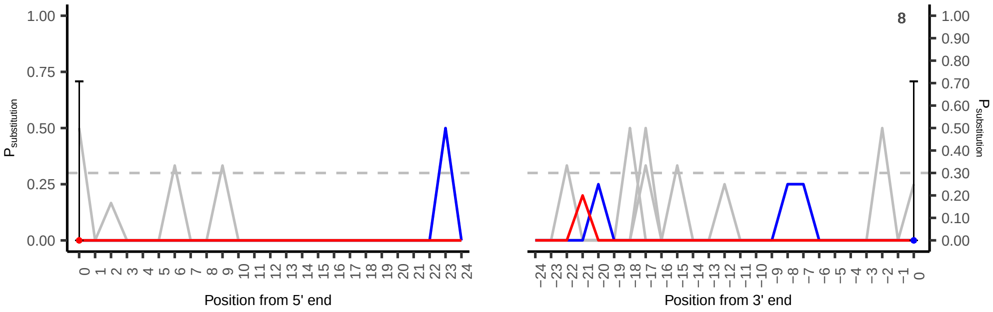


G.2 ELF059

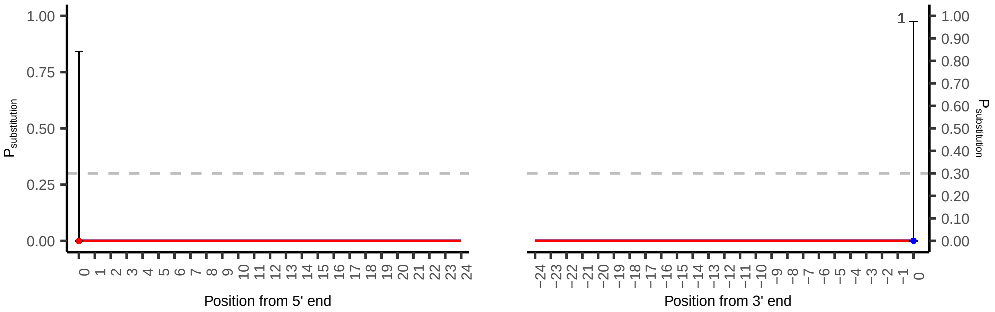
ELF059_378_Metazoa_no_Primates_ingroup



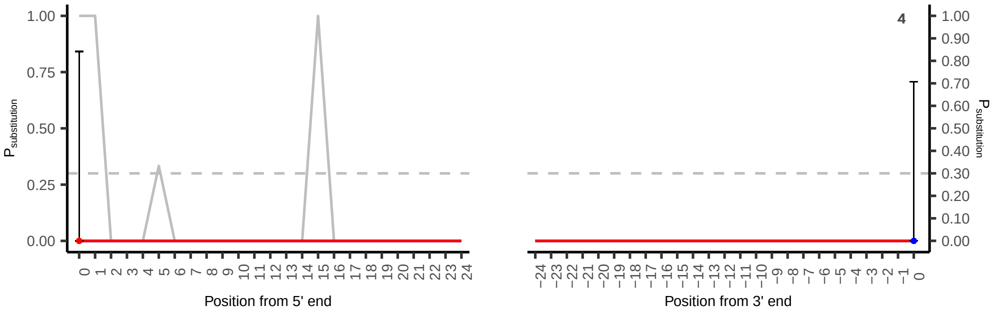
ELF059_359_Metazoa_no_Primates_ingroup



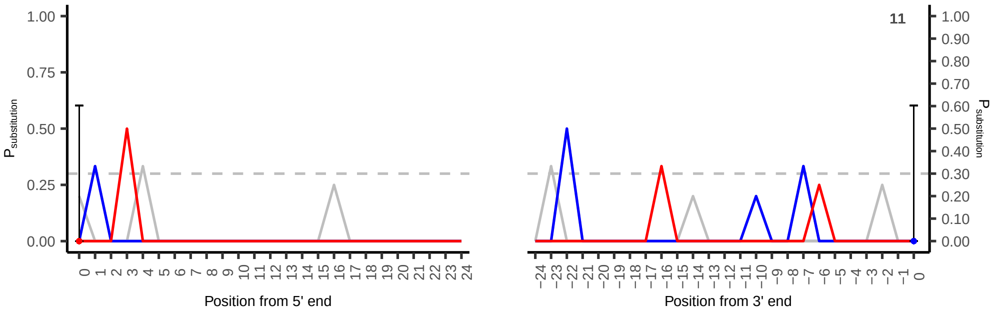
ELF059_337_Metazoa_no_Primates_ingroup



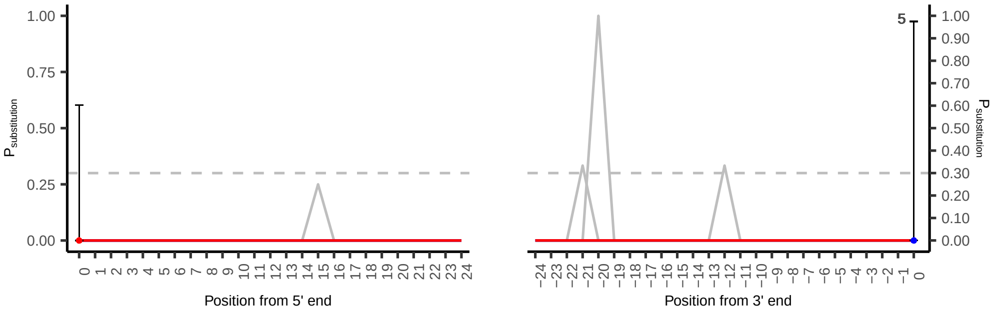
ELF059_280_Metazoa_no_Primates_ingroup



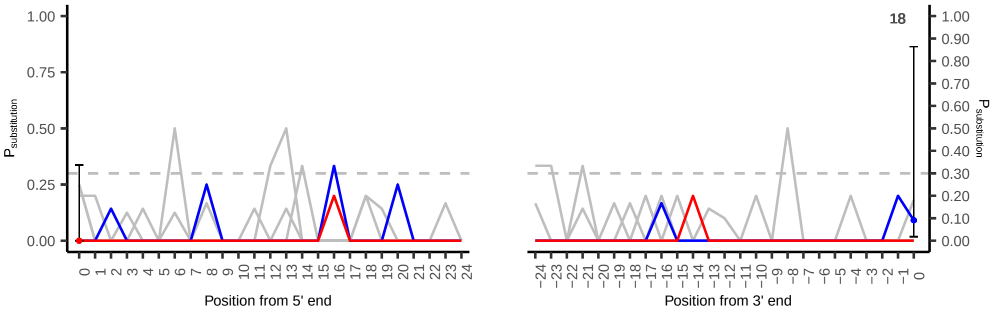
ELF059_270_Metazoa_no_Primates_ingroup



ELF059_230_Metazoa_no_Primates_ingroup

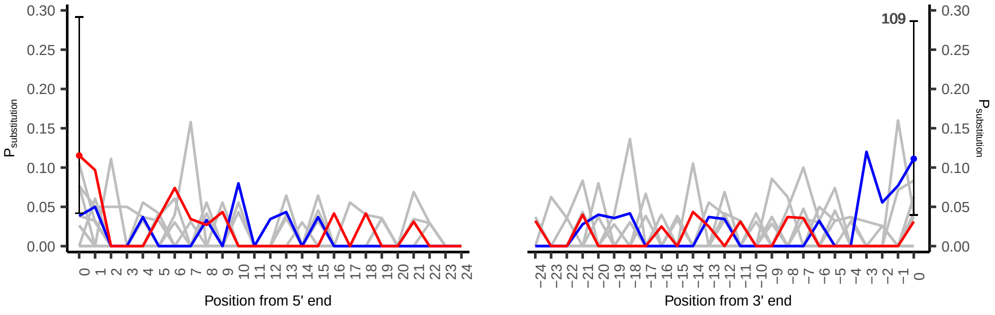


ELF059_210_Metazoa_no_Primates_ingroup

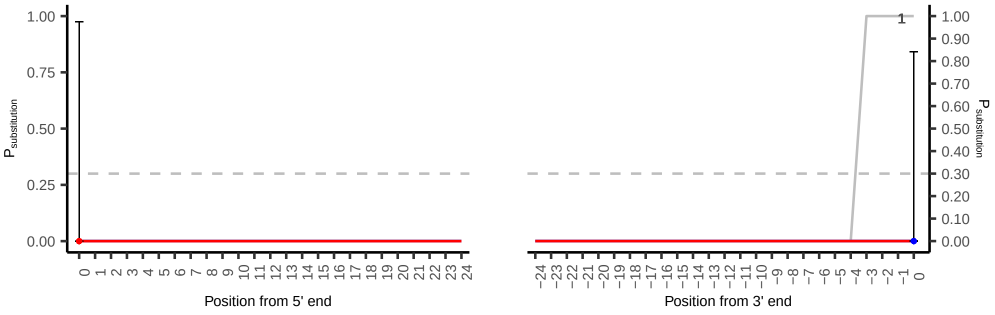


G.3 ELF059A

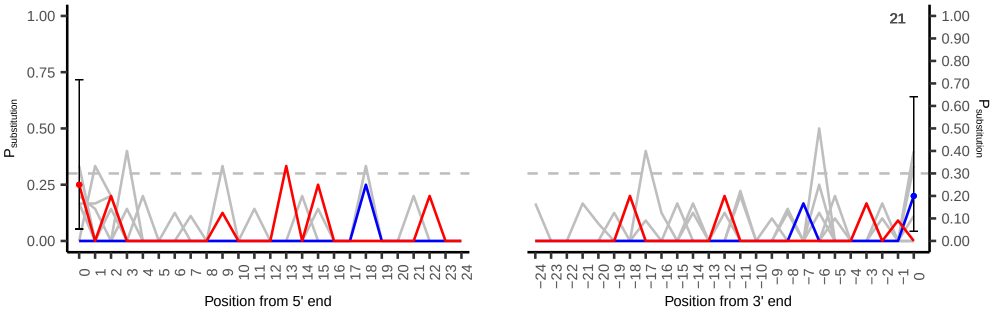
ELF059A_355_Metazoa_no_Primates_ingroup



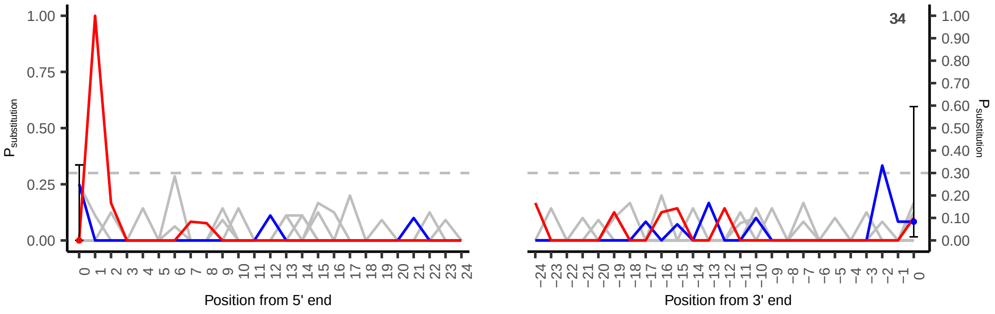
ELF059A_320_Metazoa_no_Primates_ingroup



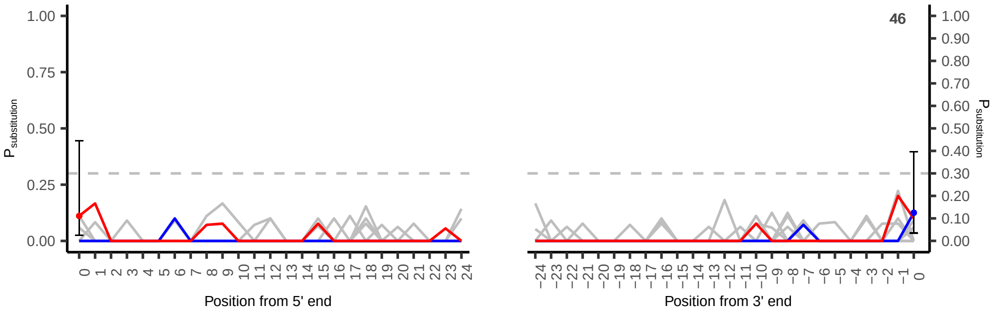
ELF059A_250_Metazoa_no_Primates_ingroup



ELF059A_190_Metazoa_no_Primates_ingroup



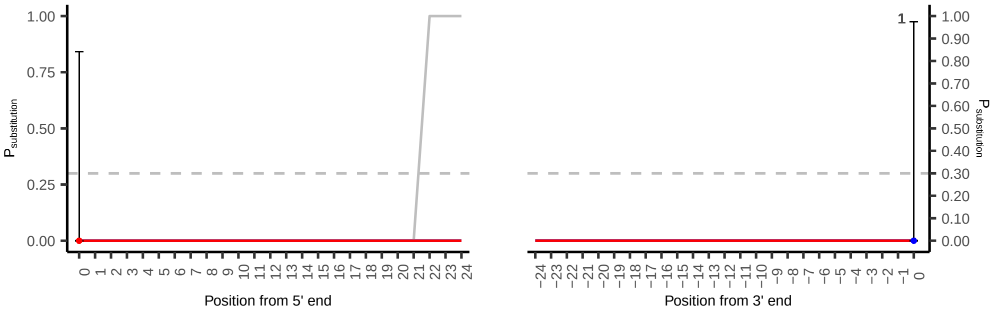
ELF059A_135_Metazoa_no_Primates_ingroup



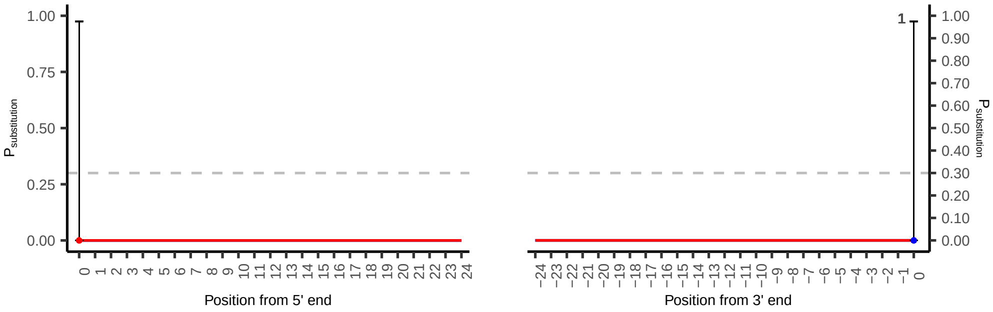
G.4 ELF060

No data for ELF060_465 or 250.

