This is a repository copy of *Phylogenetic patterns and phenotypic profiles of the species of plants and mammals farmed for food*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/138020/

Version: Accepted Version

Article:

1 **Phylogenetic patterns and phenotypic profiles of the species of plants and**

2 **mammals farmed for food**

3

4 Rubén Milla[1*]; Jesús M. Bastida[1]; Martin M. Turcotte[2]; Glynis Jones[3]; Cyrille Violle[4],

5 Colin P. Osborne[5]; Julia Chacón-Labella[1]; Ênio E.Sosinski Jr.[6]; Jens Kattge[7,8]; Daniel C.

6 Laughlin[9]; Estelle Forey[10]; Vanessa Minden[11,12]; Johannes H. C. Cornelissen[13]; Bernard

7 Amiaud[14]; Koen Kramer[15]; Gerhard Boenisch[7]; Tianhua He[16]; Valerio D. Pillar[17]; and

8 Chaeho Byun[18]

9

10 [1]Departamento de Biologia, Geología, Física y Química Inorgánica. Universidad Rey Juan

11 Carlos. Tulipán s/n, 28933 Móstoles, Spain

12 [2]Department of Biological Sciences, University of Pittsburgh, Pittsburgh PA 15260, USA

13 [3]Department of Archaeology, Minalloy House. 18-22 Regent Street, University of

14 Sheffield, Sheffield S1 3NJ, United Kingdom.

15 [4]CEFE UMR 5175, CNRS – Université de Montpellier – Université Paul-Valéry

16 Montpellier – EPHE – 1919 route de Mende, FR-34293 Montpellier, Cedex 5, France.

17 [5]Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN,

18 United Kingdom.

19 [6]Embrapa Clima Temperado, Pelotas, Rio Grande do Sul, Brazil.

20 [7]Max Planck Institute for Biogeochemistry, Hans Knöll Str. 10, 07745 Jena, Germany

21 [8]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher

22 Platz 5e, 04103 Leipzig, Germany.

23 [9]Department of Botany, University of Wyoming, 1000 E. University Ave., Laramie, Wyoming 82071 USA.

25 [10]Laboratoire Ecodiv URA/EA 1293, IRSTEA, FR CNRS 3730 SCALE, UFR Sciences et Techniques, Normandieuniversite, Universite de Rouen, 76821 Mont Saint AignanCedex, France.

28 [11]Department of Biology, Ecology and Biodiversity, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium.

30 [12]Landscape Ecology Group, Institute of Biology and Environmental Sciences, Carl von Óssietzky-Strasse 9-11, 26111 Oldenburg, Germany.

32 [13]Systems Ecology, Departament of Ecological Science, Faculty of Earth and Life Sciences, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.

34 [14]UMR Ecologie et Ecophysiologie Forestières, Faculté des Sciences et Technologies, Université de Lorraine, INRA, F-54506, Vandoeuvre-lès-Nancy Cedex, France.

36 [15]Wageningen Environmental Research (Alterra) - Wageningen University, Droevendaalsesteeg 4, 6700AA Wageningen, The Netherlands.

38 [16]School of Molecular and Life Sciences, Curtin University, GPBox U1987, WA 6845, Australia.

40 [17]Department of Ecology, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 91501-970, Brazil.

42 [18]School of Civil & Environmental Engineering, Yonsei University A460 121dong Yonsei-ro 50, Seodaemun-gu, Seoul 03722, Korea.

45 * Corresponding author: (+34) 914 884 645; E-mail: ruben.milla@gmail.com

47
48

49   **The origins of agriculture were key events in human history, during which people**

50   **came to depend for their food upon small numbers of animal and plant species.**

51   **However, the biological traits determining which species were domesticated for food**

52   **provision and which were not, are unclear. Here, we investigate the phylogenetic**

53   **distribution of livestock and crops, and compare their phenotypic traits with those of**

54   **wild species. Our results indicate that phylogenetic clustering is modest for crop**

55   **species but more intense for livestock. Domesticated species explore a reduced portion**

56   **of the phenotypic space occupied by their wild counterparts and have particular traits**

57   **in common. For example, herbaceous crops are globally characterized by traits**

58   **including high leaf nitrogen concentration and tall canopies, which make them fast**

59   **growing and proficient competitors. Livestock species are relatively large mammals**

60   **with low basal metabolic rates, which indicate moderate to slow life histories. Our**

61   **study therefore reveals ecological differences in domestication potential between**

62   **plants and mammals. Domesticated plants belong to clades with traits advantageous**

63   **in intensively-managed high-resource habitats whereas domesticated mammals are**

64   **from clades adapted to moderately productive environments. Combining comparative**

65   **phylogenetic methods with ecologically relevant traits has proven useful to unravel the**

66   **causes and consequences of domestication.**

67

68   The plant and animal species domesticated for human food supply represent only a small

69   fraction of global biodiversity. Of around 370,000 extant flowering plants[1], only 1,000-

70   2,000 have undergone some form of domestication for that purpose [2–4]. Similarly, humans

71   have domesticated 20-31 species of mammals for food [5,6], from ~ 5,400 species

72   contemporary to late Paleolithic people[7]. The taxonomic distribution of species used for

4

73   farming seems non-random[5,8], such that certain families include numerous domesticated

74   species (*e.g.* grasses and legumes among flowering plants, and bovids and camelids among

75   mammals), while many others contain none. An uneven phylogenetic distribution of the

76   species that became domesticated would imply that certain combinations of phenotypic

77   traits are more adaptive for husbandry, if these traits are phylogenetically conserved[9].

78   However, global comparative analyses between domesticates and wild relatives are rare[10–12]

79   or consider taxonomically and/or geographically restricted groups of species[13,14]. Filling

80   that gap would direct agricultural sciences towards the phylogenetic groups and traits that

81   could be pursued for new food sources. Additionally, investigating such patterns at a global

82   scale, while explicitly linking phylogenetic and trait distributions, would highlight the

83   usefulness of the tools and concepts of evolutionary ecology to address questions at its

84   interface with agricultural sciences and archaeology.

85   The phenotypes of current livestock and crops are the result of early domestication

86   processes and millennia of unconscious and deliberate selection under farming[15]. Evolution

87   under farming has caused the traits of domesticated species to change under shifting

88   selective forces[13]. For example, local breeding preserves mutants that would otherwise be

89   eliminated by natural selection and thereby offset the sampling effects of early farming[4]. In

90   fact, diversifying selection has promoted remarkable variance in the size of crop seeds or in

91   animal coat colors[5,16]. Conversely, directional selection for productivity has resulted in the

92   convergence of a number of livestock and crop traits, *i.e.* the domestication

93   syndrome[5,9,17](but see [18,19]). Domestication syndrome traits include increased docility or

94   reductions in brain size in livestock[5,6,20] and increased palatability or the loss of seed

95   dispersal mechanisms in crop plants[4,9].

96   Adaptations to the farming environment might also promote convergent evolution

97   *via* natural selection[13]. For instance, evolution in high-resource croplands should select for

98   suites of traits that enable fast resource acquisition and rapid canopy closure, according to

99   trait-based ecology[21]. A few recent studies have partially supported this view, showing that

100   several crop plants display traits indicative of high competitive ability[14,22–24]. Therefore, if

101   directional and stabilizing selection are strong, the phenotypic diversity of domesticated

102   species will be low, adding to early domestication bottlenecks. In contrast, diversifying

103   selection, associated with centuries-long geographic expansion under farming, is expected

104   to promote the widening of phenotypic spaces[16]. The net effects of the early filtering of

105   wild species, of subsequent domestication processes, and of later crop and livestock

106   evolution, on the phenotypic spaces explored by domesticated species remain unknown. In

107   this paper, we show the results of the first broad-scale phylogenetic analyses addressing

108   whether domesticates are a limited phylogenetic and phenotypic sample of wild plants and

109   animals, and uncovering traits that distinguish domesticated species from wild species.

110   We used phylogenetic comparative methods to investigate the phylogenetic patterns

111   of domestication events, and to ask whether domesticates are phenotypically distinct from

112   their wild relatives. We did this by compiling and analyzing two large datasets. First, we

113   compiled a database on the distribution of species domesticated for food across mammal

114   and angiosperm families and genera. With that dataset we investigated evolutionary

115   patterns of the relative abundance of domesticated species (proportion of all domesticated

116   species within a given genus or family), and of domestication frequencies (proportion of all

117   species in a genus or family that were domesticated) across mammal and angiosperm

118   phylogenies. Second, we put together a large-scale database of three key phenotypic traits

119   for farm mammals (size-corrected basal metabolic rate, adult body mass and neonate body

120   mass) and crop plants (leaf nitrogen content, plant canopy height and seed dry mass) across

121   a broad sample of domesticated (23 mammals, 181 angiosperms) and wild species (885

122   mammals, 2,943 angiosperms). Traits were selected based on their key functional relevance

123   for resource-use-acquisition, life history and ecological strategies, both for domesticated

124   and for wild species[21,25–28]. Using this second dataset, we compared the phenotypic spaces

125   of domesticates to those of their wild relatives. We set out to address three questions: 1)

126   How are food domesticates distributed across the phylogenies of mammals and

127   angiosperms? 2) Do livestock and crop species have particular phenotypic profiles, when

128   compared to their wild counterparts? And, if so, 3) Do the phenotypic traits of domesticated

129   species fall within the trait space exhibited by wild species, or do they extend their

130   phenotypic space beyond the boundaries set by wild plants and animals?

131

132   **Results**

133   Some families and genera contain more livestock or crop species than others. Livestock

134   species were found in only ten families. In particular, Bovidae harbour ~40 % of

135   domesticated species (Supplementary Table 1), and only 22 genera of mammals contain

136   domesticated species (Supplementary Table 1). In contrast, crop species were distributed

137   across 120 families and 453 genera of angiosperms (Supplementary Data 1). Fabaceae,

138   *Solanum*, and Poaceae are examples of taxa yielding high proportions of crops. The

139   abundances of domesticated species were far from randomly distributed across families and

140   growth forms (plants) or dietary types (mammals), both for mammals and angiosperms

141   (Supplementary Table 2). We next investigated whether the above deviations from a

142   random distribution were phylogenetically structured.

