Evolutionary Interactions of Two Colonizing Species of Large House Spider (Araneae: *Tegenaria* spp.)

~ *Testing the Reinforcement Hypothesis*

Peter James Paul Croucher

Submitted for the Degree of Doctor of Philosophy

University of York
Department of Biology

August 1998

For my parents

Per Silvia

Tu non mi basti mai

0.0 Abstract

This thesis explores the controversial evolutionary concept of reinforcement using a model spider system. Reinforcement is the process by which natural selection strengthens prezygotic isolation between incipient species, reducing the frequency of maladaptive hybridization and hence completing reproductive isolation. The model system involved house spiders of the *Tegenaria atrica* group: *T. atrica*, *T. saeva* and *T. gigantea*. Detailed surveys confirmed the mosaic distributions of *T. saeva*, *T. gigantea* and their intermediates in the York area to which they have all only recently spread. In contrast, remarkably discrete distributions were found in southern England with *T. saeva* to the west and *T. gigantea* to the east and an abrupt parapatric boundary in Dorset. The European distributions of the *T. atrica* group were determined and phylogenetic analyses of mitochondrial DNA (mtDNA) sequences supported their taxonomic sister-species grouping, described their relationships with other *Tegenaria* spp., and provided estimates of their divergence times. MtDNA sequences demonstrated long-distance, asymmetrical introgression of *T. gigantea* haplotypes into *T. saeva* populations in southern England. Morphometrical and allozyme analyses revealed a much greater degree of hybridity in the York area compared to southern England, and indicated that introgression was largely asymmetrical with *T. saeva* populations having experienced an influx of *T. gigantea* genes. Morphometrical analyses of spiders from southern England revealed no evidence of reinforcement in parapatry, despite identifying a possible mechanism, and strongly refuted the 'lock-and-key' hypothesis for the evolution of species-specific animal genitalia. Courtship behaviour analyses were suggestive of reinforcement; behavioural displacement in parapatry was more pronounced in *T. gigantea* - agreeing with patterns of introgression. Breeding studies confirmed that previously unreported female hybrids do exist but are easily overlooked in morphometrical analyses. It is concluded that reinforcement may act to reduce the frequency of interspecific hybridization in southern England, but that the greater degree of hybridity observed in the York area may result from the recentness of contact and/or the geographical population structure.

0.1 Contents

0.0 Abstract	**3**
0.1 Contents	**4**
0.2 List of Tables	**9**
0.3 List of Figures	**9**
0.4 Acknowledgements	**11**
0.5 Author's Declaration	**11**
1 Introduction	**19**
1.1 Thesis Structure	**19**
1.2 Introductory Theory	**20**
1.2.1 Species and Hybrids	20
1.2.2 Reproductive Isolation	22
1.2.3 Hybrid Zones	23
1.2.4 Reinforcement and Character Displacement	27
1.3 The System	**31**
1.3.1 The Tegenaria atrica *group*	31
1.3.2 Interspecific Hybridization in Spiders	35
1.4 Aims of the Thesis	**37**
2 Distributions	**38**
2.1 Introduction	**38**
2.1.1 Background	38
2.1.2 Aims	40
2.2 Materials and Methods	**41**
2.2.1 European Distributional Information	41
2.2.2 British Distributions	41
2.3 Results	**43**
2.3.1 European Distributions	43
2.3.2 British Distributions	45