ELF060_420_Metazoa_no_Primates_ingroup

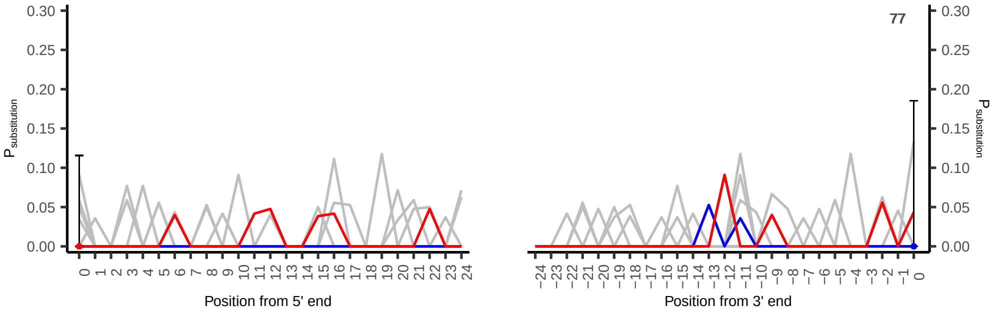


ELF060_350_Metazoa_no_Primates_ingroup

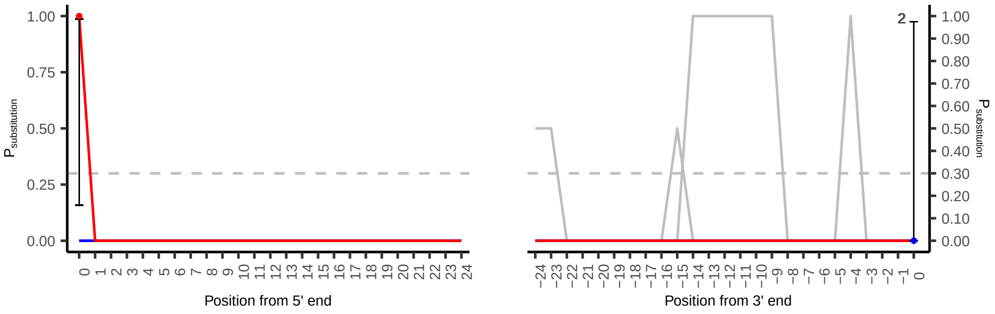


G.5 ELF032A

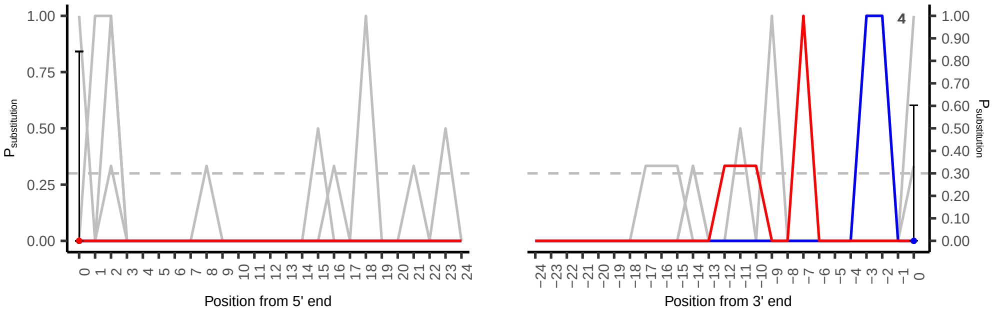
ELF032A_177_Metazoa_no_Primates_ingroup



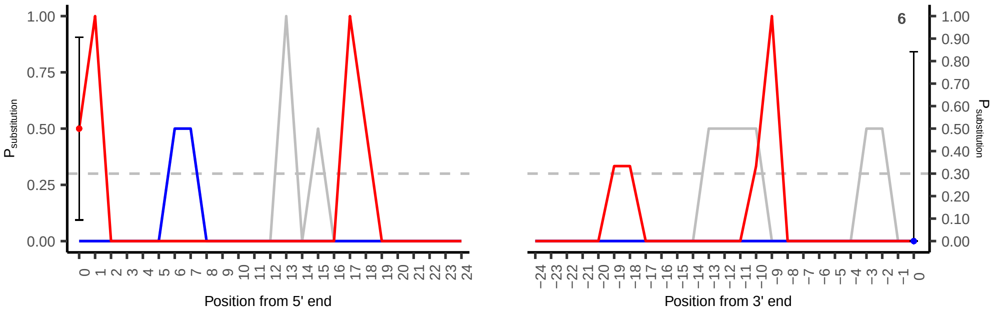
ELF032A_153_Metazoa_no_Primates_ingroup



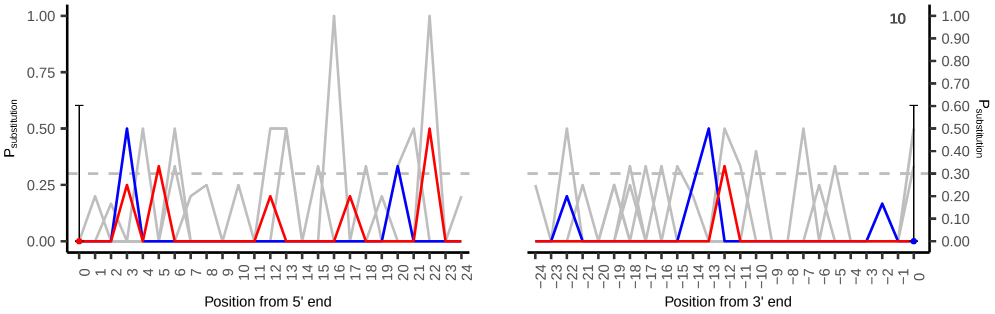
ELF032A_117_Metazoa_no_Primates_ingroup



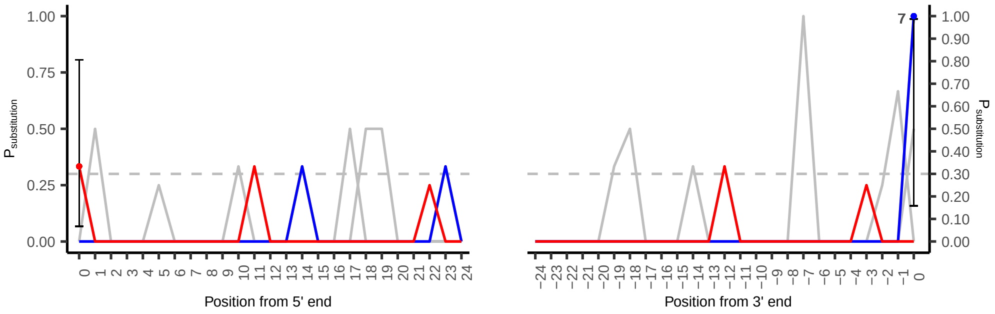
ELF032A_095_Metazoa_no_Primates_ingroup



ELF032A_074_Metazoa_no_Primates_ingroup



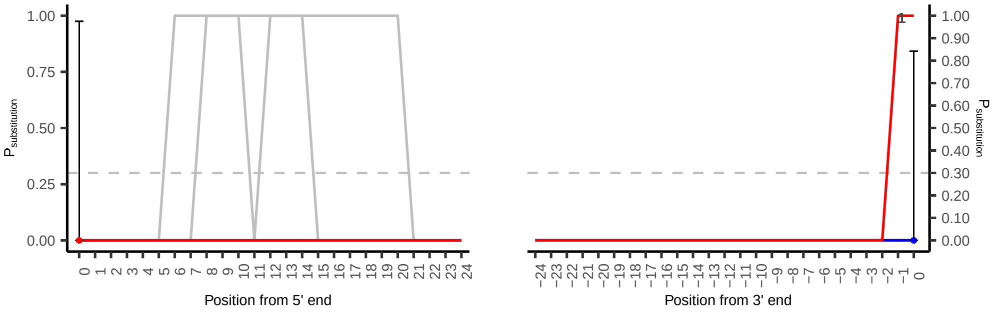
ELF032A_047_Metazoa_no_Primates_ingroup



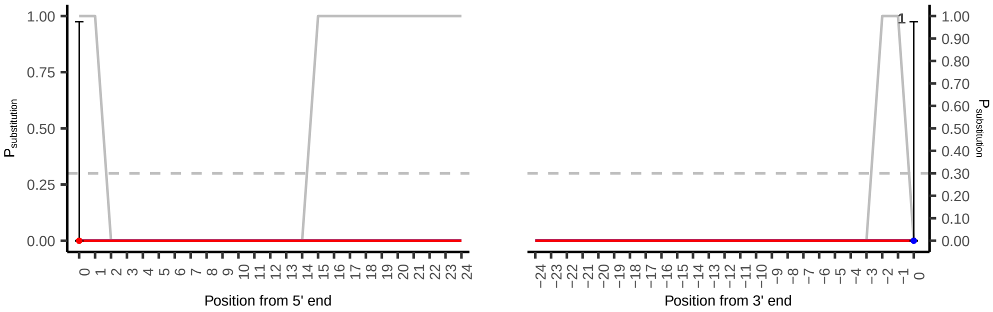
G.6 ELF033

No data for ELF033_203, 195, or 183.

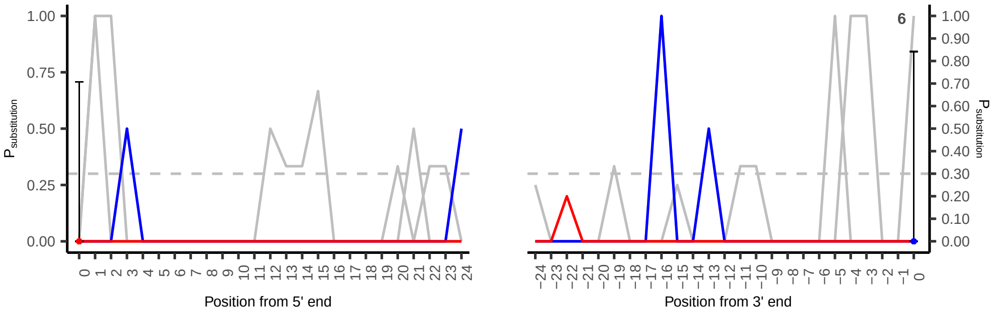
ELF033_187_Metazoa_no_Primates_ingroup



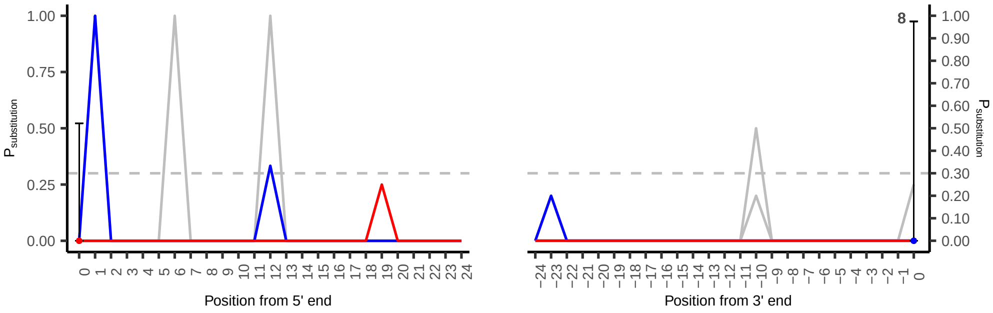
ELF033_176_Metazoa_no_Primates_ingroup



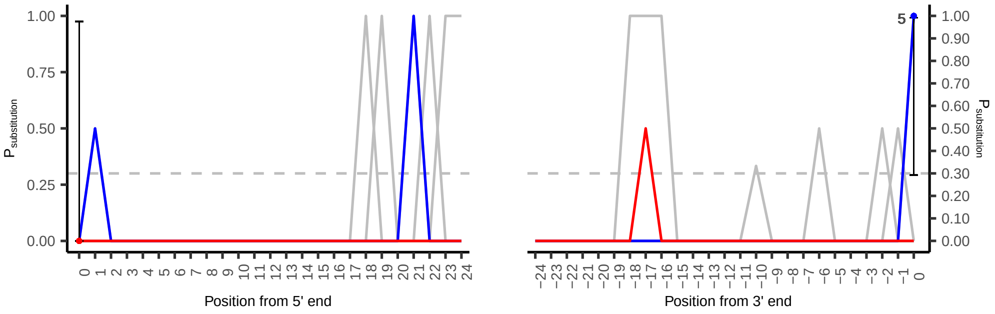
ELF033_155_Metazoa_no_Primates_ingroup



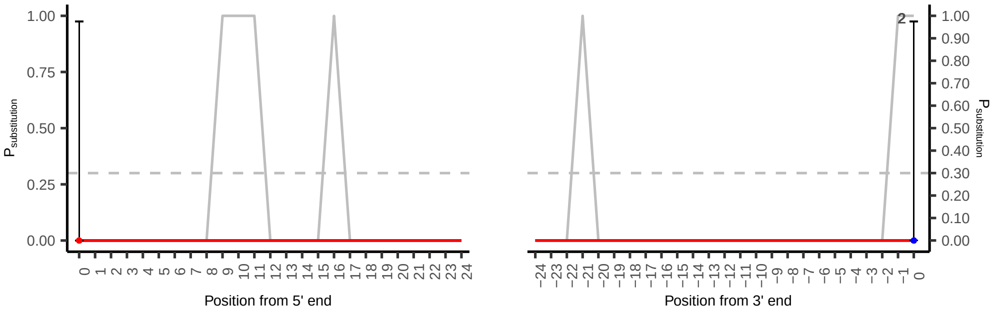
ELF033_118_Metazoa_no_Primates_ingroup



ELF033_075_Metazoa_no_Primates_ingroup



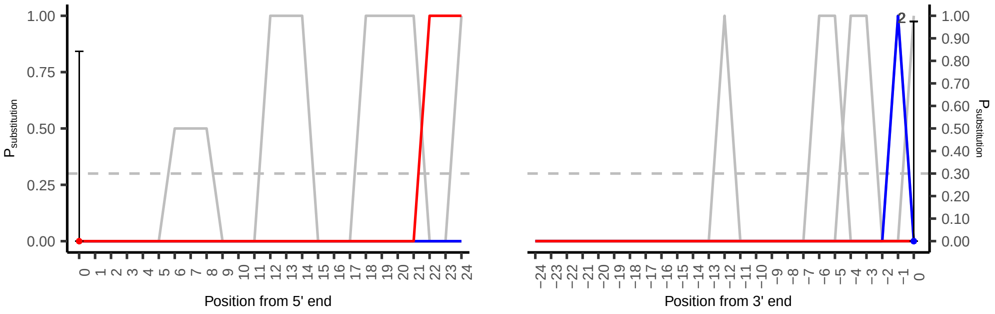
ELF033_046_Metazoa_no_Primates_ingroup



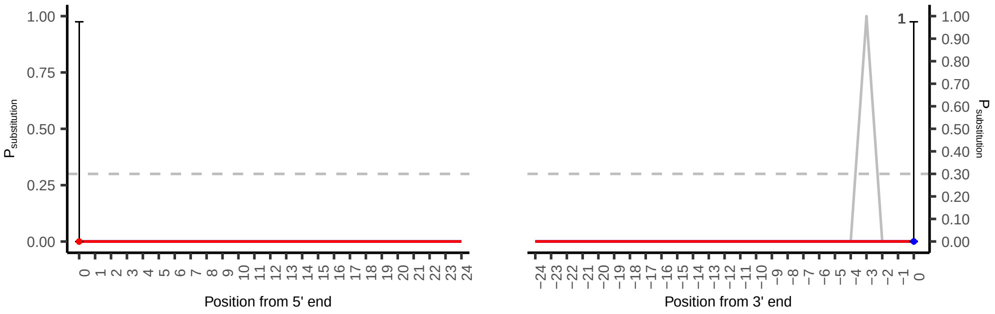
G.7 ELF033A

No data for ELF033A_126 or 070.