143   In general, we found that the phylogenetic structure of domestication abundances

144   and frequencies was modest for plants, though stronger for mammals. First, Local

145   Indicators of Phylogenetic Affinity (LIPAs) indicated that ca. 90% of the plant families

146   hosting domesticated species were randomly distributed with respect to the domestication

147   status of their phylogenetic neighbourhood (Supplementary Data 2). Approximately 10% of

148   angiosperm families departed from a non-significant LIPA score (Supplementary Data 2).

149   However, such departure signalled over-dispersion (*i.e.* focal family surrounded by families

150   without domesticated species more than the random expectation), rather than clustering

151   (Supplementary Data 2). Interestingly, that ca. 10% of angiosperm families include

152   Fabaceae, Poaceae, Rosaceae, Solanaceae or Asteraceae, which host crops of high

153   agricultural relevance. For mammals, four (Suidae, Cervidae, Caviidae, and Cunilidae) out

154   of ten families with domesticated species had at least one LIPA score indicating

155   phylogenetic clustering, either for abundances or frequencies. Second, we investigated

156   phylogenetic clustering at the scale of the whole phylogenetic tree. When two contrasting

157   evolutionary models were compared (*i.e.* a Brownian motion model of evolution,

158   representing strong phylogenetic structure, *vs.* a star phylogeny, representing full

159   phylogenetic independence), phylogenetically independent models showed better fit to the

160   data than Brownian motion models, both for mammals and angiosperms (Supplementary

161   Tables 3 and 4). Finally, global phylogenetic clustering was investigated with a gradual

162   approach (phylogenetic signal), which complemented the binary (non-phylo vs phylo)

163   comparison of phylogenetic models above. The frequency of domestication events

164   generally showed a phylogenetic signal (Figure 1 and Supplementary Figure 1, but see

165   angiosperm families in Figure 1). Domestication abundance, instead, showed low or no

166   phylogenetic signal in angiosperms, low signal in mammals at the family scale, and high

167   signal when mammals were examined separately for each dietary type (Figure 1 and

168   Supplementary Figure 1).

169   The subset of domesticated species used for phenotypic space analyses covered a

170   wide range of phylogenetic and geographic origins (Figure 2). In spite of this taxonomic

171   and geographic diversity, domesticated species were distributed across a portion of the

172   phenotypic space generally occupied by wild species, and rarely fell beyond the bounds set

173   by wild mammals and plants (Figures 3 and 5; and Supplementary Tables 5 and 6, but see

174   ruminant livestock in Figure 3). Livestock occupied a small subset of the phenotypic spaces

175   of wild mammals (Figure 3 and Supplementary Table 5). Within the common phenotypic

176   boundary occupied by wild and domesticated mammals, livestock species were, on average,

177   larger as adults and gave birth to larger neonates, but had lower basal metabolic rates,

178   compared to their wild counterparts (Figure 4 and Supplementary Tables 7 and 9). Those

179   phenotypic biases were upheld when investigated separately for ruminants and non-

180   ruminants, though domestic ruminants lied mostly outside the phenotypic boundaries of

181   wild ruminants. In contrast, domesticated crops have been selected from a wide range of

182   botanical diversity in the three focal traits (Figure 5 and Supplementary Table 6). An

183   exception to that pattern is the small phenotypic space occupied by domesticated

184   graminoids (grass-like monocot plants), in comparison with their wild analogues (Figure 5

185   and Supplementary Table 6). However, although crops were phenotypically diverse, they

186   occupied some regions of the phenotypic space more than others, which lead to phenotypic

187   differentiation when compared to wild plants. Specifically, herbaceous crops, both

188   graminoid and non-graminoid, were generally larger plants with larger seeds, and with

189   leaves with higher nitrogen content, than their wild counterparts (Figure 6 and

190   Supplementary Tables 8 and 10). Woody crops were more similar to wild woody plants,

191 though they consistently had larger seeds (Figure 6 and Supplementary Tables 8 and 10).

192 The phenotypic departure of domesticated species from the trait medians of their wild

193 counterparts was generally unrelated to the differences in geographic origin, climate at

194 geographic origin, or antiquity of domestication of crop and livestock species

195 (Supplementary Table 11).

196

197 **Discussion**

198 Our results provide the first quantitative global test of the long-standing hypothesis that

199 domestication events have a strong phylogenetic structure. We found only weak evidence

200 for phylogenetic clustering in crops, but stronger evidence of such clustering in livestock

201 species. Interestingly, the non-random phylogenetic distribution of species that became

202 domesticated was associated with non-random phenotypic spaces of crops and farm

203 mammals. Livestock species had moderate to low basal metabolic rates, gave birth to large

204 offspring and were large adults. Herbaceous crops had high leaf nitrogen content (an

205 indicator of photosynthetic rates), were large as adults, and produced large seeds. These

206 results show that domesticated mammals and plants occupy specific portions of the spectra

207 of phenotypic variation[21,25–28]. Despite such phenotypic differentiation, and in spite of

208 substantial trait evolution during domestication[16], domesticated species were rarely

209 positioned outside the bounds of the phenotypic spaces set by the wild species of their kind,

210 excluding ruminant livestock. These findings have important implications for

211 understanding the potential of wild species for farming, the patterns of phenotypic

212 convergence under domestication, and the adaptation of wild species to the environmental

213 conditions of farming habitats[13,18,20,29,30]. In addition, we demonstrate that a macro-

10

214   evolutionary approach, scarcely embraced in the domestication literature, has the strong

215   potential to advance this field.

216   Our analyses showed that domesticated mammals represent a small sub-sample of

217   the total phenotypic variation displayed by wild animals. This was expected, since livestock

218   species are a very small fraction of all contemporary mammals. More unexpectedly, farmed

219   mammals had mid-to-low basal metabolic rates and were mid-to-large adults and neonates.

220   These results portray domesticated mammals as moderate to slow life history strategists,

221   *i.e.* species with intermediate body temperatures, with moderately long juvenile periods,

222   giving birth to few but relatively large offspring, and living for reasonably long time spans,

223   accordingly to the fast-slow life-history framework[26,27]. Low basal metabolic rates, which

224   might entail slow relative growth rates[31], are adaptive in unproductive and unpredictable

225   environments in ruminants[32], rodents[33] and mammals in general[27,28]. In addition, a

226   moderate-slow lifestyle might genetically associate with behavioural traits that are critical

227   to animal domestication, such as boldness, tameness, or sociality[33–35], which remains to be

228   investigated. Interestingly, many domesticated mammals evolved body size reduction after

229   domestication[29,36]. This suggests that the moderate-slow lifestyle of livestock identified in

230   our current work is largely the result of early selection of wild animals, rather than of

231   further evolution under farming. In livestock species that show such body size reductions

232   during domestication, decreases in sexual dimorphism were also reported, which, following

233   Rensch´s rule[37], might account for their overall smaller adult size [5,29,38]. Further studies

234   advancing this line of enquiry should consider the wild progenitors of livestock species,

235   account for sexual dimorphism, and address species domesticated for purposes beyond food

236   provision, which clearly display body size reduction after domestication (*e.g.* dogs or cats).

237   For crops, our results for non-woody plants are compatible with hypotheses positing

238   that early human selection favoured traits advantageous in the fertile, disturbed habitats

239   surrounding human settlements and early agricultural fields[5,9,30]. Herbaceous crops

240   occupied only a portion of the phenotypic space of their wild counterparts, suggesting

241   habitat filtering[39]. This is in line with previous case studies reporting that crops are a

242   subsample of the phenotypic variation found in nature, and have not surpassed the

243   biological limits observed for wild plants[13,23,40,41], which suggests limitations of artificial

244   selection to move phenotypes beyond what is observed in the wild. Additionally, the

245   phenotypic profiles described here indicate that herbaceous crops are fast growing species

246   (high leaf nitrogen content) and proficient competitors in resource abundant environments

247   (tall plants and large seeds)[21,25], which would suit the ecological requirements of early

248   agricultural habitats[42]. Such phenotypic differentiation passed unnoticed in the previous

249   literature addressing smaller sets of crops, where crop-specific contrasting results were

250   commonly reported[14,22–24,41]. This highlights the relevance of a global approach and the

251   usage of the comparative method in this field. Woody crops yielded large seeds but, in

252   contrast to herbs, were neither tall species nor species with high nitrogen content in their

253   leaves. Multiple explanations might account for such discrepancy between growth forms,

254   including trade-offs in resource allocation to fruit tissue, to vegetative growth, and to the

255   maintenance of woody tissues, the clonal mode of propagation common to woody crops, or

256   crop uses[4,9,18,43].

257   A direct comparison of the phenotypic spaces of plants and mammals yielded an

258   additional relevant insight. In plants, metabolism and size are largely decoupled[25]. In

259   contrast, the evolution of metabolic rates and body size are coordinated in mammals[44]. Our

260   trait analyses were consistent with these patterns, both for wild and domesticated species

261 (Figures 3 and 5). The phenotypes of domesticates therefore evolve under different

262 constraints in plants and mammals. As a consequence, we found wider combinations of

263 traits and larger phenotypic spaces for crops than for livestock. Endothermy, design of the

264 vascular system, and size-metabolism constraints might impede the simultaneous

265 maximization of mass-specific growth rates and body size in mammals[45], both of which are

266 desirable traits for productive husbandry. Within these constraints, human selection for

267 farming favoured animals with intermediate-high sizes, although at the cost of low

268 metabolic rates, and thus probably modest relative growth rates. Breeding livestock that

269 overcome size-metabolism constraints are therefore expected to be challenging. On the

270 other hand, crop plants occupied a wider part of the tradeoff-free spectrum of metabolism

271 (leaf nitrogen) *vs.* size (plant and seed sizes), in line with their wild counterparts. Plant

272 modularity and nitrogen transfer among modules, which underpin such uncoupling between

273 metabolism and size[46], might thus promote the greater phenotypic diversity of crops than

274 livestock mammals, even when considered within plant growth-forms. Breeding to

275 simultaneously optimize variation in plant and organ sizes, and variation in metabolic rates,

276 might be tradeoff-free in plants.