2.4 Discussion 51

3 Morphology and Morphometrics 54

3.1 Introduction 54

3.1.1 The Mechanical Isolation Hypothesis 54

3.1.2 Hybridization, Introgression, and Morphometrics 55

3.1.3 Fluctuating Asymmetry 57

3.1.4 Aims 59

3.2 Materials and Methods 60

3.2.1 Samples and Measurements 60

3.2.2 Merrett Plots of Conductor + Tegulum Length against Carapace (Prosoma) Length 64

3.2.3 Measurement Error 65

3.2.4 Univariate Analysis 65

3.2.5 Multivariate Analysis 66

3.2.6 Fluctuating Asymmetry Analysis 67

3.3 Results 70

3.3.1 Merrett Plots of Conductor + Tegulum Length against Carapace (Prosoma) Length 70

3.3.2 Univariate Analysis and Measurement Error 77

3.3.3 Multivariate Analysis 81

3.3.4 Fluctuating Asymmetry 91

3.4 Discussion 94

3.4.1 Hybridization and Introgression 94

3.4.2 Reinforcement/Character Displacement of Morphology? 99

3.4.3 Fluctuating Asymmetry 100

4 Allozyme Markers in *T. saeva* and *T. gigantea* 103

4.1 Introduction 103

4.1.1 Background 103

4.1.2 Aims 108

4.2 Materials and Methods **109**

4.2.1 Samples 109

4.2.2 Homogenization 110

4.2.3 Allozyme Electrophoresis 110

4.2.4 The Enzymes 112

4.2.5 Analyses 113

4.3 Results **117**

4.3.1 Allele Frequencies and Distributions 117

4.3.2 Genotypic Overlap 126

4.3.3 Genetic Distances and 'Population Trees' 133

4.4 Discussion **138**

5 Mitochondrial DNA Species Markers, Introgression, and Phylogeny **141**

5.1 Introduction **141**

5.1.1 Aspects of Mitochondrial Genetics and Function 141

5.1.2 Mitochondrial DNA as a Phylogenetic Marker 146

5.1.3 Mitochondrial DNA as a Measure of Gene Flow and a Marker of Introgression 148

5.1.4 Aims 150

5.2 Materials and Methods **152**

5.2.1 Sample Collection and Storage 152

5.2.2 DNA Extraction 153

5.2.3 Amplification using the Polymerase Chain Reaction 154

5.2.4 Purification of PCR Products 159

5.2.5 Gel Electrophoresis of DNA 160

5.2.6 Sequencing 161

5.2.7 Phylogenetic Analysis 164

5.2.8 Relative Rate of CO1 Evolution 171

5.3 Results A: Mitochondrial DNA Phylogeny **172**

5.3.1 DNA Extraction, PCR, and Sequencing 172

5.3.2 Sequence Characteristics 175

5.3.2.1 CO1 175

5.3.2.2 16S 190

5.3.2.3 ND1 194

5.3.2.4 Combined CO1 and ND1 Fragments 205

5.3.3 mtDNA Sequence Variation 205

5.3.3.1 CO1 205

5.3.3.2 16S 207

5.3.3.3 ND1 208

5.3.3.4 Combined CO1 and ND1 Fragments 209

5.3.4 Phylogeny Reconstruction 210

5.3.4.1 CO1 - Parsimony Analysis 210

5.3.4.2 CO1 - Distance Analysis 215

5.3.4.3 Relative Rate Tests of CO1 Evolution 217

5.4.4.4 ND1 - Parsimony Analysis 218

5.4.4.5 ND1 - Distance Analysis 219

5.3.4.6 Combined CO1 and ND1 Fragments - Parsimony Analysis 222

5.3.4.7 Combined CO1 and ND1 Fragments - Distance Analysis 225

5.4 Results B: A Putative Pseudogene of CO1 in *Tegenaria parietina* **228**

5.4.1 Sequence Characteristics 228

5.4.2 Relative Rate Tests of CO1 Pseudogene Evolution 234

5.4.3 Phylogenetic Analysis 234

5.5 Discussion **238**

5.5.1 *Phylogenetic Inferences and Introgression* 238

5.5.2 *Divergence* 241

5.5.3 *The Putative Pseudogene of CO1 in* T. parietina 245

6 Conspecific and Heterospecific Crosses: Comparative Analyses of Courtship Behaviour and Reproductive Success in *T. saeva* and *T. gigantea*. **248**