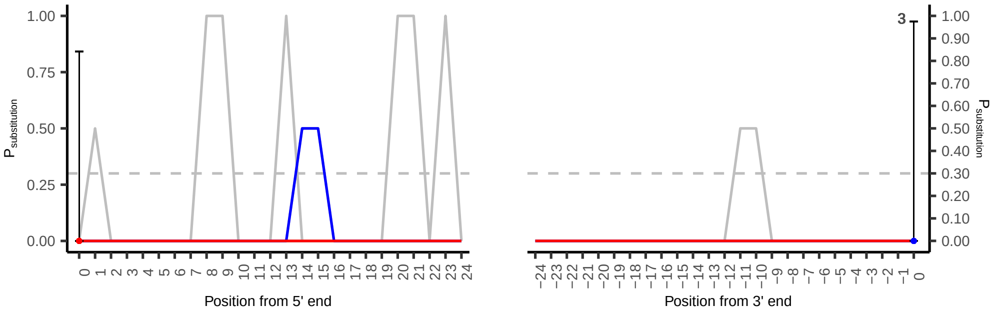
ELF033A_158_Metazoa_no_Primates_ingroup



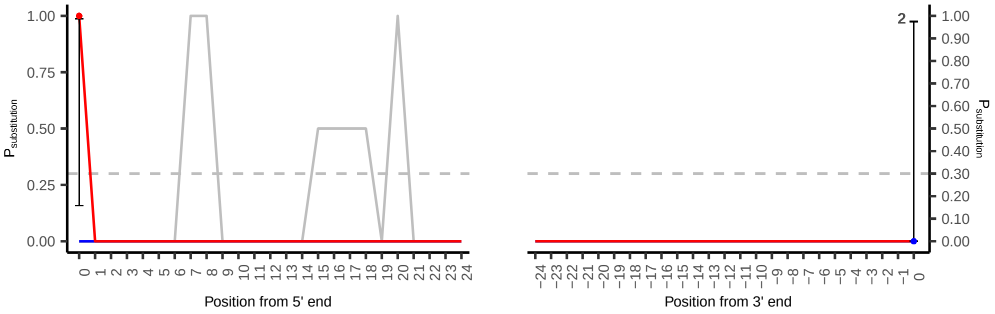
ELF033A_111_Metazoa_no_Primates_ingroup



ELF033A_098_Metazoa_no_Primates_ingroup



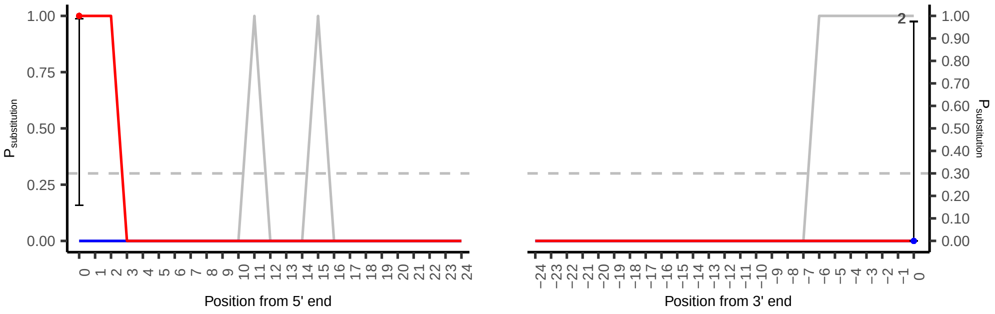
ELF033A_050_Metazoa_no_Primates_ingroup



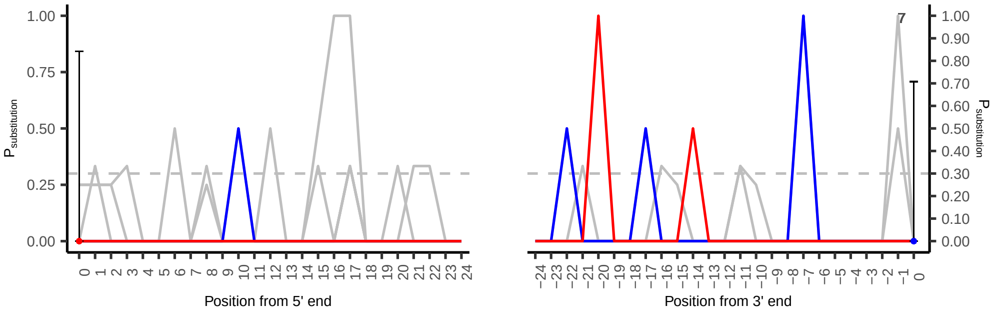
G.8 ELF034A

No data for ELF034A_282, 225, 261, 195, 183, 172, or 166.

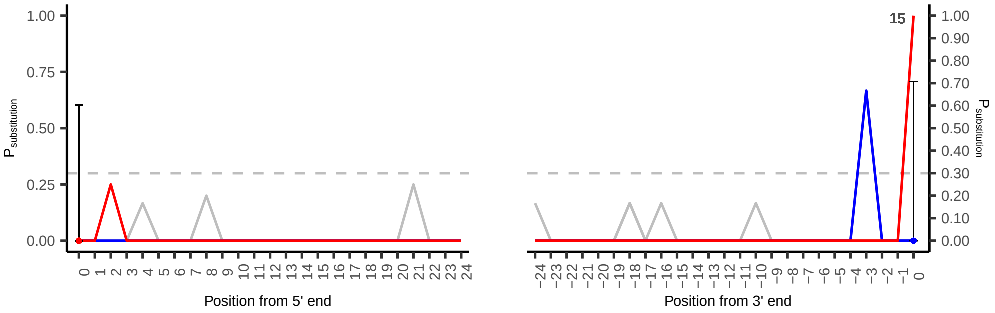
ELF034A_146_Metazoa_no_Primates_ingroup



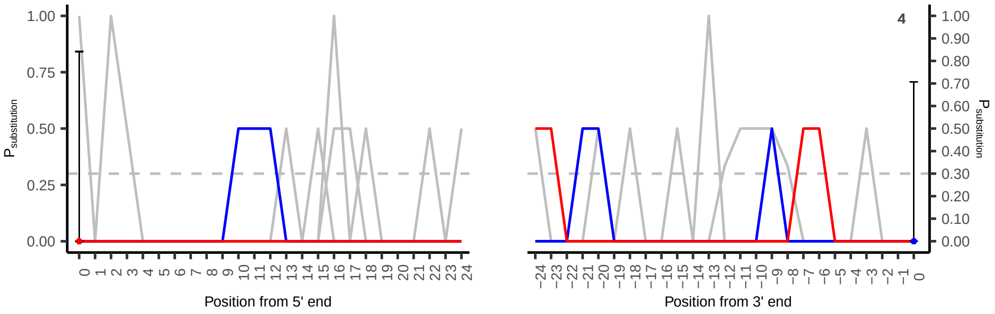
ELF034A_126_Metazoa_no_Primates_ingroup



ELF034A_081_Metazoa_no_Primates_ingroup



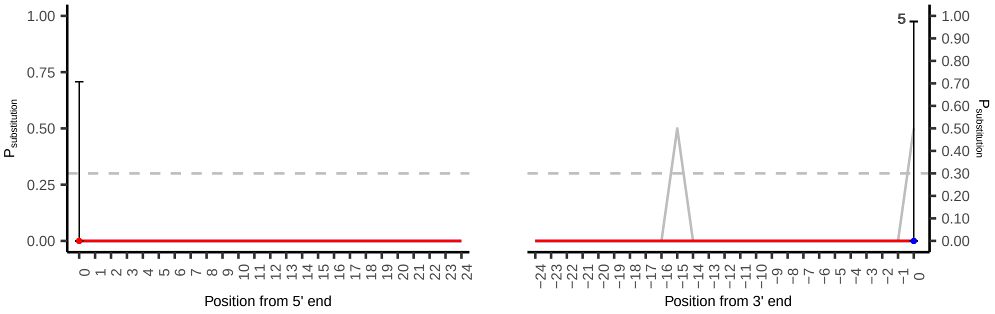
ELF034A_063_Metazoa_no_Primates_ingroup



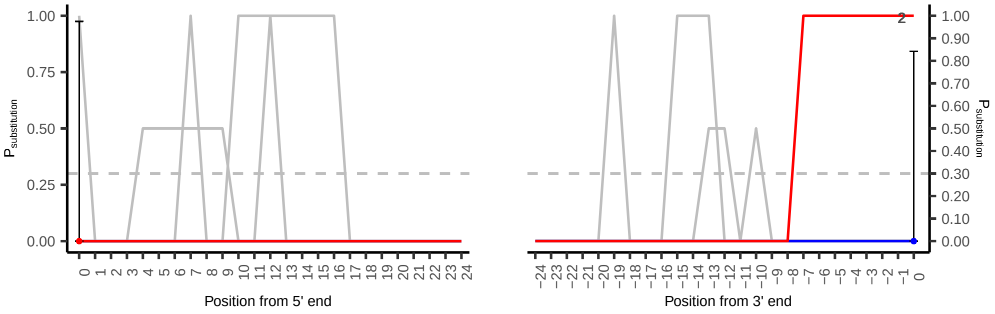
G.9 ELF034

No data for ELF034_185, 177, 079, or 061.

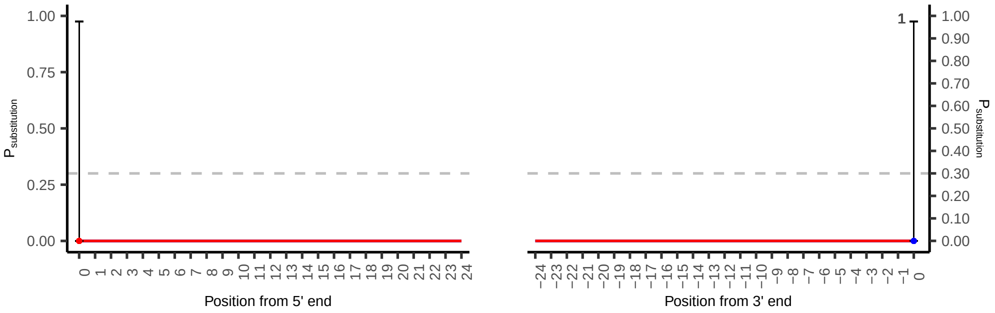
ELF034_219_Metazoa_no_Primates_ingroup



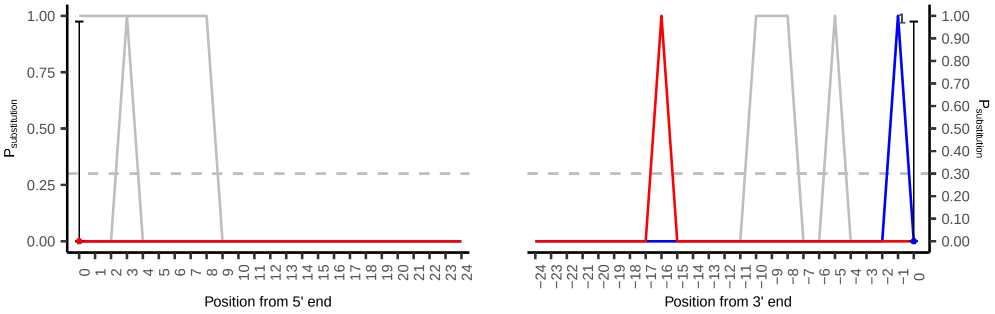
ELF034_202_Metazoa_no_Primates_ingroup



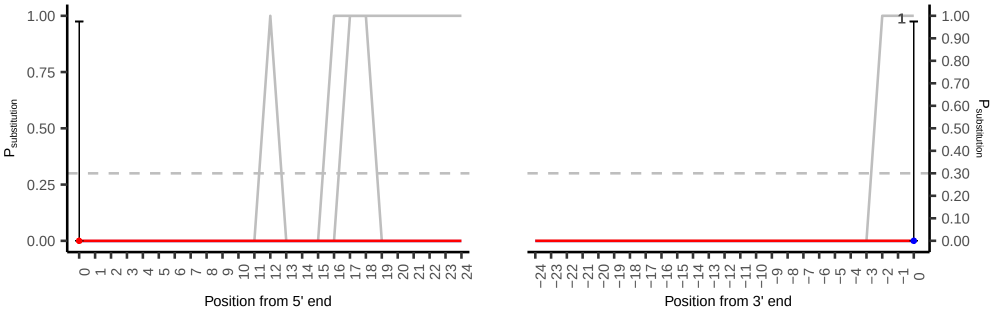
ELF034_157_Metazoa_no_Primates_ingroup



ELF034_132_Metazoa_no_Primates_ingroup



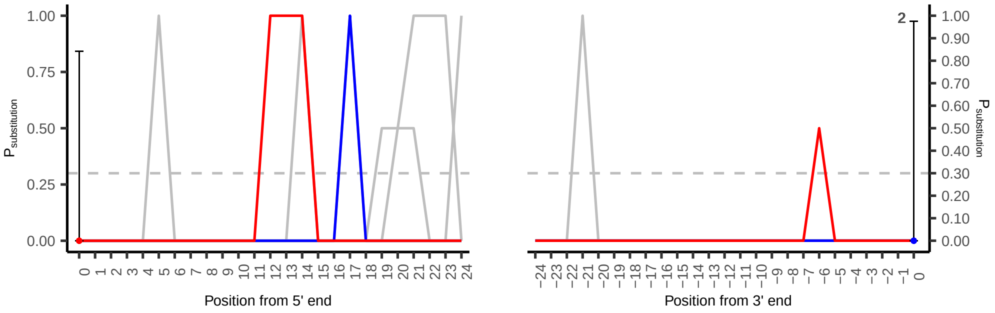
ELF034_094_Metazoa_no_Primates_ingroup



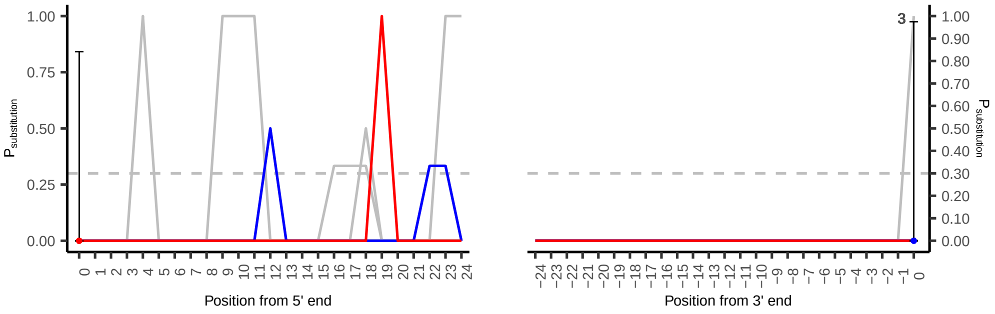
G.10 ELF054

No data from ELF054_315.