277 Finally, we highlight two limitations of the current work. First, trait data come from

278 plants and animals sampled in their typical habitats, which are different among species, and

279 are undoubltley so among organisms living in farm vs wild habitats. Thus, the phenotypic

280 patterns encountered here came from a mixture of genotypic and environmental drivers,

281 whose relative importance is plainly indistinguisible for large scale macroecological

282 studies. However, the few experiments that grew sets of crops and of their wild progenitors

283 in common gardens, and phenotyped the types of traits that we measured here, tended to

284 concur with our results, which suggests a strong genetic component at play [22,40,47].

285 Therefore, it will be necessary to take these experimental approaches one step further in

286 terms of number of species and phylogenetic breadth, both for wild and domesticated

287 organisms. Second, we note that, for both domesticated and wild species, intraspecific

288 variation was not considered here. Such variation, even if unmanageable to explore

289 systematically at the macroevolutionary scale of the current study, might expand trait

290 spaces greatly[48]. Leaf trait intraspecific variation in sunflower, wheat, maize and coffee was

291 recently reported, and occupies a fair portion of wild phenospaces, though very rarely

292 spreading outside wild boundaries [40,49,50]. We supplemented those leaf trait patterns

293 available in the literature with *ad hoc* analyses for seed and plant sizes (Supplementary

294 Figure 2). Similarly to leaf traits, intraspecific variation in size traits is constrained within

295 wild envelopes in sunflower, soybean, and barley (Supplementary Figure 2). However,

296 maize, as a crop species in which the seed-plant size centroid is outside the phenosphace of

297 its corresponding wild analogs (Figure 5), expands most of its intraspecific variation

298 outside wild boundaries (Supplementary Figure 2). In our view, these analyses, and the

299 available literature, are still too scant to reach solid solid conclusions on the role of

300 intraspecific variation in trait differences between wild and domesticated organisms. Thus,

301 investigating how and to what extent diversifying breeding of domesticates expands

302 phenospaces is needed to bridge the macro- and the micro-evolutionary scales.

303 This study placed domesticates within their broader botanical and zoological

304 context, which facilitates appreciation of the qualities and potentials of the species that

305 support our food system, and could help in the search for suitable future domesticates.

306 Suitable phenotypes among mammals include moderate-slow life histories, while fast

307 growth traits and large size dominate among herbaceous crops. Further, the usage of a

308 phylogenetic comparative approach, which was seldom embraced in the domestication

309   literature (but see[12,14,51]), provided unique insights, that are unattainable at smaller scales.

310   Overall, our work indicates that certain phylogenetic clades and phenotypic profiles have

311   been more exploited than others for provisioning human food, and that such filtering was

312   based on strong, conscious or unconscious, early selection at human settlements. Future

313   work should investigate biogeographic and historical determinants, disentangle genotypic

314   from environmental drivers, and address the microevolutionary scale, of the broad

315   phylogenetic and phenotypic patterns of differences between domesticated and wild kins

316   revealed here.

317

318   **Methods**

319

320   **Study system**

321   We explored macroevolutionary patterns of the distributions of species domesticated for

322   human food, and compared their phenotypic trait space occupancy with that of wild

323   species. We included the broadest possible diversity of mammals and angiosperms farmed

324   for human food provision, with distinct domestication histories and intensities,

325   phylogenetic affinities, and phenotypic profiles (see Supplementary Table 1,

326   Supplementary Data 1, and Figures 1-6).

327

328   **Collection of data on the distribution and abundance of angiosperms and mammals**

329   **farmed for food**

330   In order to assess patterns of distribution and abundance of food domesticates across

331   phylogenies, we compiled the abundances and identities of domesticated species within

332   mammals and angiosperms, at the family and genus levels. We aimed to build a

333   comprehensive working list of all putative species domesticated for food. Therefore we

334   included all species for which farming could be substantiated, as an indicator of some

335   degree, even if incipient, of domestication.

336   We used the literature to build a preliminary list of mammals farmed for food[5,6,52,53].

337   We excluded most anectodal evidence of deer and antelopes farming, but for the sake of

338   inclusivity we considered species like reindeer, sika deer, moose, bison, muskox or

339   common eland, which are regularly farmed regionally, and thus should include incipiently

340   domesticated populations. We also included recent incipient domesticated species for food

341   like the African giant rat, or the South American paca. The set of mammals comprised 27

342   domesticated species, distributed across 22 genera and 10 families. Taxonomy was checked

343   using the *taxize* R package[54].

344   For plants, we compiled a list of all putative domesticated species from the

345   literature[2,4,18,55]. From that list, we extracted taxa for which cultivation could be

346   demonstrated, and filtered that extract by species used for human food or forage. We

347   include forage species because human food supply depends on livestock feeding, and a

348   substantial part of the agricultural land is devoted to that usage. To assign usage for food or

349   forage we primarily used two studies[2,56], supplemented with http://www.pfaf.org and with

350   other sources when needed. Plant taxonomy was standardized using the *Taxonstand* R

351   package[57], synonymous names were cleaned, and binomials were attributed to families

352   using The Plant List (http://www.theplantlist.org/). Thirty-five species were taxonomically

353   unresolved, and their provisional binomials were used. Our final 944 species list should

354   include the vast majority of angiosperms known to have been cultivated for human food.

355   The species of our list belonged to 453 genera and 120 angiosperm families.

356   Domesticated mammals were grouped into two dietary types, *i.e.* ruminants and

357   non-ruminants, based on the MammalDIET database[58], and on information at

358   http://www.ultimateungulate.com/ungulates.html - Ruminantia. The ruminant category also

359   included pseudo-ruminants, *i.e.* Camelidae and Hippopotamidae. Domesticated plants were

360   grouped by growth form into herbaceous, graminoids (Poaceae, Cyperaceae and

361   Juncaceae), and woody plants (shrubs, trees, woody vines and tree-like species). Growth

362   forms were assigned using the TRY database[59], the Global Woodiness Database[60], and were

363   supplemented species-wise with primary literature when not available in those sources.

364

365   **Selection and compilation of phenotypic traits**

366   We selected a set of three traits for mammals and three traits for angiosperms that are

367   functionally analogous. Selection of traits used the following criteria: 1) previous evidence

368   of domestication effects on those traits[14,22,29,30,36]; 2) functional relevance for basic

369   metabolism, resource-use, competition and reproductive strategies; and 3) availability of

370   data, both for domesticated and for wild species. By functional analogy, the selected traits

371   can be grouped into: 1) plant canopy height (m) and mammal adult body mass (g) as

372   proxies for adult size and competitive ability for resources[61,62]; 2) leaf nitrogen content (mg

373   N mg^{-1}), and size-corrected mammalian basal metabolic rate (ml O_2 h^{-1} g^{-1}) as proxies of

374   photosynthetic and metabolic rate, respectively[26,63]; and 3) seed dry mass (mg) and neonate

375   body mass (g) as proxies for offspring size and likelihood of survival [64,65].

376   The two trait datasets were assembled separately for mammals and angiosperms.

377   The dataset on mammal traits was compiled from the PanTHERIA database for adult and

378   neonate body mass[66], and from [67] for basal metabolic rate (supplemented with PanTHERIA

379   if basal metabolic rate was unavailable at [67]). The dataset on mammal traits comprised 480

380   species, including 23 domesticated species (see Supplementary Table 1 for the identity of

381   domesticated species). For domesticated plants, trait data were not available for all 944

382   crops species. Thus, we started by retrieving data from the 203 species list of domesticated

383   plants published in Meyer et al.[18], which was built to maximize crop diversity. This helped

384   to avoid biases in growth form, taxonomy, or agricultural relevance. That list of crop taxa

385   was further supplemented, again using criteria to maximize diversity and filtering by the

386   availability of trait data. Then, data for wild angiosperms were added. Overall, the sources

387   for angiosperm trait data were: 1) the TRY database[59] (https://www.try-db.org, accessed

388   2016-11-13; 2) literature searches for wild species incompletely recorded or not present in

389   the TRY database; 3) literature searches for trait data of crop species, which are mostly

390   absent from TRY[68]; 4) our own data already collected on crops and other wild species (see

391   Data availability section). The final angiosperm trait dataset comprised 3,124 species,

392   including 181 domesticated species (see Supplementary Data 1 for the identity of

393   domesticated species).

394   Plant species names were standardized using the *Taxonstand* R package[57], and were

395   attributed to families according to The Plant List (http://www.theplantlist.org/). A majority

396   of crop binomials are synonymous to the wild genotypes of their wild progenitors.

397   Therefore, to decide whether a given observation of a crop related binomial was attributable

398   to a crop or a synonymous wild species we used the following criteria. First, we looked for

399   explicit statements in the original publication or database on whether the studied taxa were

400   crop or wild. If uncertain, an observation was assigned to "wild" if the study was

401   observational and was conducted under natural field conditions, or if the seeds for an

402   experiment were collected from wild stocks. In contrast, an observation was assigned to

403   "crop" if the seeds came from commercial companies, cultivars or varieties or if studies had

404   been conducted in an agricultural setting, and no explicit mention to wild status was found.