6.1 Introduction **248**

6.1.1 Spider Mating Behaviour 248

6.1.2 Courtship 248

6.1.3 Copulation 253

6.1.4 Courtship and Reproductive Isolation 253

6.1.5 Aims 257

6.2 Materials and Methods **258**

6.2.1 Sample Collection, Husbandry, and Experimental Procedure 258

6.2.2 Transition Matrix Analyses of Pre-copulatory Courtship 261

6.2.3 Duration of Courtship and Copulatory Parameters 264

6.3 Results **266**

6.3.1 Courtship and Mating Behaviour 266

6.3.2 Conspecific and Heterospecific Crosses 271

6.3.3 Transition Matrix Analysis 276

6.3.4 Duration of Courtship and Copulatory Parameters 302

6.4 Discussion **309**

6.4.1 Courtship 309

6.4.2 Mechanical Isolation 313

6.4.3 Character Displacement/Reinforcement? 315

7 General Discussion **317**

7.1 The Tegenaria atrica *group* 317

7.2 Hybridization and Introgression in the T. atrica *group* 318

7.3 Postzygotic Isolation 327

7.4 Prezygotic Isolation, Reinforcement, and Character Displacement 330

7.5 Conclusions 337

Appendices **339**

A.1 Allozyme Reagents, Procedures, and Genetic Distances (Chapter 4) **339**

A.2 Molecular Reagents (Chapter 5) **344**

A.2.1 DNA Extraction 344

A.2.2 Gel Electrophoresis 345

A.3 Behavioural Transitions (Chapter 6) **346**

Glossary **349**

References **354**

0.2 List of Tables

1.1	British representatives of the Agelenidae (after Roberts, 1993).	32
3.1	Regression equations derived from plots of tegulum + conductor length (y) against prosoma length (x).	73
3.2	Comparison of residual mean squares from regression for each of the tegulum + conductor length against prosoma length data sets.	73
3.3	Tegulum + conductor length means and standard errors adjusted by ANCOVA on prosoma length, with 95% comparison limits.	75
3.4	Summary statistics for all variables measured on males.	78
3.5	Summary statistics for all variables measured on females.	80
3.6	Multiple group principle component coefficients (PCs) for the allopatric *T. saeva* and allopatric *T. gigantea* data sets from southern England: MALES.	82
3.7	Multiple group principle component coefficients (PCs) for the allopatric *T. saeva* and allopatric *T. gigantea* data sets from southern England: FEMALES.	83
3.8	Discriminant function coefficients derived from the allopatric *T. saeva* and allopatric *T. gigantea* data sets from southern England after transformation to principle components.	85
3.9	Descriptive statistics for the discriminant function scores of each data set.	90
3.10	Summary of fluctuating asymmetry in four variables measured on males and results of Levene's Tests for heterogeneity of variances between samples.	92

4.1	Allele frequencies at the four variable loci in the sample zones from southern England.	122
4.2	Allele frequencies at the four variable loci in the samples from the York area.	122
4.3	Log-likelihood ratio tests of heterogeneity for the four variable loci.	125
4.4	Observed and expected genotype frequencies in the sample zones from southern England.	127
4.5	Observed and expected genotype frequencies in the samples from the York area.	128
4.6	Pairwise measures of genotypic overlap (S_g) and probability of a correct diagnosis (P_c) between any two samples.	128
4.7	Mantel tests on different measures of genetic distance.	134
5.1	Individuals sequenced, identification, location of origin, and haplotypes for the CO1, 16S, ND1, and combined CO1 and ND1 fragments.	176
5.2	Alignment for the 325 bp of CO1 in the 5'->3' direction for the J strand.	178
5.3	Nucleotide frequencies at each codon position for the CO1 fragment.	188
5.4	The frequency of unambiguous nucleotide changes in the single MP tree for CO1.	188
5.5	Alignment for the 438 bp of 16S in the 5'->3' direction for the N strand.	191
5.6	Alignment for the 543 bp of ND1 in the 5'->3' direction for the N strand.	195
5.7	Nucleotide frequencies at each codon position for the coding region of the ND1 fragment.	204