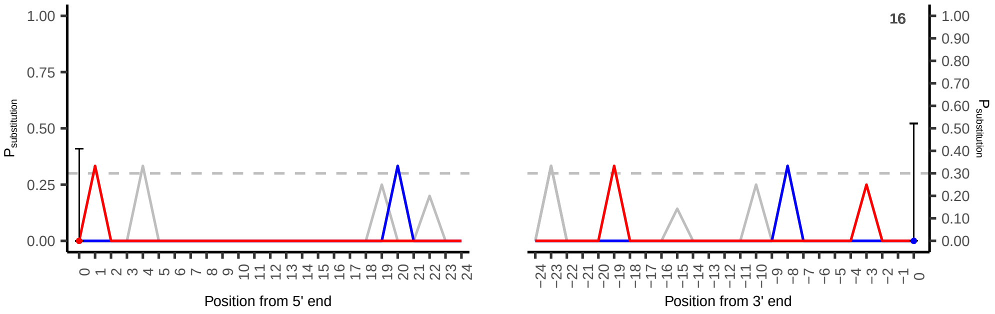
ELF054_356_Metazoa_no_Primates_ingroup



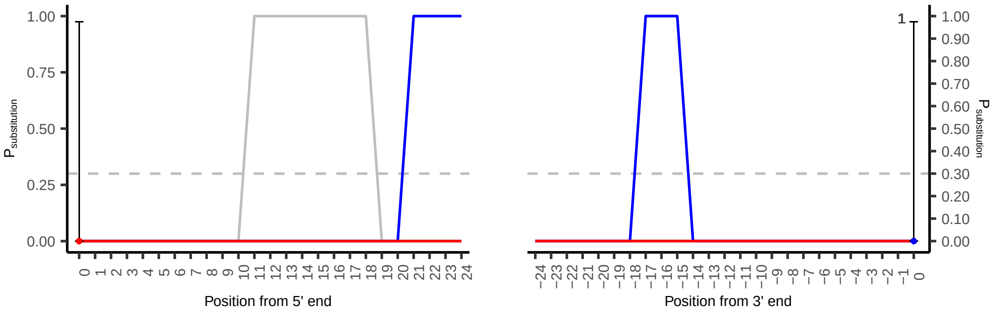
ELF054_330_Metazoa_no_Primates_ingroup



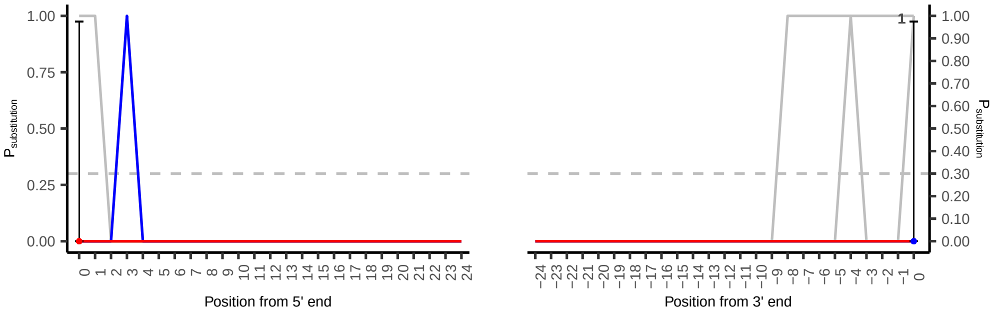
ELF054_291_Metazoa_no_Primates_ingroup



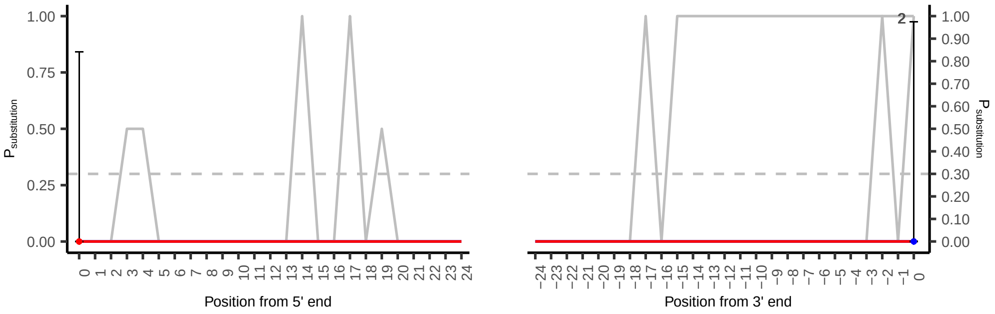
ELF054_268_Metazoa_no_Primates_ingroup

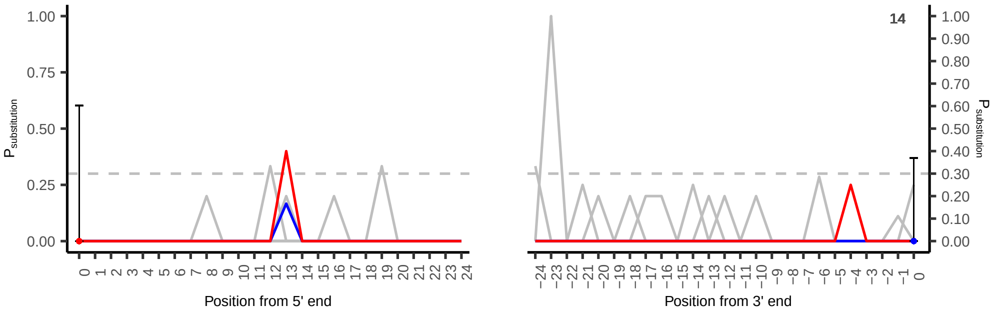


ELF054_182_Metazoa_no_Primates_ingroup



ELF054_140_Metazoa_no_Primates_ingroup

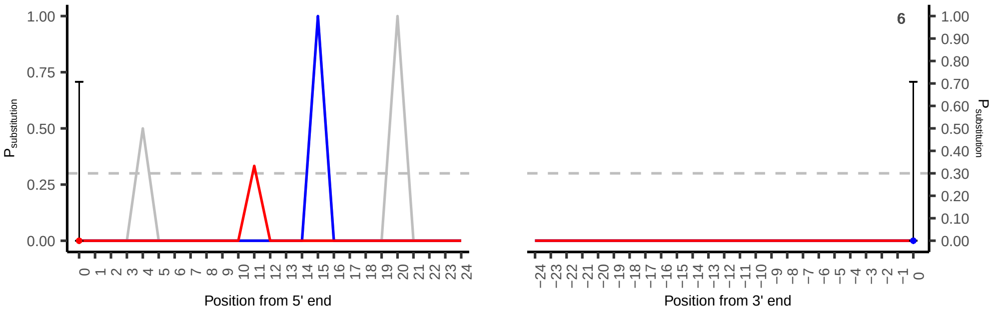



ELF054_058_Metazoa_no_Primates_ingroup

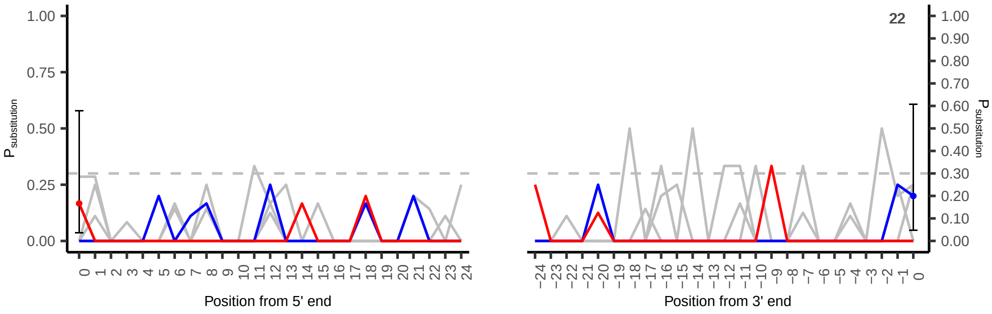
G.11 ELF039

No data for ELF039_145.

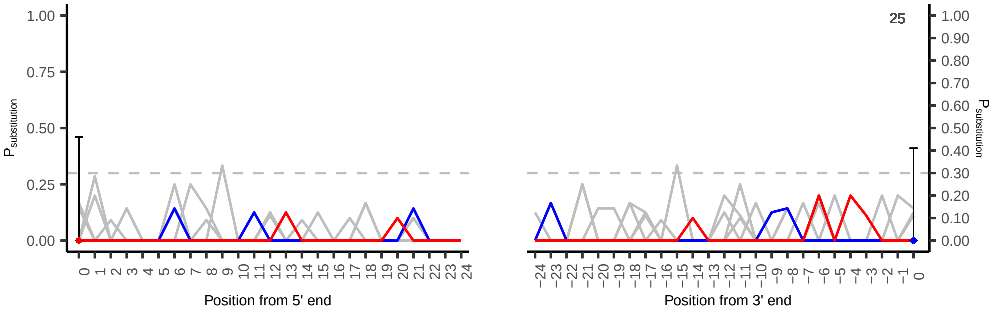
ELF039_485_Metazoa_no_Primates_ingroup



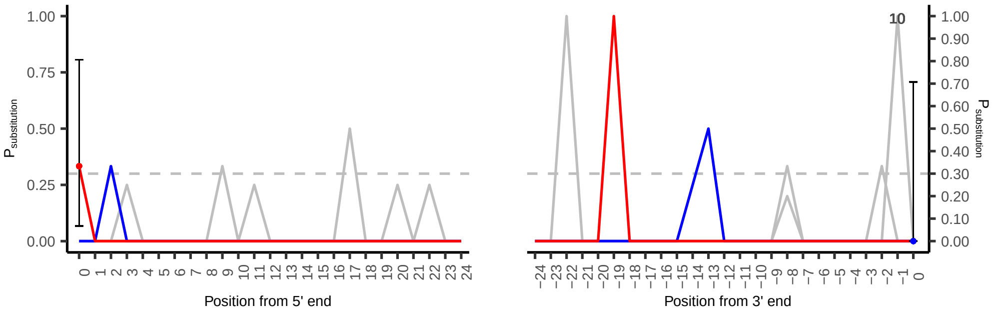
ELF039_460_Metazoa_no_Primates_ingroup



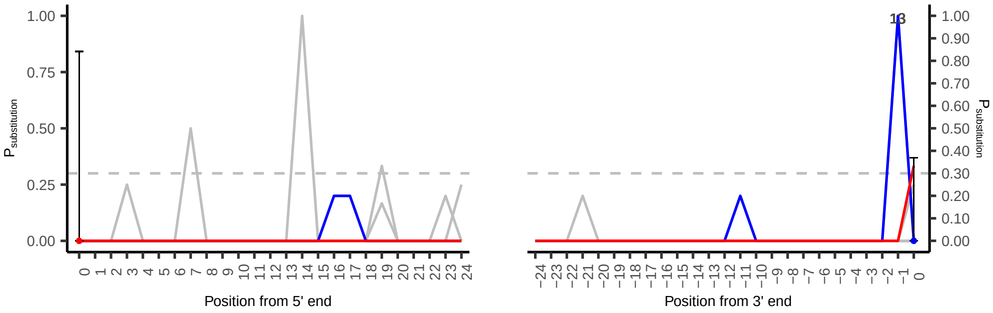
ELF039_415_Metazoa_no_Primates_ingroup



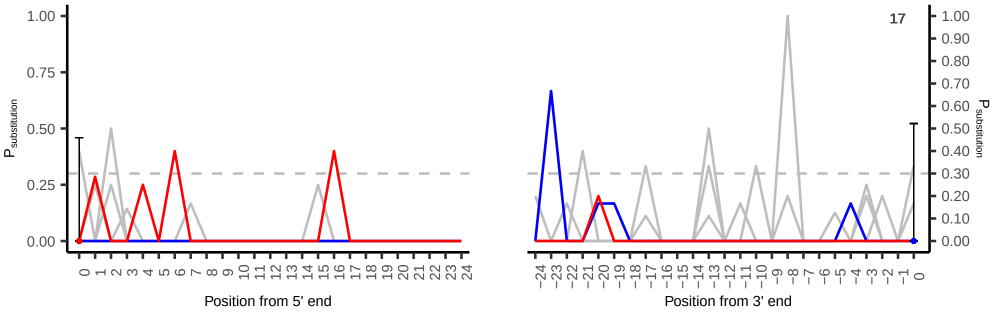
ELF039_384_Metazoa_no_Primates_ingroup



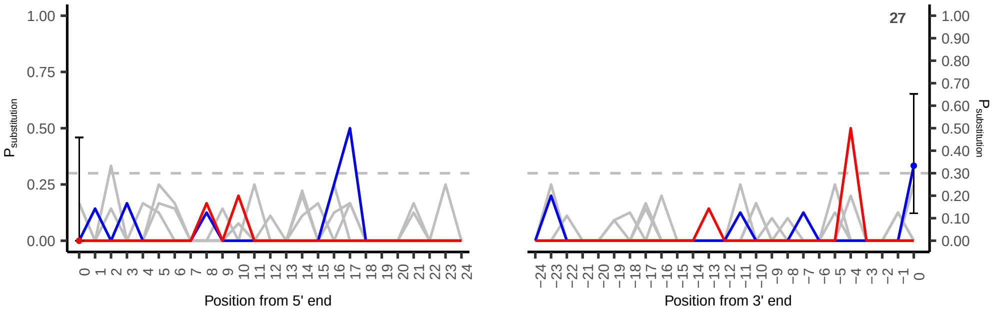
ELF039_355_Metazoa_no_Primates_ingroup



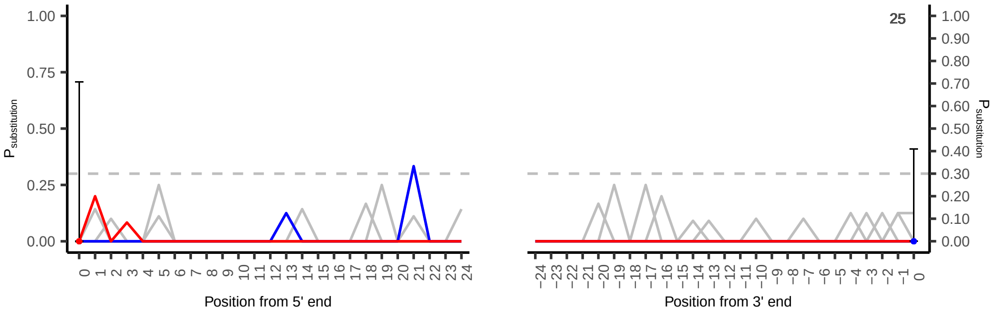
ELF039_341_Metazoa_no_Primates_ingroup



ELF039_321_Metazoa_no_Primates_ingroup



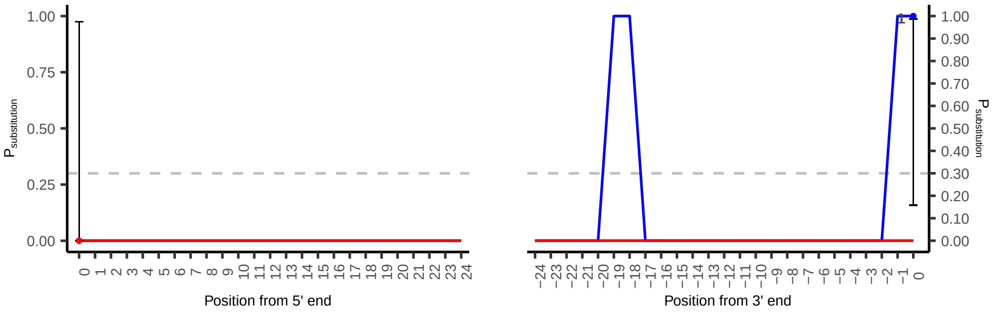
ELF039_250_Metazoa_no_Primates_ingroup



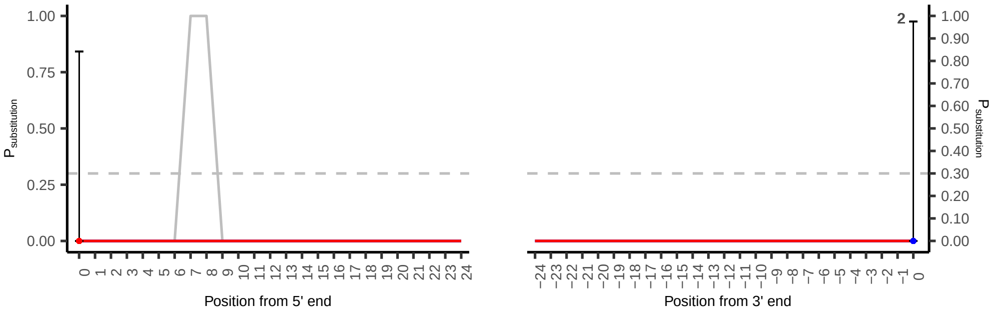
G.12 ELF040A

No data from ELF040A_487 or 350.