405

406   **Data handling prior to analyses**

407   Our angiosperm dataset had 1.51 % missing data (leaf nitrogen 1.70 %, plant height 1.27%,

408   and seed mass 1.58 %). Since several plots and analyses involved the joint use of two or

409   more traits, we adopted a multiple imputation approach to deal with missing data, following

410   recommended procedures[69,70]. We generated ten complete datasets using the Predictive

411   Mean Matching method (PMM) of the *MICE* package[71]. Phylogenetic relatedness (built as

412   described below) was incorporated into the imputation procedure as phylogenetic

413   orthogonal eigenvectors[72]. Results reported in the main body of the paper are from

414   averaged imputed data of those ten complete datasets. A dataset without imputed data, and

415   thus with a slightly reduced sample size, was used to test for robustness and sensitivity to

416   our data handling procedures. The results of sensitivity analyses were consistent with those

417   shown in the main text (Supplementary Note, and Supplementary Tables 12-14). The

418   mammalian traits dataset lacked basal metabolic rate data for six species, which were

419   estimated using the phylogenetically-corrected allometric scaling of adult body mass to

420   basal metabolic rate available at [67] ($BMR = 2.382m^{0.729}$). All continuous variables were

421   log-transformed prior to analyses. An exception was seed mass, which was log-generalized

422   transformed because a few crops do not yield seeds (*Musa acuminata, Vaccinium*

423   *corymbosum and Allium sativum*). This latter procedure is recommended when data contain

424   zeros, and the smallest positive value is not close to one.

425

426   **Macroevolutionary patterns in the abundance and frequency of food domesticates**

427 We performed separate analyses on the relative abundance of domesticated species, and

428 domestication frequency, at the family and genus levels. Relative abundance, at family and

429 genus level, was calculated as the number of domesticated species in a particular family or

430 genus divided by the total number of domesticated species. Because this metric does not

431 account for species richness within clades, we also calculated a domestication frequency

432 metric as the number of domesticated species in a focal family or genus per total number of

433 species in that same family or genus. These two metrics inform about different features of

434 the distribution of domesticated species. Species richness at family and genus level, needed

435 to compute domestication frequencies, was retrieved from lists of accepted names from

436 Wilson   and   Reeder   (2005)   for   mammals,   and   from   The   Plant   List   v1.1

437 (http://www.theplantlist.org/),   making   use   of   the   *taxonlookup*   R   package

438 (https://github.com/traitecoevo/taxonlookup), for plants.

439   Phylogenetic hypotheses were built at the family and genus levels for mammals and

440 for angiosperms separately. Mammal phylogenies were based on Bininda-Emonds et al's[73]

441 megaphylogeny as a backbone. The mammalian family-level tree included 142 families

442 (91% of total mammalian families), while the genus-level tree included 498 genera (39.6 %

443 of total mammalian genera). Seventy-two genera were ruminants and 337 were non-

444 ruminant herbivores (http://tolweb.org). Angiosperm phylogenies were based on the

445 PhytoPhylo megaphylogeny[60,74]. The angiosperm family-level tree included 404 families

446 (97% of total angiosperm families), while the genus-level tree included 7,233 genera (ca.

447 56 % of total angiosperm genera) (http://www.theplantlist.org/). All families, and all but

448 seven genera with domesticated species, were present in the megaphylogeny. Those seven

449 genera (*Gigantochloa, Nopalea, Parmentiera, Polianthes, Sphenostylis*, *Stizolobium,* and

450 *Vitellaria*) were bound to the PhytoPhylo megaphylogeny based on published phylogenies

451 of their respective families. The genus level tree included 2,745 genera of herbs, 421 of

452 graminoids, 3,500 of woody plants and 567 genera hosting both herbs and woody plants.

453 Genera containing herbs and woody plants were included in both of their respective growth

454 form analyses. Angiosperm phylogenies were fully resolved, but mammalian phylogenies

455 contained some internal polytomies (12 % of nodes in the family-level tree, 24 % in

456 ruminants and 38 % in non-ruminant herbivores). Therefore, analyses were run across 100

457 randomly resolved family and genus-level mammalian trees.

458   To assess whether the abundance and frequency of domestication are randomly

459 distributed across mammal and angiosperm families and genera, we performed four

460 complementary analyses. First, we conducted randomisation analyses to test whether the

461 distribution of the abundances of domesticated species per family, and of the frequencies of

462 domestication events, differed from random expectations. Observed kurtosis and skewness

463 of the distribution of abundances were compared to that of 1,000 randomised distributions

464 at each respective level. Second, we computed Local Indicators of Phylogenetic

465 Association (LIPA) based on Local Moran's I [75], to detect families surrounded by

466 phylogenetic neighbourhoods with similar or distinctive (positive or negative

467 autocorrelation, respectively) relative abundances of domesticated species or domestication

468 frequencies. For each LIPA score, statistical significance was assigned by performing non-

469 parametric two-sided tests with 999 randomisations. For mammals, LIPAs were averaged

470 across the 100 randomly resolved trees. Third, we calculated the phylogenetic signal of the

471 relative abundances of domesticated species, and of the frequencies of domestication, at the

472 genus and family levels, and separately for mammals and angiosperms. Provided that our

473 data followed either zero-inflated log-normal (abundances), or binomial (frequencies),

474 distributions, we did not calculate standard Pagel's λ or Blomberg's K metrics, which are

475   meant for continuous data with normal or log-normal distributions. Instead, we computed

476   phylogenetic signal as the phylogenetic heritability parameter (λ) of phylogenetic mixed

477   models (PMMs), where our metrics of domestication were the response, an intercept was

478   set as the sole fixed-effects predictor, and inverse matrices of the phylogenetic distances

479   matrices were the covariance structure terms[76,77]. PMMs allow the specification of family

480   distributions of data deviating from Gaussian. The lambda parameter of such models,

481   specified without meaningful fixed-effect predictors, and without additional covariance

482   structures in the random term, is the phylogenetic signal of the response variable,

483   analogously to a null Phylogenetic Generalized Least Square (PGLS) model[76]. Fourth, we

484   fitted and compared two evolutionary models to test whether relative abundances and

485   domestication frequencies were phylogenetically structured or phylogenetically

486   independent. We used a Brownian motion (BM) model to approximate neutral drift

487   evolution or randomly fluctuating selection[78]. Under BM, relative abundances and

488   frequencies evolve as a random walk through trait space along the branches of the

489   phylogeny, and thus represents strong phylogenetic structuring. BM was compared to a

490   non-phylogenetic model (a star phylogeny), which was used to approximate a phylogenetic

491   independent distribution. To compare both models we used the bias-corrected Akaike

492   Information Criterion (AIC_c), and calculated the difference between the AIC_c of the best

493   (smallest AIC_c) and the alternative model[79]. In addition, for each model we calculated the

494   AIC_c weights (AIC_c-w), with a high AIC_c-w indicating a low relative AIC_c for that model

495   and hence higher support[79]. For mammals, all parameters were averaged across 100

496   randomly resolved trees and the percentage of preferred models was calculated.

497   Phylogenetic signals were computed using the *MCMCglmm* function of the R package

498   *MCMCglmm*[80], setting family distribution as zero inflated poisson for domestication

22

499 abundances, and as binomial for frequencies of domestication. Evolutionary model fitting

500 was performed with the *FitContinuous* function of the R package *geiger*[81]. Local Moran's *I*

501 was calculated using the *lipaMoran* function of the R package *phylosignal*[75].

502

503 **Comparative analyses of phenotypic trait space occupancy of wild and domesticated**

504 **species**

505 New sets of phylogenetic trees were built at the species level for those mammals and

506 angiosperms included in our traits datasets. The mammal phylogeny for the 480 species

507 with trait data was built from Bininda-Emonds et al's[73] megaphylogeny using Phylomatic

508 v.3[82] (http://phylodiversity.net/phylomatic/pmws). The angiosperm phylogeny for the 3,124

509 species with trait data was based on the PhytoPhylo megaphylogeny[60,74], and was built

510 using scenario three of the R package *S.PhyloMaker*[74]. To account for phylogenetic

511 uncertainty (20.3 % of unresolved nodes for mammals and 15.3 % for angiosperms), all

512 analyses were performed on 100 randomly resolved trees by using the *multi2di* function of

513 the *ape* R package[83].

514 To visualize the phenotypic spaces explored by wild and domesticated species we

515 used bivariate phenospaces. Additionally, we used convex hulls to draw the minimum

516 convex envelope for each pair of traits, domestication status and growth form or dietary

517 type[84]. In addition, for each growth form or dietary type, we calculated the area and volume

518 of each three-traits convex hull. To test for significant differentiation in trait space between

519 domesticated and wild species, we performed phylogenetic-corrected MANOVAs and

520 ANOVAs, separately for each growth form or dietary type. Convex hull calculations were

521 performed using function *convhulln* of the R package *geometry*[85]. Phylogenetic

522 MANOVAs/ANOVAs were run with the function *aov.phylo* in the R package geiger[81]. See

523   Supplementary Methods for tests on whether the phenotypic departure of domesticated

524   species from the trait medians of their wild counterparts was related to differences between

525   domesticates in geographic origin, climate at geographic origin, or antiquity of

526   domestication. All analyses of the paper were conducted in R v3.4.3[86].

527

528   **Data availability**

529   All phenotypic traits of mammalian species included in this study are available from the

530   literature (see Methods section). For plants, most data are available from the database

531   TRY[59] (https://www.try-db.org), and all original sources of TRY data are listed in

532   Supplementary References 1. All references for data not included in TRY are available in

533   the Supplementary References 2. Unpublished data owned by R.M. and J.M.B. are

534   available from Supplementary Data 3. Unpublished data from the University of Sheffield

535   database of weed functional attributes can be requested from G.J. Lists of livestock and

536   crop taxa are available at Supplementary Table 1 and Supplementary Data 1, respectively.

537   Phylogenetic trees used in this study are available from Supplementary Data 4. Data on

538   geography and climate at domestication sites are available as Supplementary Data 5.