5.8	The frequency of unambiguous nucleotide changes in the single MP tree for ND1.	204
5.9	The frequency of unambiguous nucleotide changes in the MP tree for the combined CO1 and ND1 data.	205
5.10	Pairwise distances among the nine mt CO1 haplotypes analysed.	206
5.11	Pairwise distances among the three mt 16s haplotypes analysed.	208
5.12	Pairwise distances among the six mt ND1 haplotypes analysed.	209
5.13	Pairwise distances among the six combined mt CO1 and ND1 haplotypes analysed.	210
5.14	Relative rate test for CO1.	217
5.15	Alignment of the putative *T. parietina* CO1 pseudogene against the entire *T. parietina* CO1 fragment.	229
5.16	Pairwise distances among the nine mt and one pseudogene CO1 haplotypes analysed.	233
5.17	Comparison of the 53 repeated base pairs in the CO1 pseudogene.	233
5.18	Relative rate test for the CO1 pseudogene.	234
6.1	Male and female responses in each pairing combination.	273
6.2	Reproductive success in each pairing combination.	274
6.3	ANOVA: Mean number of eggs per egg-sac for each pairing combination.	275
6.4	Allopatric *T. saeva* female x *T. saeva* male (TSf x TSm) transition matrix (n = 10).	278
6.5	Allopatric *T. gigantea* female x *T. gigantea* male (TGf x TGm) transition matrix (n = 12).	279

6.6 Allopatric *T. gigantea* female x *T. saeva* male (TGf x TSm) transition matrix (n = 11). 280

6.7 Allopatric *T. saeva* female x *T. gigantea* male (TSf x TGm) transition matrix (n = 9). 281

6.8 Parapatric *T. saeva* female x *T. saeva* male (TSf x TSm) transition matrix (n = 12). 282

6.9 Parapatric *T. gigantea* female x *T. gigantea* male (TGf x TGm) transition matrix (n = 8). 283

6.10 Parapatric *T. gigantea* female x *T. saeva* male (TGf x TSm) transition matrix (n = 8). 284

6.11 Parapatric *T. saeva* female x *T. gigantea* male (TSf x TGm) transition matrix (n = 10). 285

6.12 Summary of courtship behaviours and their acronyms. 286

6.13 Pairwise comparisons between the transition matrices based on columns totals. 295

6.14 Significant transitions between transition matrices. 299

6.15 Mean duration of courtship and copulatory parameters for different types of cross. 303

A.1.1 The chemicals used in the electrophoretic analysis. 340

A.1.2 The electrophoretic method for each of the four polymorphic enzymes. 341

A.1.3 Pairwise genetic distances from allozyme data: D_g and Roger's distances. 342

A.1.4 Pairwise genetic distances from allozyme data: Nei's and Cavalli-Sforza distances. 342

A.1.5 Pairwise genetic distances from allozyme data: Reynolds distances. 343

A.3.1 ALLOPATRIC PAIRINGS: Preceding acts that significantly facilitate or inhibit following acts ($P < 0.01$). 347

A.3.2 PARAPATRIC PAIRINGS: Preceding acts that significantly facilitate or inhibit following acts (P < 0.01). 348