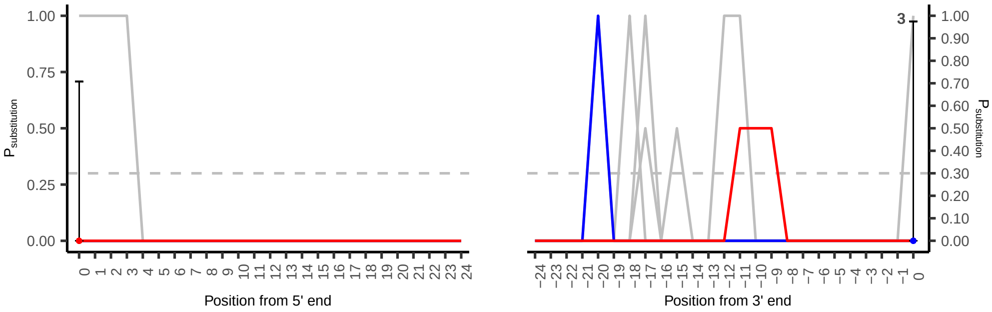
ELF040A_298_Metazoa_no_Primates_ingroup



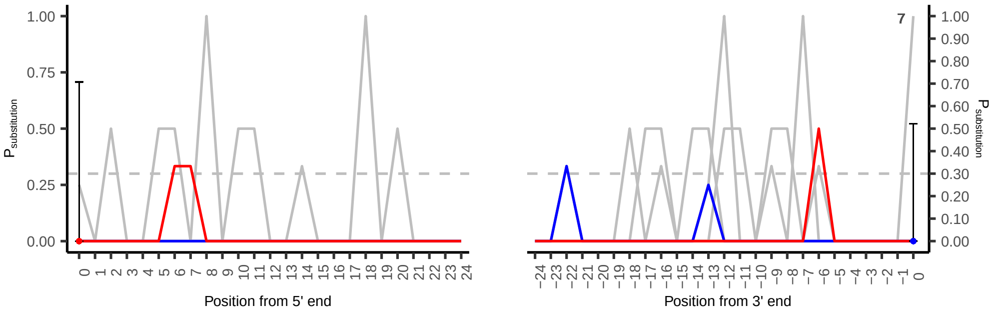
ELF040A_208_Metazoa_no_Primates_ingroup



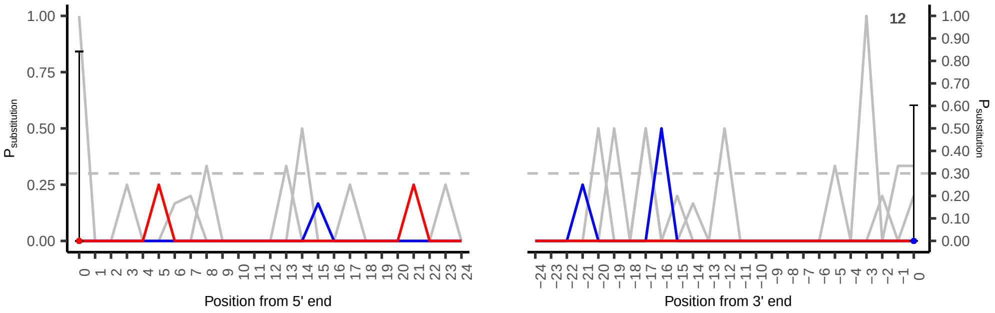
ELF040A_192_Metazoa_no_Primates_ingroup



ELF040A_112_Metazoa_no_Primates_ingroup



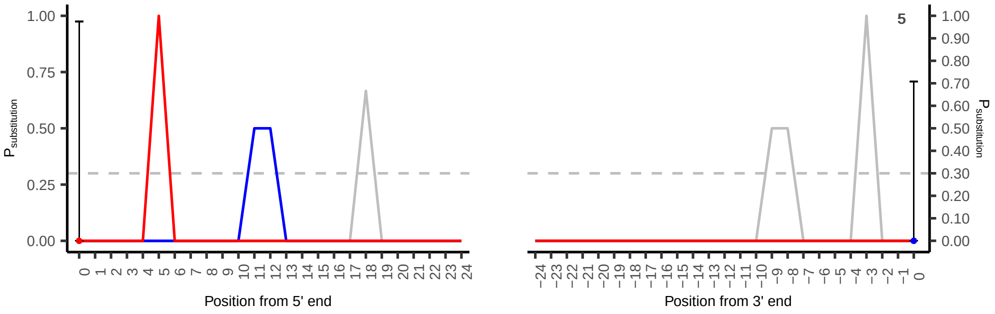
ELF040A_095_Metazoa_no_Primates_ingroup



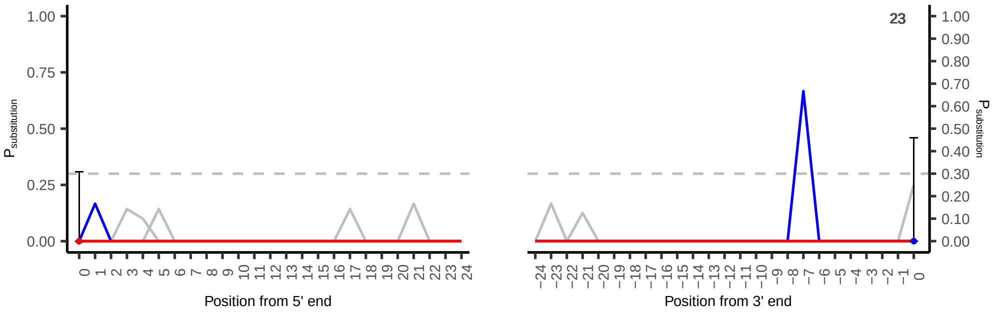
G.13 ELF041

No data from ELF041_295.

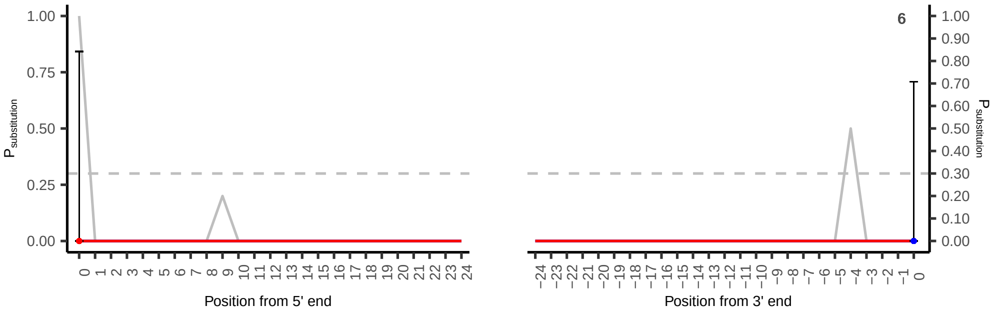
ELF041_180_Metazoa_no_Primates_ingroup



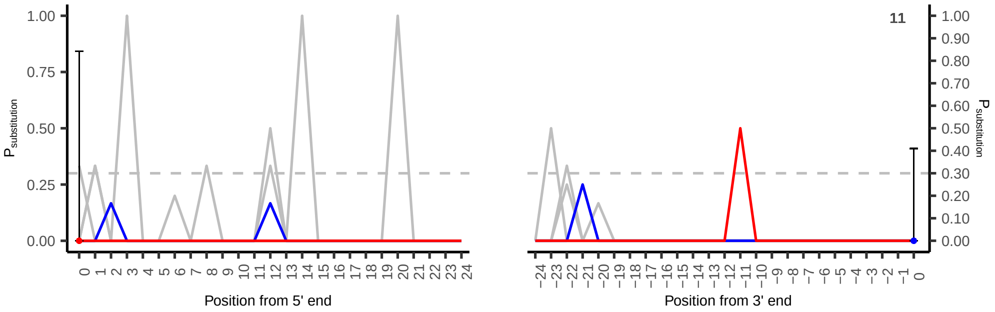
ELF041_110_Metazoa_no_Primates_ingroup

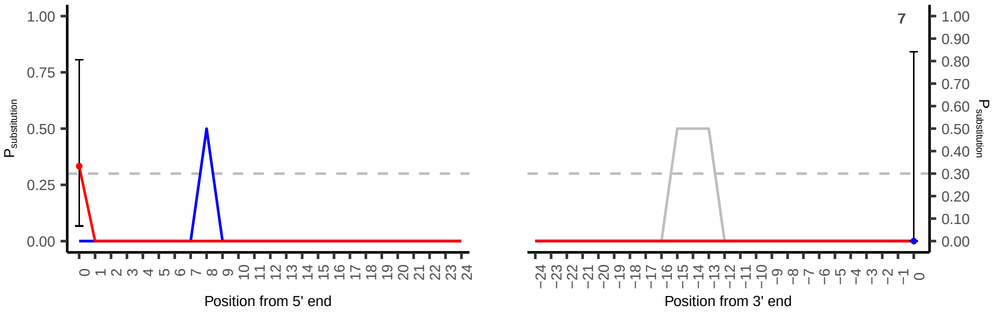


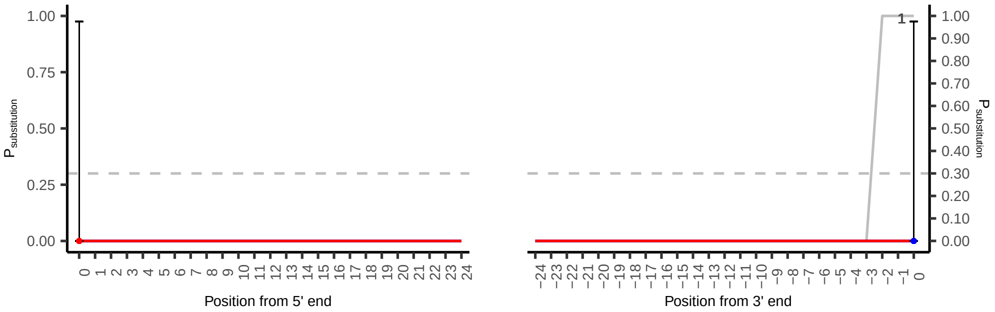
ELF041_087_Metazoa_no_Primates_ingroup

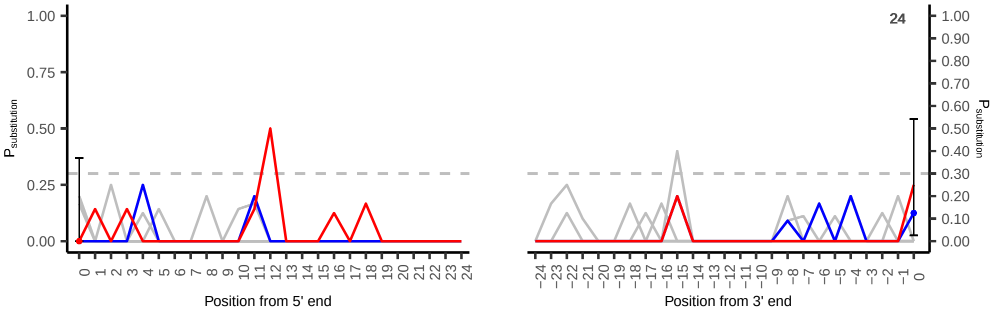


G.14 ELF042


ELF042_350_Metazoa_no_Primates_ingroup

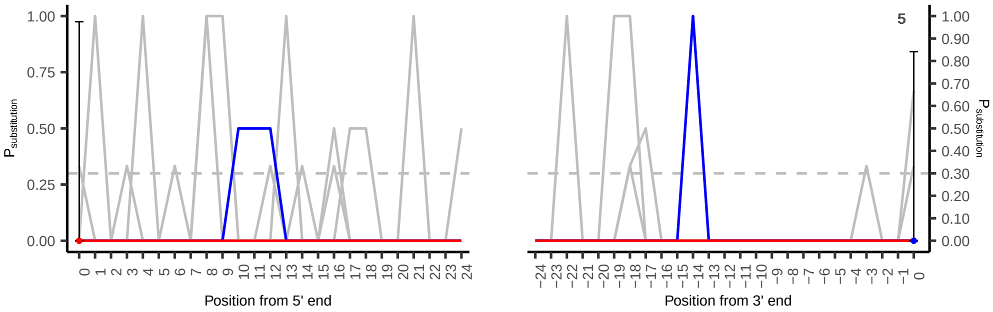

ELF042_250_Metazoa_no_Primates_ingroup


ELF042_151_Metazoa_no_Primates_ingroup


ELF042_065_Metazoa_no_Primates_ingroup

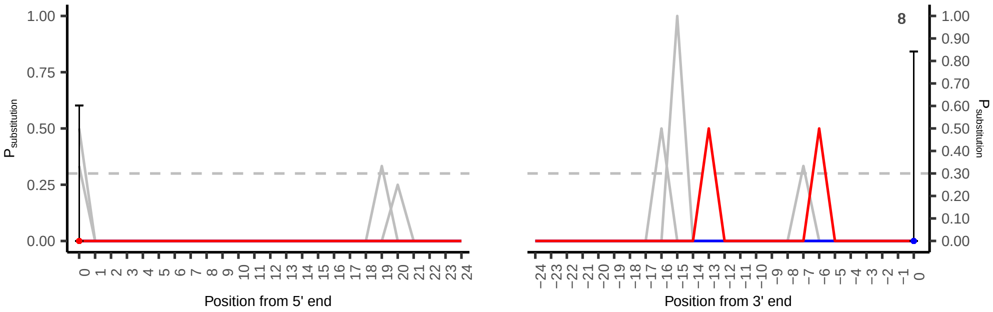
G.15 ELF044

No data from ELF044_137.



ELF044_090_Metazoa_no_Primates_ingroup

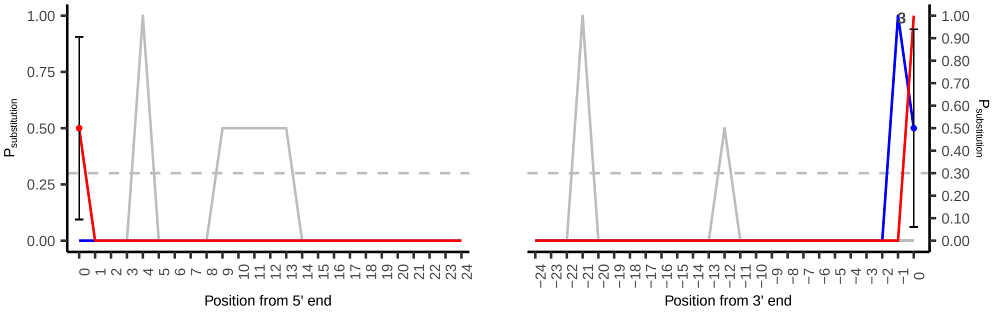
G.16 ELF044A



ELF044A_097_Metazoa_no_Primates_ingroup

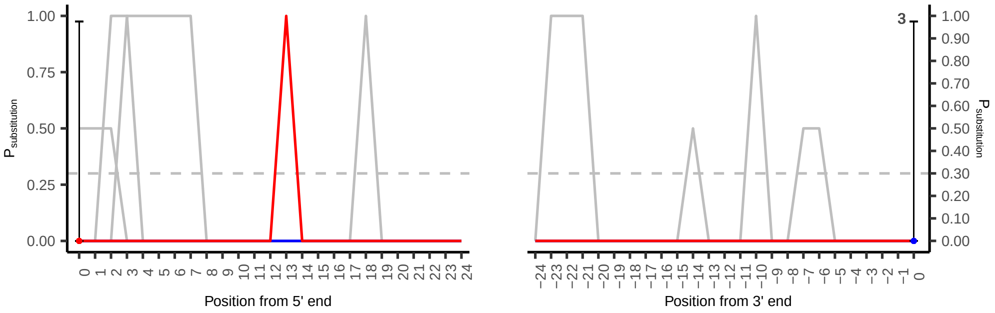
G.17 ELF031

No data for ELF031_033.