539
540   **References**
541

542   1.   RBG Kew. *The State of the World's Plants Report - 2016*. (Royal Botanic Gardens,
543   Kew, 2016).
544   2.   Zeven, A. C. & De Wet, J. M. J. *Dictionary of cultivated plants and their regions of*
545   *diversity: excluding most ornamentals, forest trees and lower plants*. (Pudoc, 1982).
546   at <http://agris.fao.org/agris-search/search.do?recordID=XE8282385>
547   3.   Dirzo, R. & Raven, P. H. Global state of biodiversity and loss. *Annu. Rev. Environ.*
548   *Resour.* **28,** 137–67 (2003).
549   4.   Meyer, R. S. & Purugganan, M. D. Evolution of crop species: genetics of
550   domestication and diversification. *Nat. Rev. Genet.* **14,** 840–852 (2013).
551   5.   Zeder, M. A. in *Biodiversity in Agriculture: Domestication, Evolution, and*
552   *Sustainability* (eds. Gepts, P., Famula, T. R., Bettinger, R. L. & Al, E.) (Cambridge
553   University Press, 2012).

554 6.   Clutton-Brock, J. *A natural history of domesticated mammals*. (Cambridge
555   University Press, 1999).

556 7.   Wilson, D. & Reeder, D. A. *Mammal Species of the World*. (Johns Hopkins
557   University Press, 2005).

558 8.   Gepts, P. *Crop domestication as a long-term selection experiment. Plant breeding
559   reviews* **24,** (John Wiley & Sons, Inc, 2010).

560 9.   Blomberg, S. P. & Garland Jr, T. Tempo and mode in evolution: phylogenetic
561   inertia, adaptation and comparative methods. *J. Evol. Biol.* **15,** 899–910 (2002).

562 10.   Diamond, J. Evolution, consequences and future of plant and animal domestication.
563   *Nature* **418,** 700–707 (2002).

564 11.   Salman-Minkov, A., Sabath, N. & Mayrose, I. Whole-genome duplication as a key
565   factor in crop domestication. *Nat. Plants* **2,** 1–4 (2016).

566 12.   Whitehead, S. R., Turcotte, M. M. & Poveda, K. Domestication impacts on plant–
567   herbivore interactions: a meta-analysis. *Philos. Trans. R. Soc. B Biol. Sci.* **372,**
568   20160034 (2017).

569 13.   Milla, R., Osborne, C. P., Turcotte, M. M. & Violle, C. Plant domestication through
570   an ecological lens. *Trends Ecol. Evol.* **30,** 463–469 (2015).

571 14.   Preece, C. *et al.* Were Fertile Crescent crop progenitors higher yielding than other
572   wild species that were never domesticated? *New Phytol.* **207,** 905–913 (2015).

573 15.   Darwin, C. *The variation of animals and plants under domestication.* **2,** (O. Judd,
574   1868).

575 16.   Abbo, S. *et al.* Plant domestication versus crop evolution: a conceptual framework
576   for cereals and grain legumes. *Trends Plant Sci.* **19,** 351–360 (2014).

577 17.   Harlan, J. R. *Crops and man*. (American Society of Agronomy, 1992).

578 18.   Meyer, R. S., DuVal, A. E. & Jensen, H. R. Patterns and processes in crop
579   domestication: an historical review and quantitative analysis of 203 global food
580   crops. *New Phytol.* **196,** 29–48 (2012).

581 19.   Sánchez-Villagra, M. R., Geiger, M. & Schneider, R. A. The taming of the neural
582   crest: a developmental perspective on the origins of morphological covariation in
583   domesticated mammals. *R. Soc. open Sci.* **3,** 160107 (2016).

584 20.   Larson, G. & Fuller, D. Q. The evolution of animal domestication. *Annu. Rev. Ecol.
585   Evol. Syst.* **45,** 115–136 (2014).

586 21.   Reich, P. B. The world-wide 'fast–slow' plant economics spectrum: a traits
587   manifesto. *J. Ecol.* **102,** 275–301 (2014).

588 22.   Milla, R., Morente-López, J., Alonso-Rodrigo, J. M., Martín-Robles, N. & Chapin
589   III, F. S. Shifts and disruptions in resource-use trait syndromes during the evolution
590   of herbaceous crops. *Proceedings. Biol. Sci.* **281,** 20141429 (2014).

591 23.   Tribouillois, H. *et al.* A functional characterisation of a wide range of cover crop
592   species: growth and nitrogen acquisition rates, leaf traits and ecological strategies.
593   *PLoS One* **10,** e0122156 (2015).

594 24.   Milla, R. & Matesanz, S. Growing larger with domestication: a matter of physiology,
595   morphology or allocation? *Plant Biol.* **19,** 475–483 (2017).

596 25.   Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529,** 167–171
597   (2015).

598 26.   Ricklefs, R. E. & Wikelski, M. The physiology/life-history nexus. *Trends Ecol. Evol.*
599   **17,** 462–468 (2002).

600 27.   Lovegrove, B. G. The zoogeography of mammalian basal metabolic rate. *Am. Nat.*

601   **156,** 201–219 (2000).

602   28.   Lovegrove, B. G. The influence of climate on the basal metabolic rate of small
603   mammals: a slow-fast metabolic continuum. *J. Comp. Physiol. B* **173,** 87–112
604   (2003).

605   29.   Zohary, D., Tchernov, E. & Horwitz, L. K. The role of unconscious selection in the
606   domestication of sheep and goats. *J. Zool.* **245,** 129–135 (1998).

607   30.   Cunniff, J. *et al.* Functional traits differ between cereal crop progenitors and other
608   wild grasses gathered in the Neolithic Fertile Crescent. *PLoS One* **9,** e87586 (2014).

609   31.   Biro, P. & Stamps, J. Are animal personality traits linked to life-history
610   productivity? *Trends Ecol. Evol.* **23,** 361–368 (2008).

611   32.   Dittmann, M. T. *et al.* Characterising an artiodactyl family inhabiting arid habitats by
612   its metabolism: low metabolism and maintenance requirements in camelids. *J. Arid
613   Environ.* **107,** 41–48 (2014).

614   33.   Careau, V., Bininda-Emonds, O. R. P., Thomas, D. W., Réale, D. & Humphries, M.
615   M. Exploration strategies map along fast–slow metabolic and life-history continua in
616   muroid rodents. *Funct. Ecol.* **23,** 150–156 (2009).

617   34.   Réale, D. *et al.* Personality and the emergence of the pace-of-life syndrome concept
618   at the population level. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **365,** 4051–4063
619   (2010).

620   35.   Found, R. & St. Clair, C. C. Ambidextrous ungulates have more flexible behaviour,
621   bolder personalities and migrate less. *Open Sci.* **4,** 160958 (2017).

622   36.   Tchernov, E. & Horwitz, L. K. Body size diminution under domestication:
623   unconscious selection in primeval domesticates. *J. Anthropol. Archaeol.* **10,** 54–75
624   (1991).

625   37.   Rensch, B. *Evolution above the species level*. (Columbia University Press, 1960).

626   38.   Vigne, J.-D. The origins of animal domestication and husbandry: a major change in
627   the history of humanity and the biosphere. *C. R. Biol.* **334,** 171–181 (2011).

628   39.   Li, Y. *et al.* Habitat filtering determines the functional niche occupancy of plant
629   communities worldwide. *J. Ecol.* 1–9 (2017). doi:10.1111/1365-2745.12802

630   40.   Donovan, L. A., Mason, C. M., Bowsher, A. W., Goolsby, E. W. & Ishibashi, C. D.
631   A. Ecological and evolutionary lability of plant traits affecting carbon and nutrient
632   cycling. *J. Ecol.* **102,** 302–314 (2014).

633   41.   Rotundo, J. L. & Cipriotti, P. A. Biological limits on nitrogen use for plant
634   photosynthesis: a quantitative revision comparing cultivated and wild species. *New
635   Phytol.* **214,** 120–131 (2017).

636   42.   Denison, R. F. *Darwinian agriculture: how understanding evolution can improve
637   agriculture*. (Princeton University Press, 2012).

638   43.   Miller, A. J. & Gross, B. L. From forest to field: perennial fruit crop domestication.
639   *Am. J. Bot.* **98,** 1389–1414 (2011).

640   44.   Brown, J. & West, G. *Scaling in Biology*. (Oxford Univ Press, 2000).

641   45.   West, G. B., Brown, J. H. & Enquist, B. J. General model for the origin of allometric
642   scaling laws in biology. *Science (80-.).* **276,** 122 (1997).

643   46.   Reich, P., Tjoelker, M., Machado, J.-L. & Oleksyn, J. Universal scaling of
644   respiratory metabolism, size and nitrogen in plants. *Nature* **439,** 457–461 (2006).

645   47.   Preece, C. *et al.* How did the domestication of Fertile Crescent grain crops increase
646   their yields? *Funct. Ecol.* **31,** 387–397 (2017).

647   48.   Vasseur, F., Violle, C., Enquist, B. J., Granier, C. & Vile, D. A common genetic

648   basis to the origin of the leaf economics spectrum and metabolic scaling allometry.
649   *Ecol. Lett.* **15,** 1149–1157 (2012).

650   49.   Martin, A. R. *et al.* Inter- and intraspecific variation in leaf economic traits in wheat
651   and maize. *AoB Plants* **10,** 1–14 (2018).

652   50.   Martin, A. *et al.* Intraspecific trait variation across multiple scales: the leaf
653   economics spectrum in coffee. *Funct. Ecol.* **31,** 604–612 (2017).

654   51.   Diamond, J. *Guns, Germs, and Steel: the Fates of Human Societies.* (W.W. Norton,
655   1997).

656   52.   Smythe, N. & Brown de Guanti, O. *Domestication and husbandry of the paca*
657   *(Agouti paca) (FAO conservation guide: 26).* (Food and Agriculture Organization of
658   the United Nations, 1995).