ELF031_056_Metazoa_no_Primates_ingroup

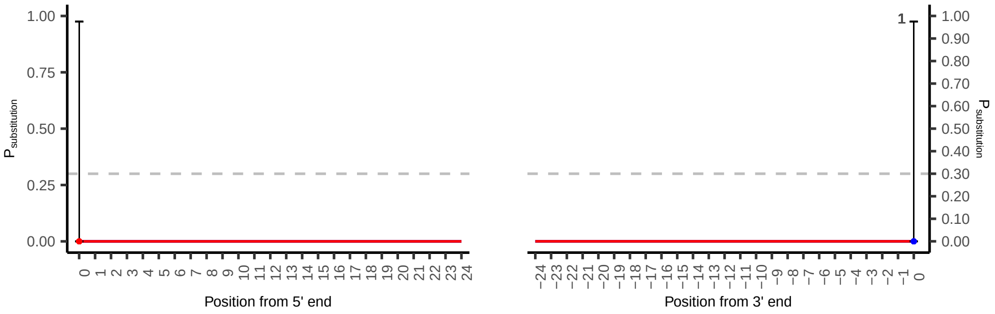
ELF031_043_Metazoa_no_Primates_ingroup



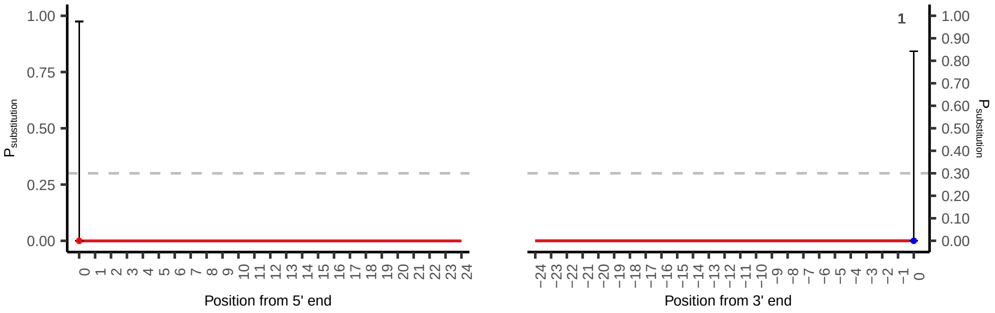
G.18 ELF031A

No data for ELF031A_310, 219, or 088.

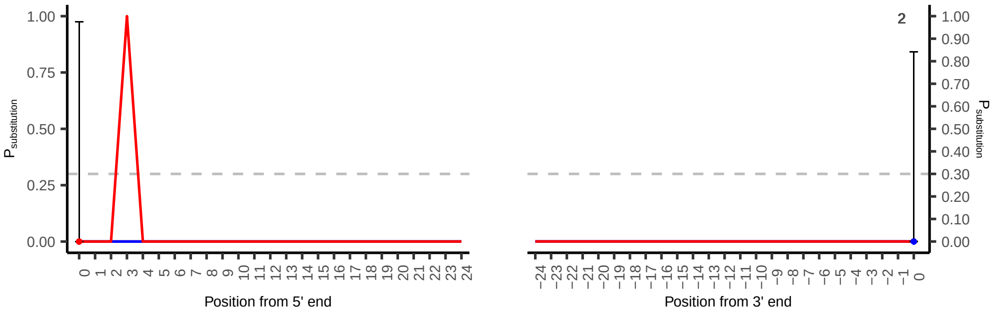
ELF031A_281_Metazoa_no_Primates_ingroup



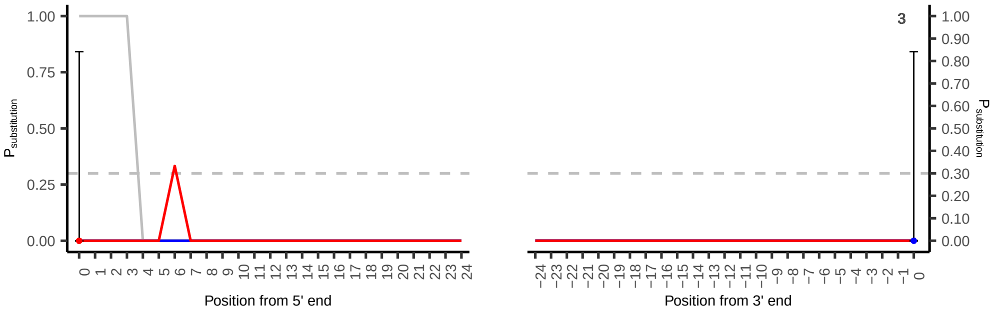
ELF031A_202_Metazoa_no_Primates_ingroup



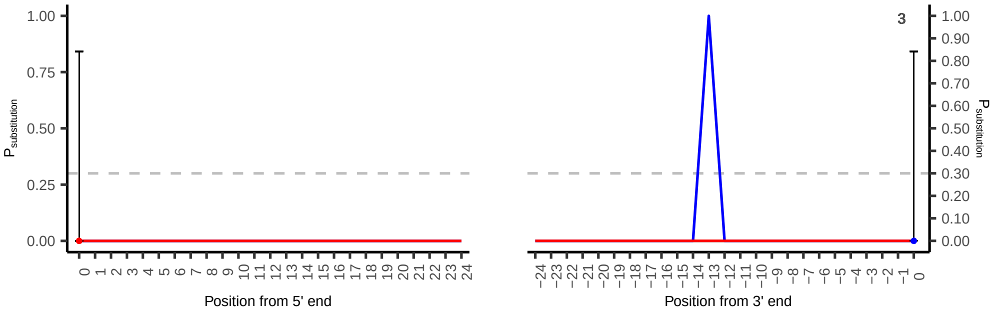
ELF031A_177_Metazoa_no_Primates_ingroup



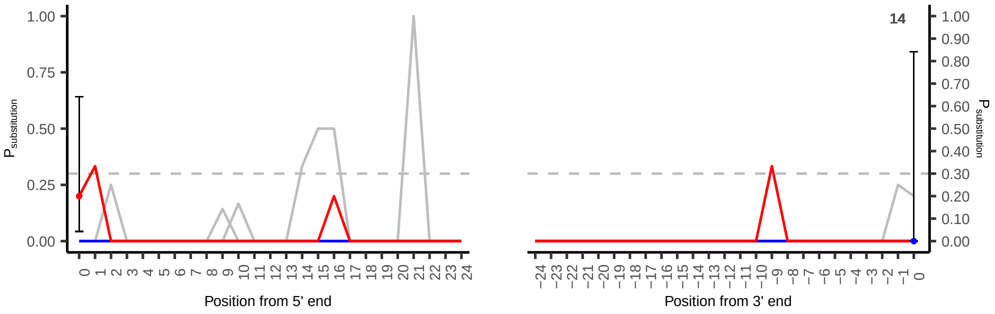
ELF031A_152_Metazoa_no_Primates_ingroup



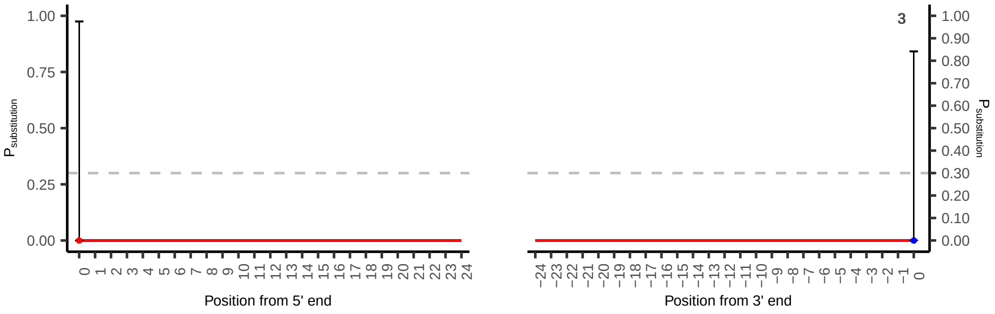
ELF031A_123_Metazoa_no_Primates_ingroup



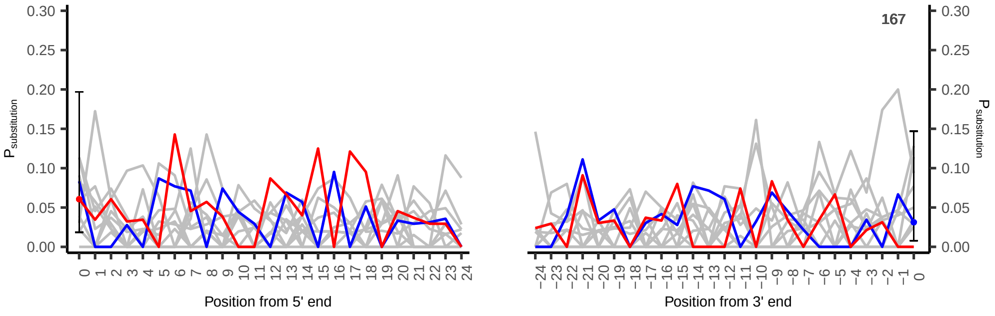
ELF031A_107_Metazoa_no_Primates_ingroup



ELF031A_072_Metazoa_no_Primates_ingroup

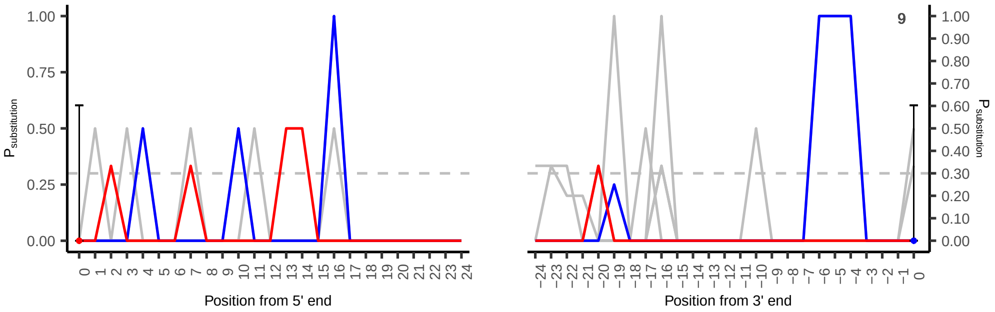


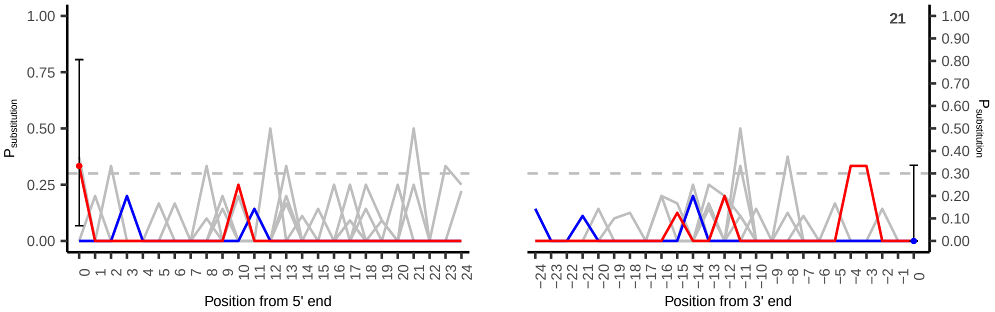
ELF031A_058_Metazoa_no_Primates_ingroup

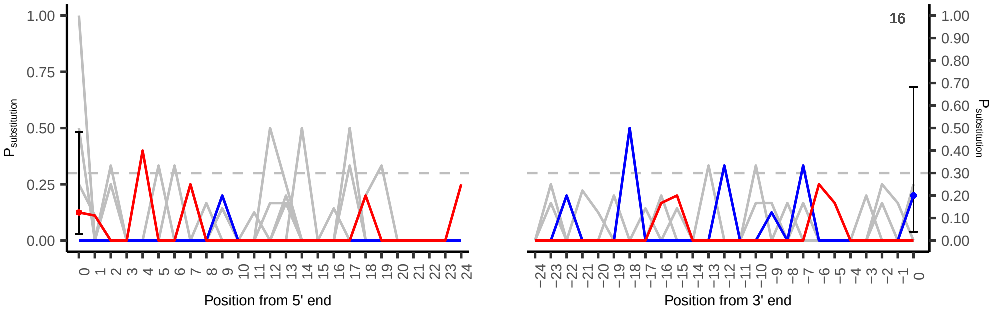


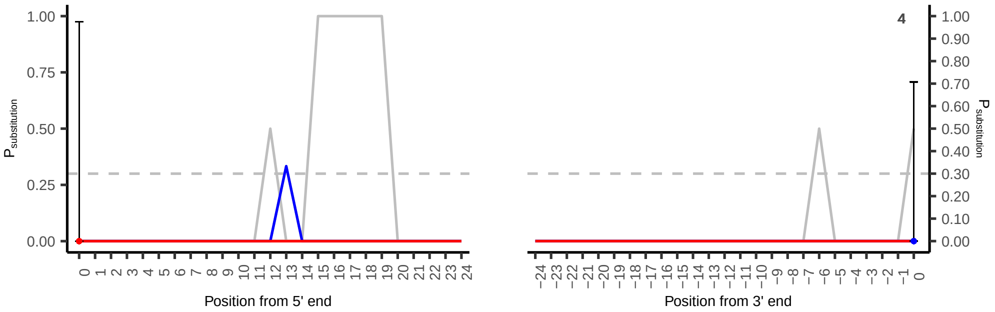
G.19 ELF051

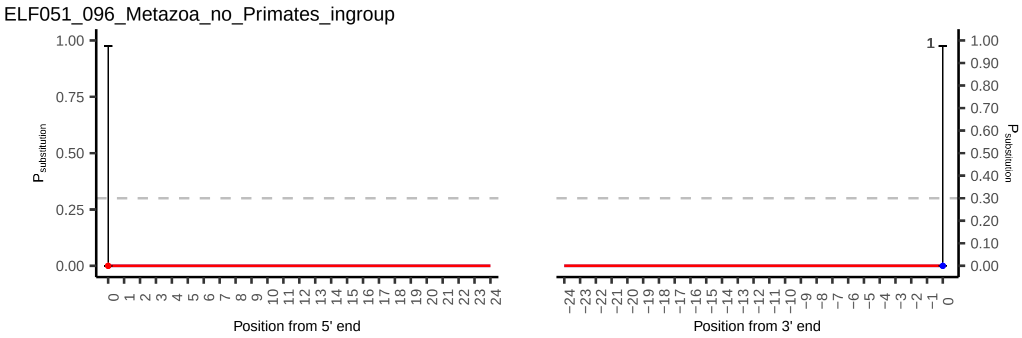
No data from ELF051_120.


ELF051_292_Metazoa_no_Primates_ingroup

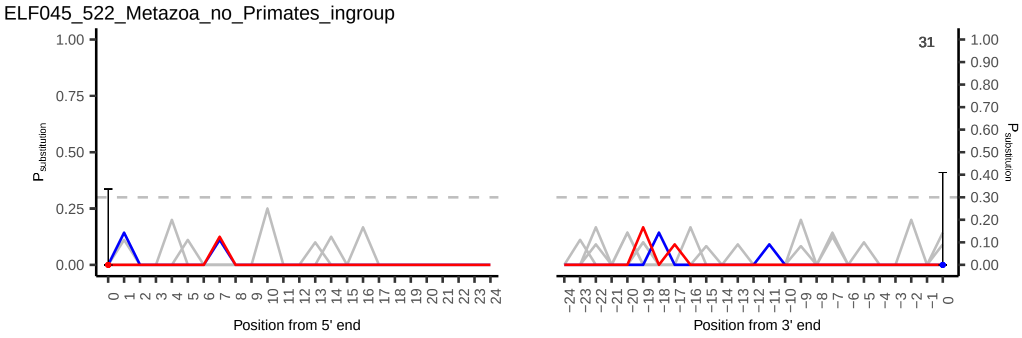

ELF051_255_Metazoa_no_Primates_ingroup

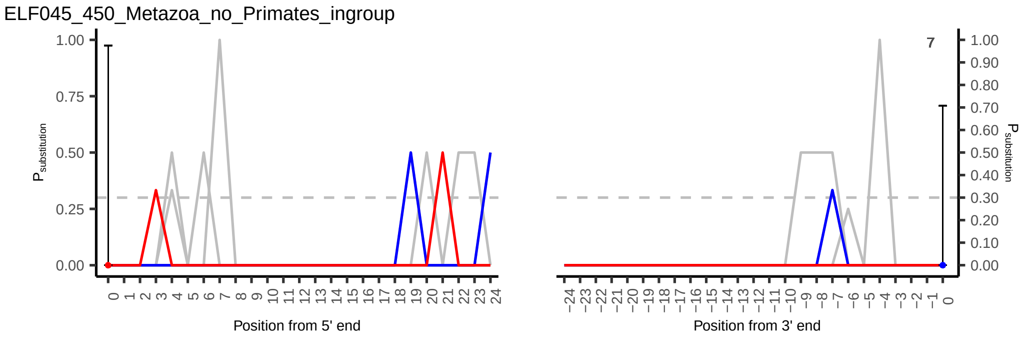

ELF051_196_Metazoa_no_Primates_ingroup

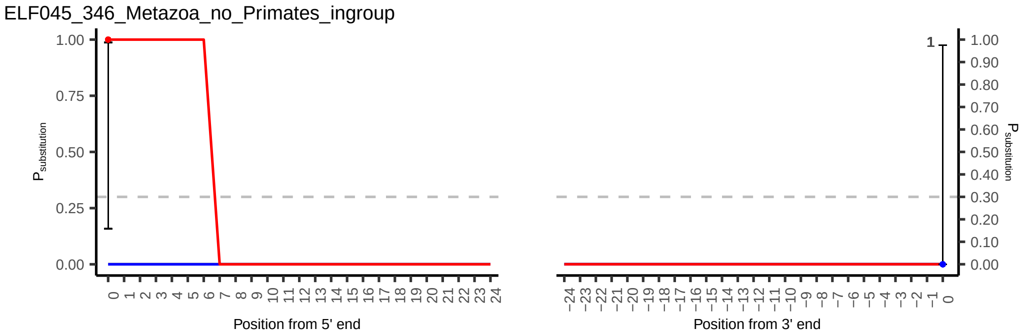

ELF051_151_Metazoa_no_Primates_ingroup


ELF051_096_Metazoa_no_Primates_ingroup
1
P substitution
1.00
0.75
0.50
0.25
0.00
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

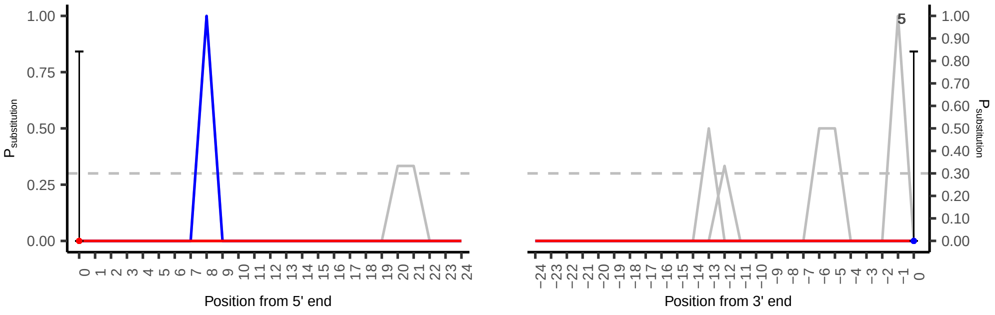
G.20 ELF045


ELF045_522_Metazoa_no_Primates_ingroup
31
P substitution
1.00
0.75
0.50
0.25
0.00
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

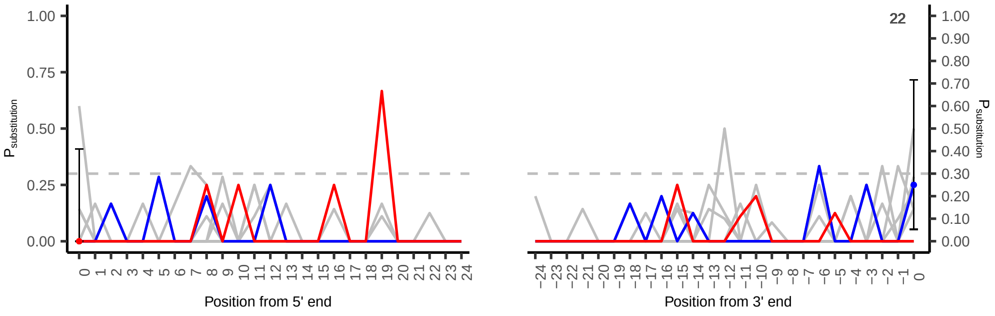

ELF045_450_Metazoa_no_Primates_ingroup
7
P substitution
1.00
0.75
0.50
0.25
0.00
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end


ELF045_346_Metazoa_no_Primates_ingroup
1
P substitution
1.00
0.75
0.50
0.25
0.00
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end