659   53.   Moreira, J. R. & Pinheiro, M. S. in *Capybara: Biology, Use and Conservation of an*
660   *Exceptional Neotropical Species* (ed. Moreira, J. R. et al.) 333–344 (Springer, 2013).

661   54.   Chamberlain, S. *et al.* taxize: Taxonomic information from around the web. (2016).

662   55.   Vincent, H. *et al.* A prioritized crop wild relative inventory to help underpin global
663   food security. *Biol. Conserv.* **167,** 265–275 (2013).

664   56.   Mansfeld, R. & Hanelt, P. *Mansfeld's Encyclopedia of Agricultural and*
665   *Horticultural Crops.* (Springer, 2001).

666   57.   Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F. S. & Golicher, D. J.
667   Taxonstand: An R package for species names standardisation in vegetation
668   databases. *Methods Ecol. Evol.* **3,** 1078–1083 (2012).

669   58.   Kissling, W. D. *et al.* Establishing macroecological trait datasets: digitalization,
670   extrapolation, and validation of diet preferences in terrestrial mammals worldwide.
671   *Ecol. Evol.* **4,** 2913–2930 (2014).

672   59.   Kattge, J. *et al.* TRY–a global database of plant traits. *Glob. Chang. Biol.* **17,** 2905–
673   2935 (2011).

674   60.   Zanne, A. E. *et al.* Three keys to the radiation of angiosperms into freezing
675   environments. *Nature* **506,** 89–92 (2014).

676   61.   Moles, A. T. *et al.* Global patterns in plant height. *J. Ecol.* **97,** 923–932 (2009).

677   62.   Glazier, D. S. & Eckert, S. E. Competitive ability, body size and geographical range
678   size in small mammals. *J. Biogeogr.* **29,** 81–92 (2002).

679   63.   Wright, I. J. *et al.* The worldwide leaf economics spectrum. *Nature* **428,** 821–827
680   (2004).

681   64.   Moles, A. T. *et al.* A brief history of seed size. *Science (80-.).* **307,** 576–580 (2005).

682   65.   Riesch, R., Martin, R. A., Lerp, H., Plath, M. & Wronski, T. Size and sex matter:
683   reproductive biology and determinants of offspring survival in Gazella marica. *Biol.*
684   *J. Linn. Soc.* **110,** 116–127 (2013).

685   66.   Jones, K. E. *et al.* PanTHERIA: a species-level database of life history, ecology, and
686   geography of extant and recently extinct mammals. *Ecology* **90,** 2648 (2009).

687   67.   Genoud, M., Isler, K. & Martin, R. Comparative analyses of basal rate of
688   metabolism in mammals: data selection does matter. *Biol. Rev.* **93,** 404–438 (2018).

689   68.   Martin, A. R. & Isaac, M. E. Plant functional traits in agroecosystems: a blueprint
690   for research. *J. Appl. Ecol.* **52,** 1425–1435 (2015).

691   69.   Nakagawa, S. & Freckleton, R. P. Missing inaction: the dangers of ignoring missing
692   data. *Trends Ecol. Evol.* **23,** 592–596 (2008).

693   70.   Penone, C. *et al.* Imputation of missing data in life-history trait datasets: which
694   approach performs the best? *Methods Ecol. Evol.* **5,** 961–970 (2014).

695 71. van Buuren, S. & Groothuis-Oudshoorn, K. mice: Multivariate imputation by
696   chained equations in R. *J. Stat. Softw.* **45,** 1–67 (2011).

697 72. Santos, T., Diniz-Filho, J. A. F. & Rangel, T. F. PVR: Computes phylogenetic
698   eigenvectors regression (PVR) and phylogenetic signal-representation curve (PSR)
699   (with null and Brownian expectations). (2013).

700 73. Bininda-Emonds, O. R. P. *et al.* The delayed rise of present-day mammals. *Nature*
701   **446,** 507–512 (2007).

702 74. Qian, H. & Jin, Y. An updated megaphylogeny of plants, a tool for generating plant
703   phylogenies and an analysis of phylogenetic community structure. *J. Plant Ecol.* **9,**
704   233–239 (2016).

705 75. Keck, F., Rimet, F., Bouchez, A. & Franc, A. phylosignal: an R package to measure,
706   test, and explore the phylogenetic signal. *Ecol. Evol.* **6,** 2774–2780 (2016).

707 76. de Villemereuil, P. & Nakagawa, S. in *Modern Phylogenetic Comparative Methods*
708   *and Their Application in Evolutionary Biology* (ed. Garamszegi, L.) 287–304
709   (Springer-Verlag, 2014).

710 77. Hadfield, J. & Nakagawa, S. General quantitative genetic methods for comparative
711   biology: phylogenies, taxonomies and multi-trait models for continuous and
712   categorical characters. *J. Evol. Biol.* **23,** 494–508 (2010).

713 78. Felsenstein, J. Phylogenies and quantitative characters. *Annu. Rev. Ecol. Syst.* **19,**
714   445–471 (1988).

715 79. Burnham, K. P. & Anderson, D. R. *Model selection and multimodel inference: a*
716   *practical information-theoretic approach.* (Springer-Verlag, 2002).
717   doi:10.1007/b97636

718 80. Hadfield, J. MCMC methods for multi-response generalised linear mixed models:
719   The MCMCglmm R package. *J. Stat. Softw.* **33,** 1–22 (2010).

720 81. Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E. & Challenger, W. GEIGER:
721   investigating evolutionary radiations. *Bioinformatics* **24,** 129–131 (2007).

722 82. Webb, C. O. & Donoghue, M. J. Phylomatic: tree assembly for applied
723   phylogenetics. *Mol. Ecol. Resour.* **5,** 181–183 (2005).

724 83. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution
725   in R language. *Bioinformatics* **20,** 289–290 (2004).

726 84. Cornwell, W. K., Schwilk, D. W. & Ackerly, D. D. A trait-based test for habitat
727   filtering: Convex hull volume. *Ecology* **87,** 1465–1471 (2006).

728 85. Barber, C. B. *et al.* geometry: mesh generation and surface tessellation. Version 0.3-
729   4. (2014).

730 86. Core Team, R. R: A language and environment for statistical computingNo Title.
731   (2016). at <https://www.r-project.org/>

732

733

734

735 **Acknowledgements**

747

748   Author contributions

749   R.M. and J.M.B. designed the study and compiled the data, R.M., J.M.B., J. C-L. and

750   M.M.T. performed statistical analyses. R.M. and J.M.B. wrote a first draft of the paper.

751   M.M.T., G.J., C.P.O. and C.V. extensively revised drafts. All authors contributed to the

752   writing of, and approved, the final version.

753

754   Competing interests

755   The authors declare no competing financial interests.

756

757 **FIGURE LEGENDS**

758 **Figure 1. Distribution of the abundance of food domesticates and frequency of**

759 **domestication events across mammalian and angiosperm families.** Length of blue bars

760 are relative abundance of domesticated species (proportion of all domesticated species that

761 are found within a given family), and of domestication frequencies (proportion of all

762 species in a family that were domesticated). Raw data can be found in Supplementary Table

763 1 and Supplementary Data 1, and family identities in the phylogeny can be browsed online

764 in the high resolution version of the Figure. Colors of tree edges correspond to

765 domestication abundances or frequencies, according to a gradient of increasing rates from

766 zero (red) to one (blue). Within each phylogeny, the inset indicates the posterior mode of

767 the phylogenetic signal (λ), together with its 95% credible interval.

768

769 **Figure 2: Phylogeographic distribution of the putative place of origin of food**

770 **domesticates included in phenotypic space analyses.** Locations in the map are medians

771 of GBIF coordinate records for the putative wild progenitor of each domesticate. Data were

772 retrievable for 168 wild progenitors of crops, out of 181 crop species, and for all of the 23

773 wild progenitors of mammal domesticates. Insets: Mantel test statistics for the correlation

774 between phylogenetic and geographic distance matrices. See Supplementary Methods for

775 further details.

776

777 **Figure 3. Phenosphospace occupancy of livestock and wild mammals.** Separate plots

778 are shown for all mammals, ruminants, and non-ruminant herbivores. Grey dots and red

779 dots are wild and domesticated mammals, respectively. Black and red polygons are convex

780 hulls for wild and domesticated mammals, respectively. Numbers in the insets are % of

781   convex hull area of domesticates outside the wild boundary (light red), of domesticates

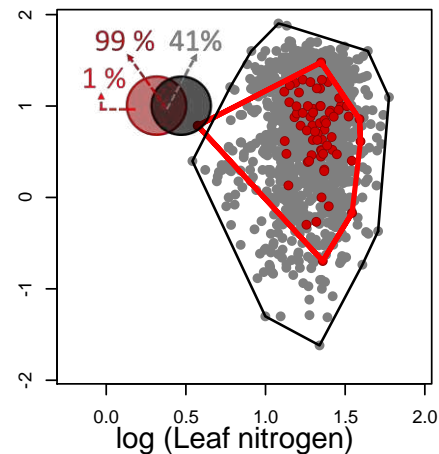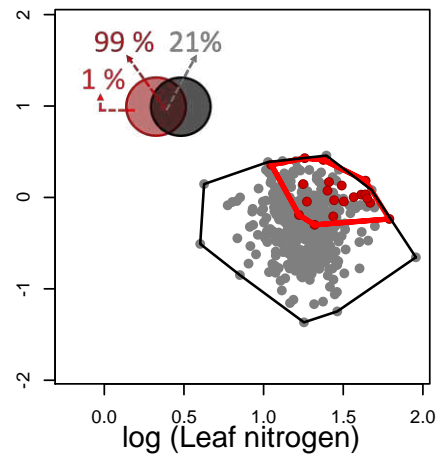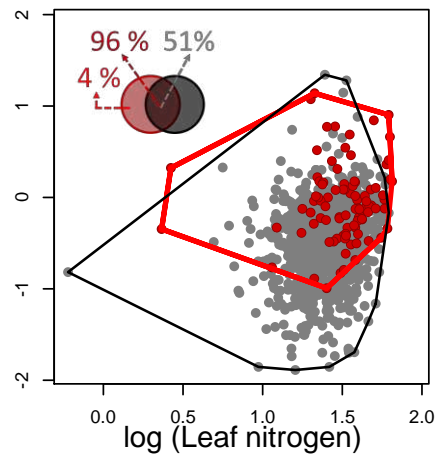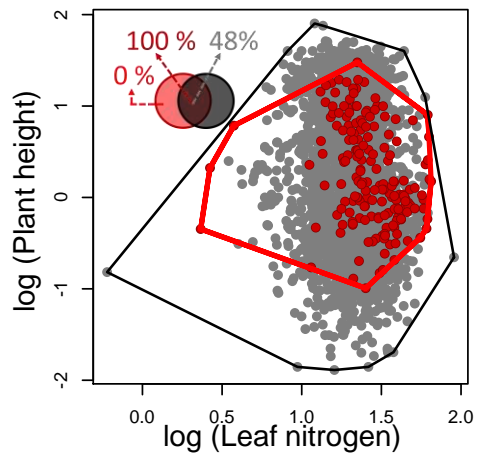782   inside the wild boundary (strong red-grey), and of wild space occupied by domesticates

783   (grey).

784

785   **Figure 4. Phenotypic differentiation between livestock and wild mammals.** Separate

786   panels are shown for all mammals, non-ruminants, and ruminants. Asterisks indicate

787   statistically significant differences at phy-P ≤ 0.05 between domestication statuses

788   according to phylogenetic Anovas (Supplementary Table 7). Center line, median; box

789   limits, upper and lower quartiles; whiskers, lowest/highest datum still within 1.5*IQR;

790   points, data exceeding whisker bounds.

791

792   **Figure 5. Phenospace occupancy of crops and wild angiosperms.** Separate plots are

793   shown for all angiosperms and for each growth form (*i.e.* herbaceous, graminoids and

794   woody). Grey dots and red dots are wild and domesticated angiosperms, respectively. Black

795   and red polygons are convex hulls for wild and domesticated angiosperms, respectively.

796   Numbers in the insets are % of convex hull area of domesticates outside the wild boundary

797   (light red), of domesticates inside the wild boundary (strong red-grey), and of wild space

798   occupied by domesticates (grey).

799

800   **Figure 6. Phenotypic differentiation between crops and wild angiosperms.** Separate

801   plots are shown for all angiosperms and each growth form (*i.e.* herbaceous, graminoids and

802   woody). Asterisks indicate statistically significant differences at phy-P ≤ 0.05 between

803   domestication statuses according to phylogenetic Anovas (Table 2). Center line, median;

804   box limits, upper and lower quartiles; whiskers, lowest/highest datum still within 1.5*IQR;

805   points, data exceeding whisker bounds.