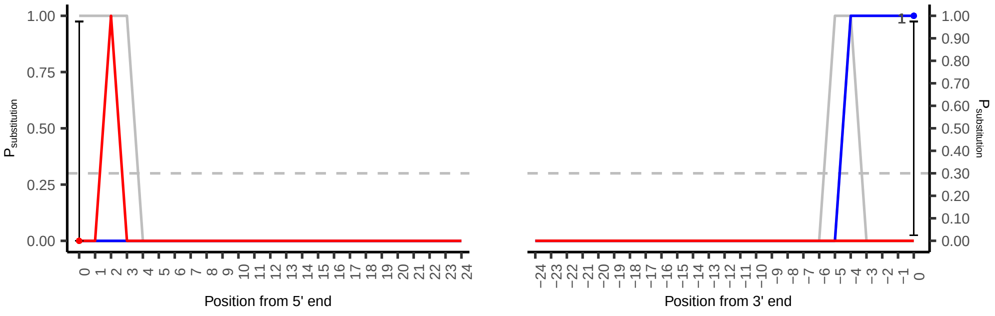
ELF045_252_Metazoa_no_Primates_ingroup



ELF045_145_Metazoa_no_Primates_ingroup



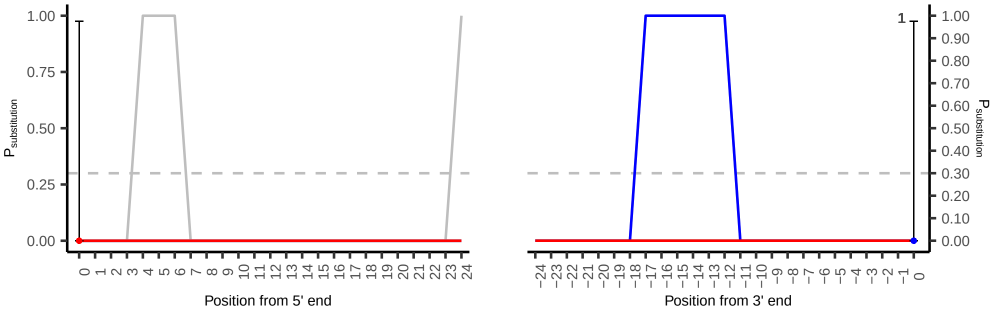
ELF045_090_Metazoa_no_Primates_ingroup



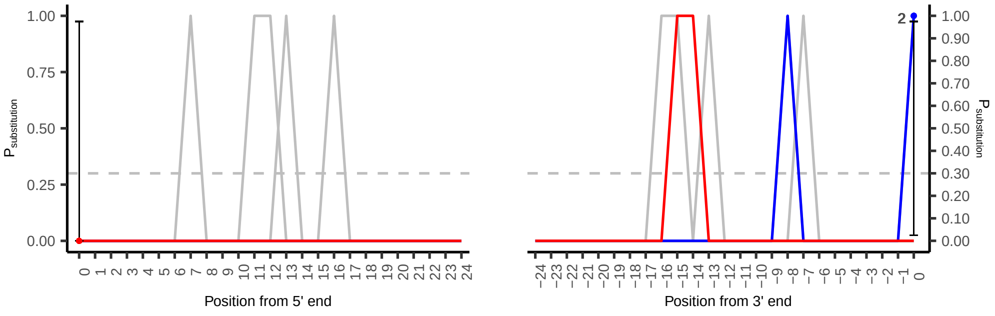
G.21 ELF047

No data from ELF047_241

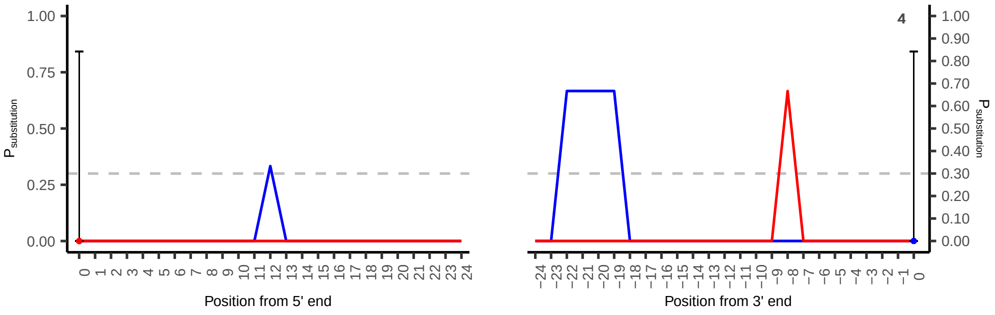
ELF047_386_Metazoa_no_Primates_ingroup



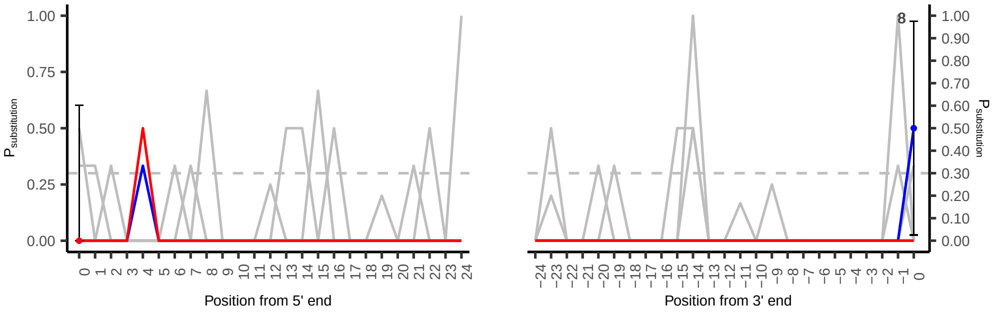
ELF047_325_Metazoa_no_Primates_ingroup



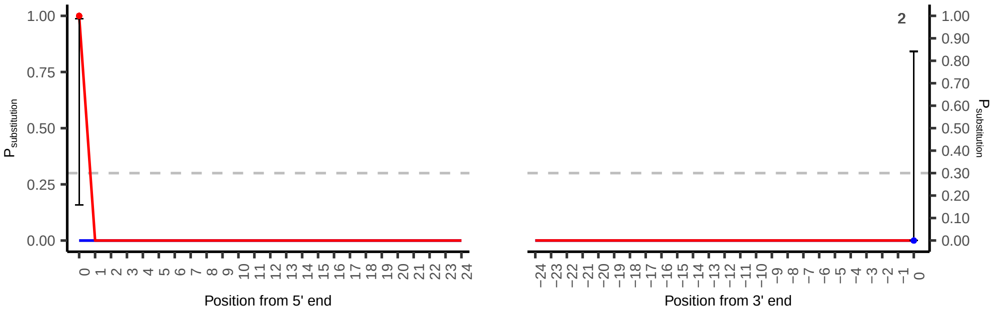
ELF047_274_Metazoa_no_Primates_ingroup



ELF047_150_Metazoa_no_Primates_ingroup

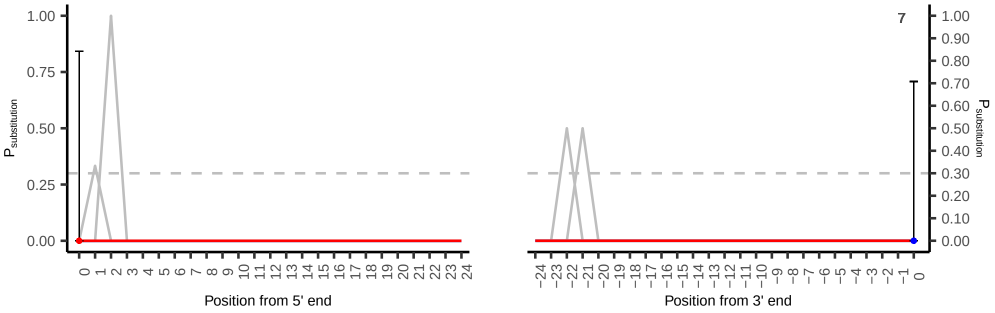


ELF047_070_Metazoa_no_Primates_ingroup

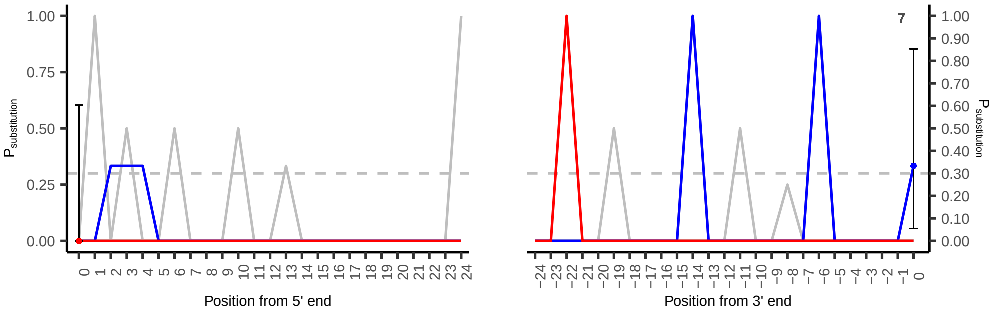


G.22 ELF047A

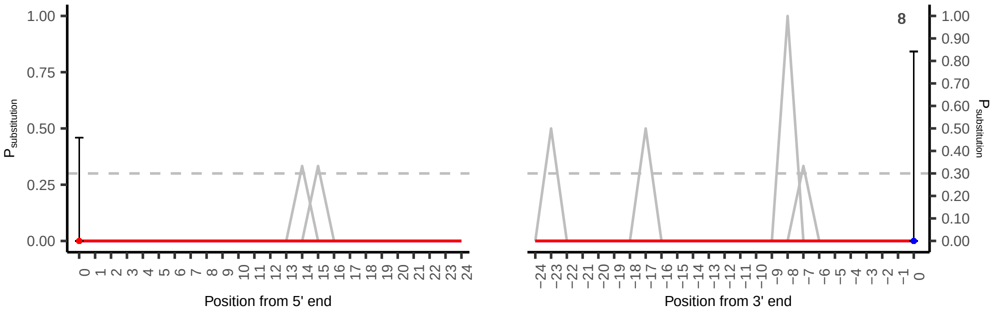
ELF047A_354_Metazoa_no_Primates_ingroup



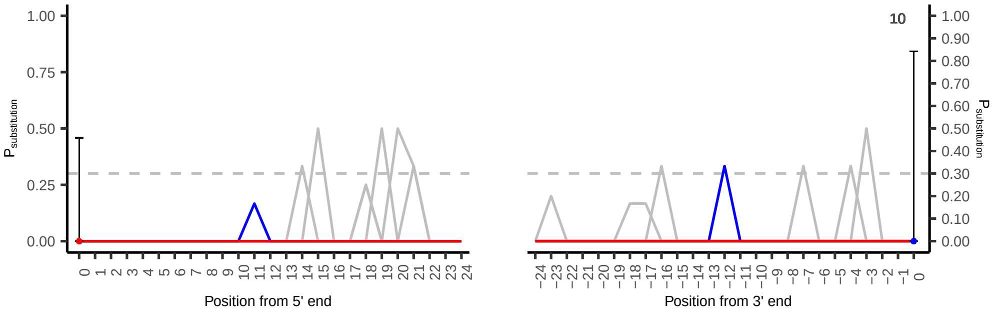
ELF047A_256_Metazoa_no_Primates_ingroup



ELF047A_150_Metazoa_no_Primates_ingroup

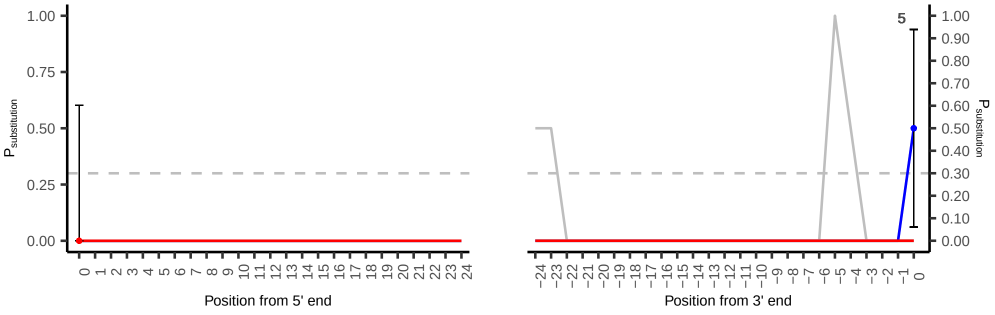


ELF047A_050_Metazoa_no_Primates_ingroup

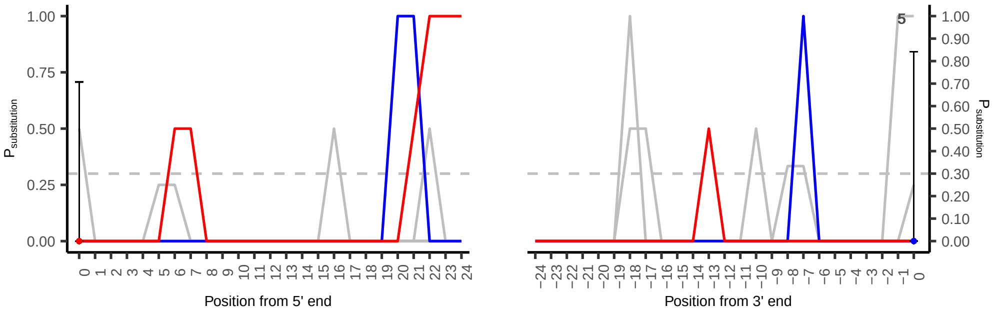


G.23 ELF049

ELF049_361_Metazoa_no_Primates_ingroup

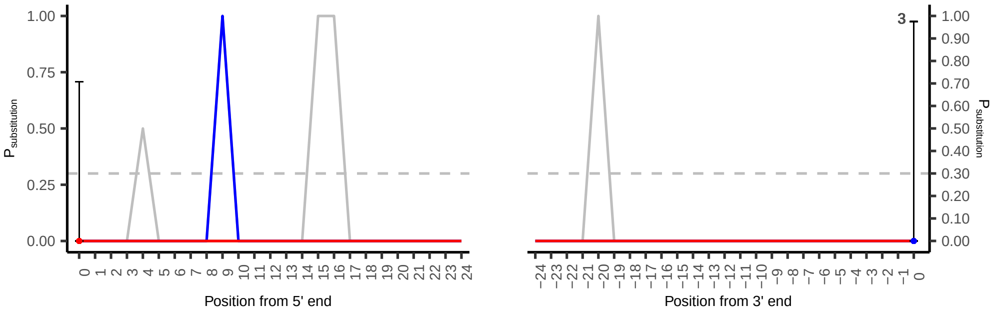


ELF049_295_Metazoa_no_Primates_ingroup



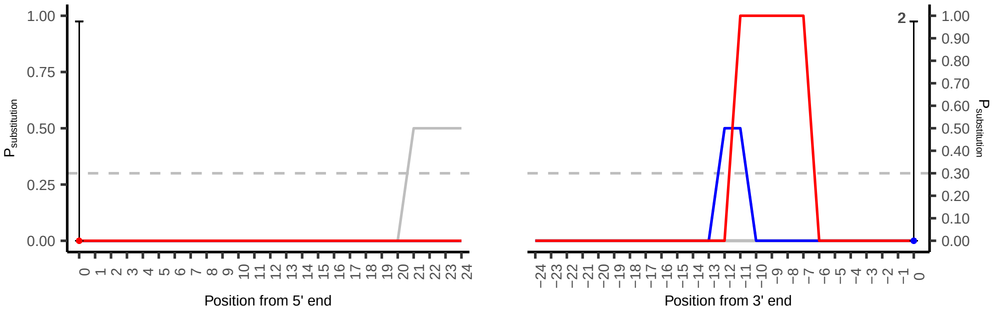
G.24 ELF046A

ELF046A_270_Metazoa_no_Primates_ingroup

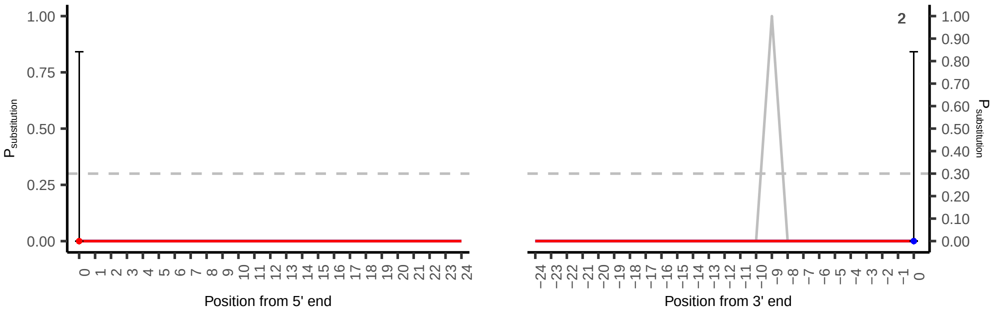


G.25 ELF050

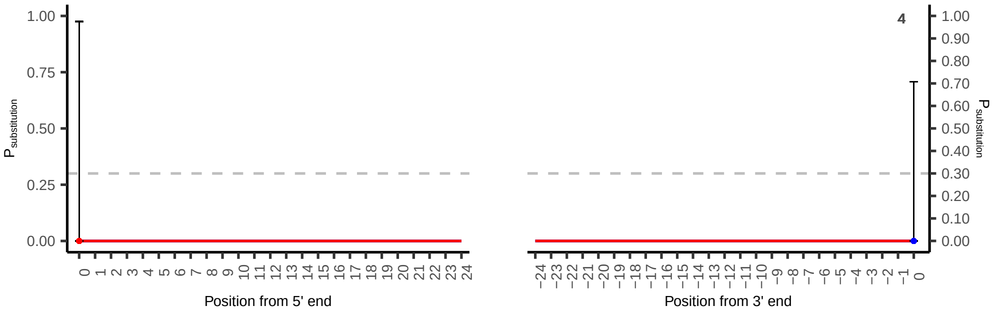
ELF050_595_Metazoa_no_Primates_ingroup



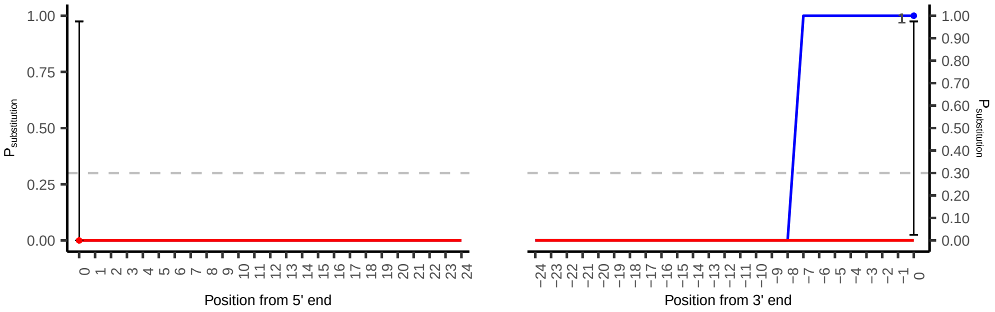
ELF050_450_Metazoa_no_Primates_ingroup



ELF050_354_Metazoa_no_Primates_ingroup

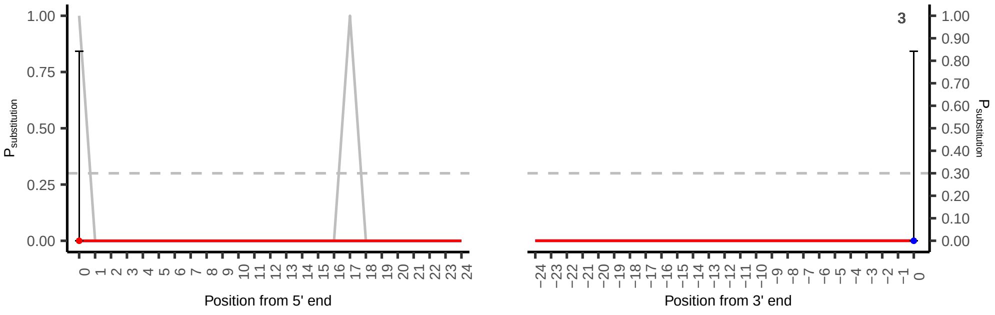


ELF050_250_Metazoa_no_Primates_ingroup

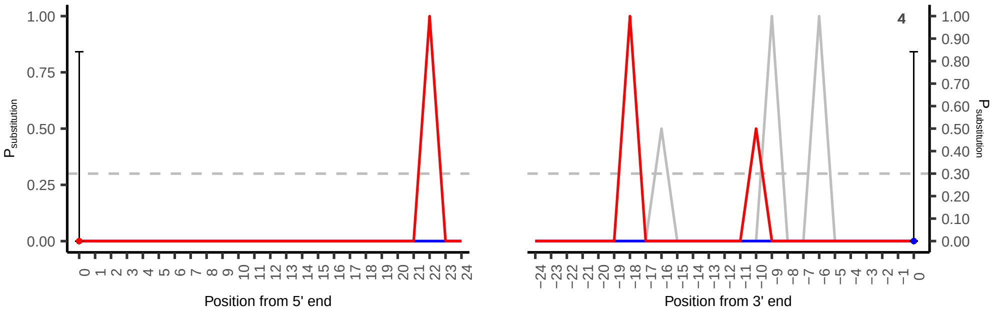


G.26 ELF053

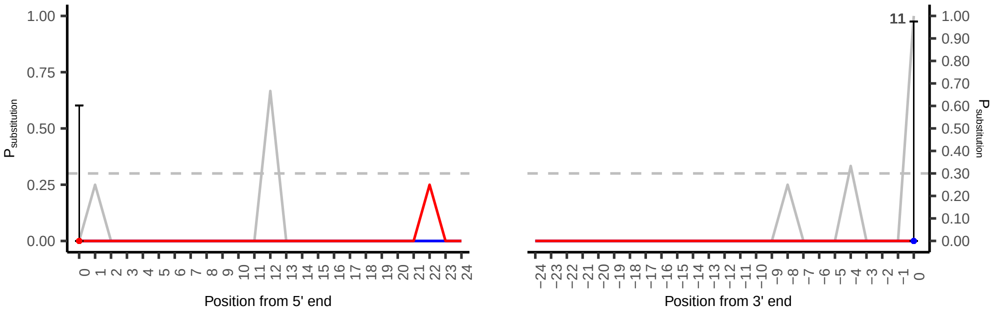
ELF053_336_Metazoa_no_Primates_ingroup



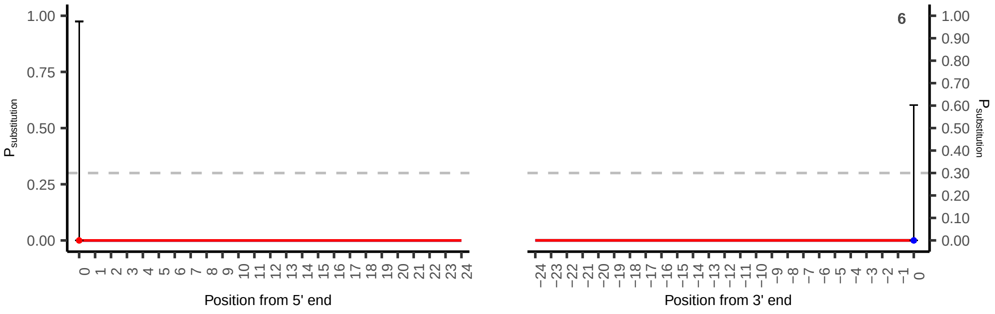
ELF053_289_Metazoa_no_Primates_ingroup

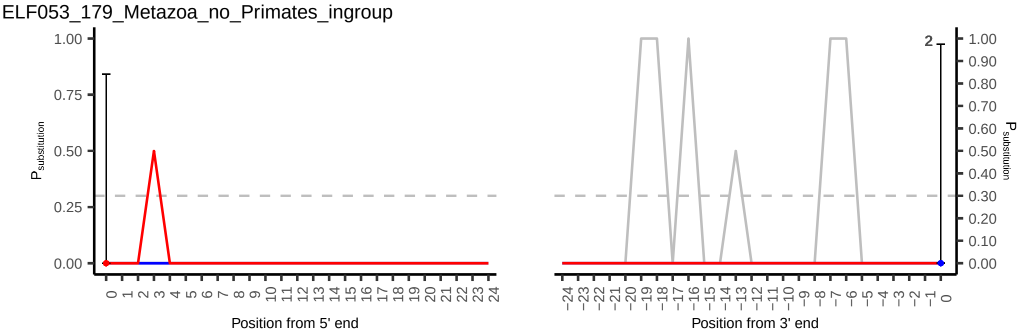


ELF053_275_Metazoa_no_Primates_ingroup



ELF053_214_Metazoa_no_Primates_ingroup




ELF053_179_Metazoa_no_Primates_ingroup
1.00
0.75
0.50
0.25
0.00
P_substitution
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from 5' end
−24 −23 −22 −21 −20 −19 −18 −17 −16 −15 −14 −13 −12 −11 −10 −9 −8 −7 −6 −5 −4 −3 −2 −1 0
Position from 3' end
2
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
P_substitution