ANGIOSPERMS

MAMMALS

ABUNDANCE OF DOMESTICATES

$\lambda = 0.006$
$(0.001 - 0.655)$

$\lambda = 0.029$
$(0.000 - 0.736)$

DOMESTICATION FREQUENCY

$\lambda = 0.010$
$(0.002 - 0.753)$

$\lambda = 0.996$
$(0.005 - 0.999)$
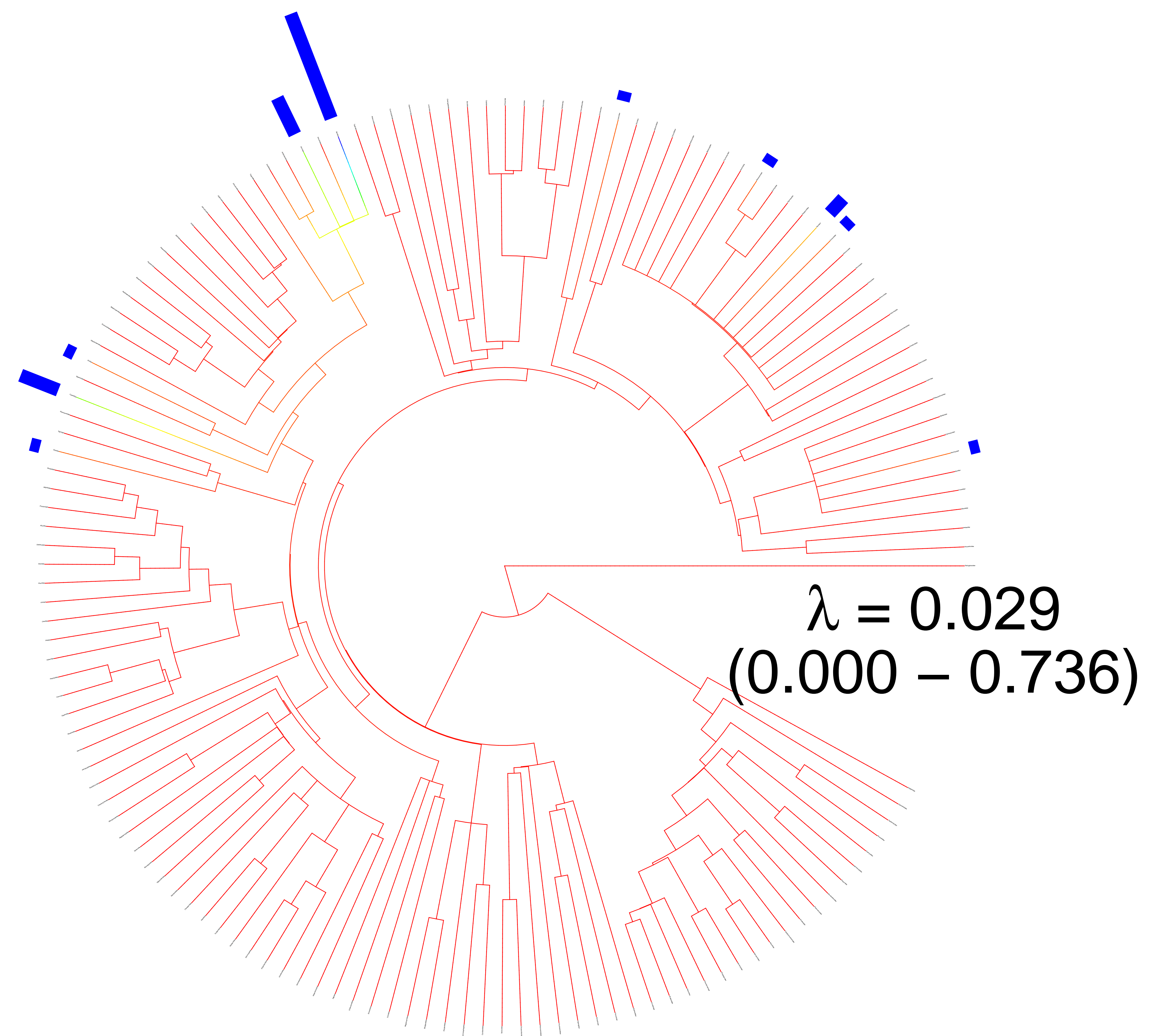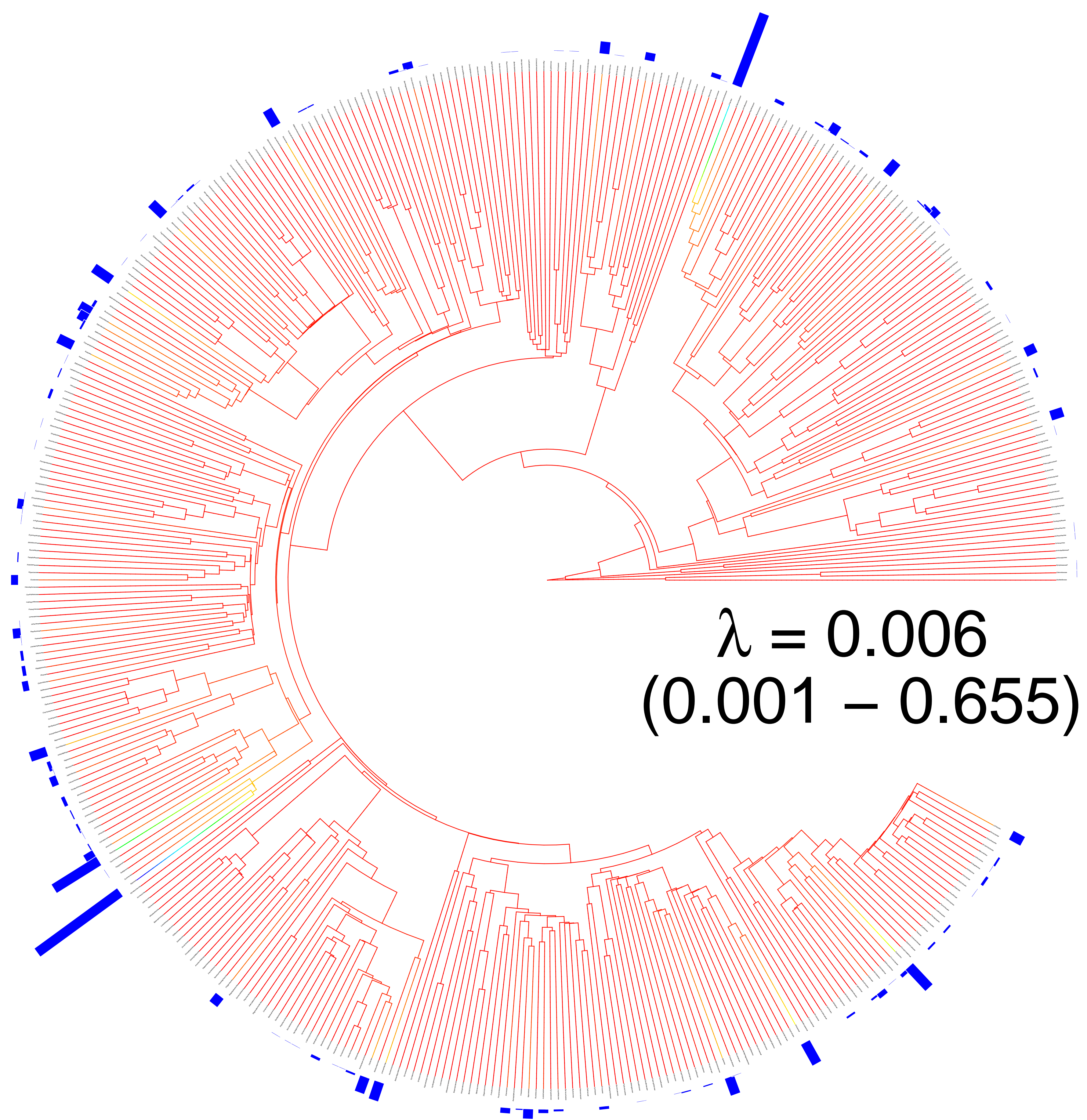
MAMMALS
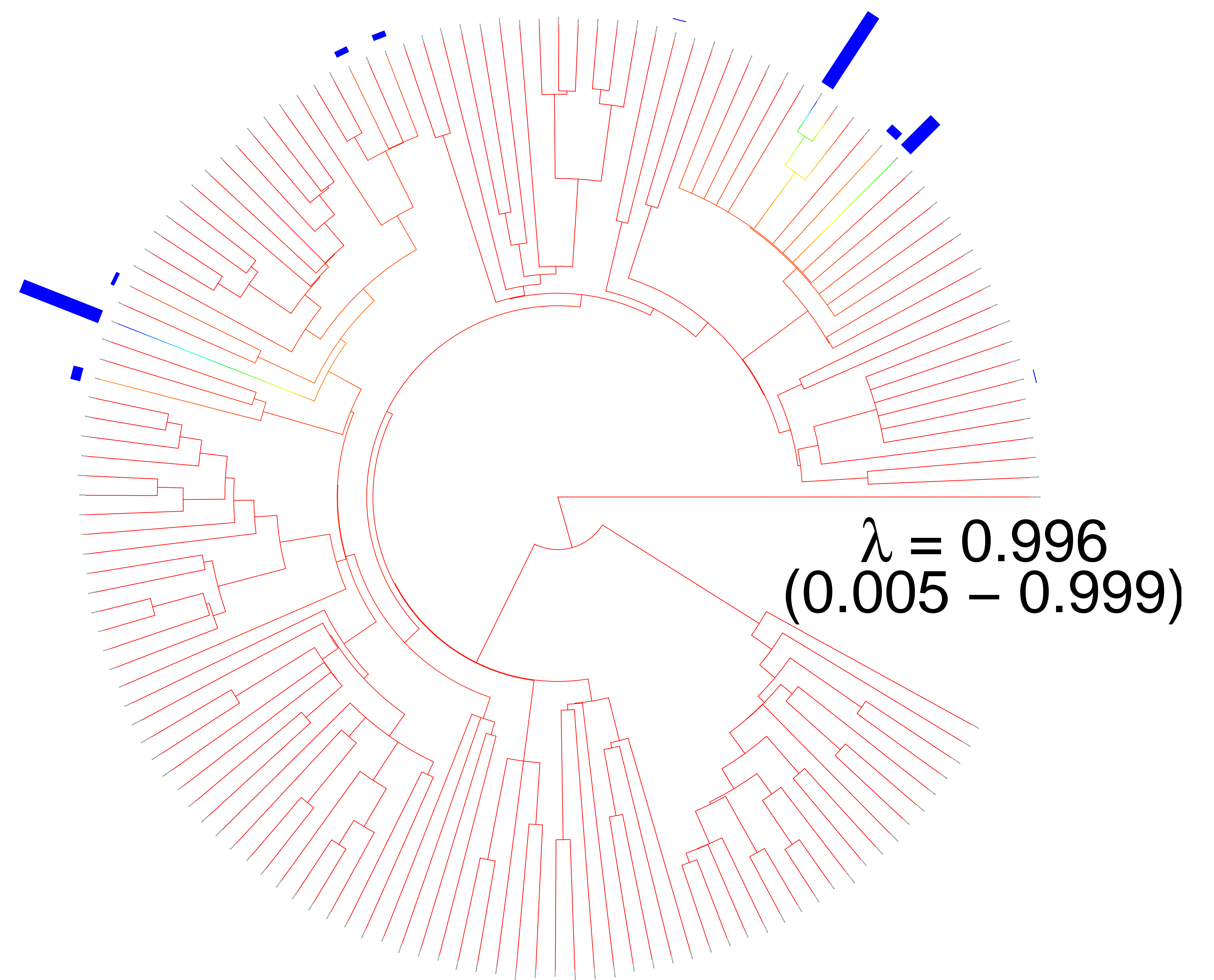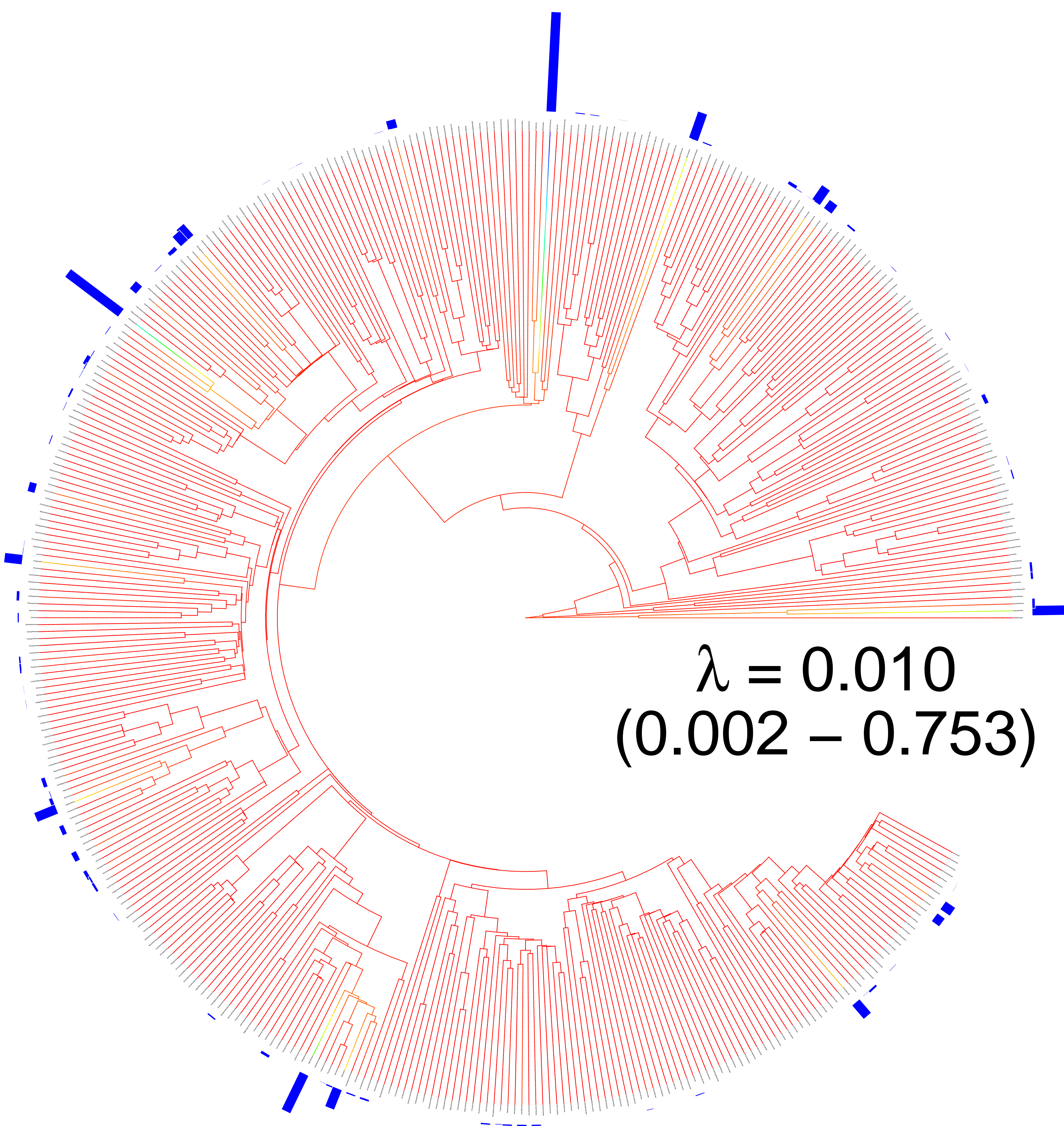
r = 0.17
p-val = 0.048

ANGIOSPERMS

r = 0.06
p-val = 0.027

	All angiosperms	Herbaceous	Graminoids	Woody

Top row (log(Plant height) vs log(Leaf nitrogen)):
- All angiosperms: 100 %, 0 %, 48%
- Herbaceous: 96 %, 4 %, 51%
- Graminoids: 99 %, 1 %, 21%
- Woody: 99 %, 1 %, 41%

Middle row (log(Seed mass) vs log(Leaf nitrogen)):
- All angiosperms: 85 %, 15 %, 71%
- Herbaceous: 79 %, 21 %, 68%
- Graminoids: 94 %, 6 %, 34%
- Woody: 97 %, 3 %, 66%

Bottom row (log(Seed mass) vs log(Plant height)):
- All angiosperms: 86 %, 14 %, 60%
- Herbaceous: 95 %, 5 %, 56%
- Graminoids: 88 %, 12 %, 24%
- Woody: 97 %, 3 %, 55%

Figure with four columns (All angiosperms, Herbaceous, Graminoids, Woody) and three rows of boxplots (log (Leaf Nitrogen), log (Plant height), log (Seed mass)) comparing Dom and Wild categories.