0.3 List of Figures

1.1 Tension zones move to minimize their length. 26

1.2 Coincident and concordant clines. 26

1.3 Scanning electron micrographs showing the combined tegulum + conductor structure of the palps. 34

2.1 Distributions of A. *T. gigantea*; B. *T. saeva*; C. *T. atrica* in Britain and Ireland. 39

2.2 European distribution of the *T. atrica* group. 44

2.3 Distribution of *T. saeva* and *T. gigantea* samples in southern England. 46

2.4 Distribution of *T. saeva* and *T. gigantea* samples at the contact zone in southern England. 47

2.5 Distribution of *T. saeva* and *T. gigantea* samples in the York area. 49

2.6 Distribution of *T. saeva* and *T. gigantea* samples in the York area (York detail). 50

3.1 Male measurements used in the morphometrical analyses. 62

3.2 Female measurements used in the morphometrical analyses. 63

3.3 Scatterplots of conductor + tegulum length against prosoma length. 71

3.4 Plots of regression lines of conductor + tegulum length against prosoma length. 72

3.5 Plot of tegulum + conductor length means, adjusted for prosoma length by ANCOVA, with 95% comparison intervals. 76

3.6	Histograms of the discriminant functions scores for males.	87
3.7	Histograms of the discriminant functions scores for females.	88
3.8	Plot of fluctuating asymmetry (FA) in male palpal ventral tegulum length (TLV) for all data sets.	93
4.1	Southern England sample sites, sample zones, and allele frequencies.	118
4.2	Southern England sample sites in the contact area.	119
4.3	York area sample sites and allele frequencies.	120
4.4	York area sample sites (close-up on York).	121
4.5	Electromorph patterns for the four variable loci.	124
4.6	The behaviour of the genotypic overlap, S_g, between two theoretical populations, 1 and 2, for a single locus with two alleles, *A* and *B*.	130
4.7	Genotypic overlaps of samples with zone 1&2 and zone 5&6 (southern England samples), and with the 'good' *T. saeva* and the 'good' *T. gigantea* (York area samples).	132
4.8	Neighbor-joining trees for all sample groups generated from A) Dg, and B) Rogers' (1972) genetic distance with Wright's (1978) modification.	135
4.9	Neighbor-joining trees for all sample groups generated using A) Cavalli-Sforza and Edwards chord distance (1967), B) Nei's (1972) genetic distance, C) Reynolds, Weir and Cockerham's (1983) genetic distance.	136
5.1	The organization of the *Drosophila yakuba* mitochondrial DNA molecule.	142
5.2	Strategy used for the amplification of the mtDNA gene fragments.	157
5.3	Agarose gel electrophoresis of DNA extractions.	172

5.4	Agarose gel electrophoresis of PCR products (after purification).	173
5.5	A section from a typical manual sequencing autoradiograph.	173
5.6	A section from a typical automated sequencer electropherogram.	174
5.7	Distribution of substitutions among the three codon positions in the mt CO1 fragment.	186
5.8	Number of substitutions of each class in the mt CO1 fragment.	188
5.9	A two-dimensional model of the invertebrate (insect) CO1 gene.	189
5.10	Distribution of substitutions among the three codon positions in the mt ND1 fragment.	203
5.11	The number of substitutions in each class in the mt ND1 fragment in members of the *T. atrica* group.	203
5.12	Frequency distribution of the lengths of 10 000 random trees for the CO1 data.	211
5.13	Maximum parsimony (MP) tree for the CO1 data.	212
5.14	Neighbor-joining tree for the CO1 data.	216
5.15	Frequency distribution of the lengths of 10 000 random trees for the ND1 data.	218
5.16	Maximum parsimony (MP) tree for the ND1 data.	220
5.17	Neighbor-joining tree for the ND1 data.	221
5.18	Frequency distribution of the lengths of 10 000 random trees for the combined CO1 and ND1 data.	223
5.19	Maximum parsimony (MP) tree for the combined CO1 and ND1 data.	224
5.20	Neighbor-joining tree for the combined CO1 and ND1 data.	226
5.21	Frequency distribution of the lengths of 10 000 random trees for the CO1 pseudogene data.	235

5.22	Maximum parsimony (MP) tree for the CO1 pseudogene data.	236
5.23	European Quaternary climatic stages and divergence dates for the *T. atrica* group from mitochondrial CO1 divergence data.	243
6.1	A female *Tegenaria saeva* displaying an aggressive warning posture to an approaching male *T. gigantea*.	270
6.2	A pair of *T. saeva 'in copula'.*	270
6.3	Sequence diagrams of pre-copulatory courtship transitions in allopatric and parapatric TSf x TSm pairings.	287
6.4	Sequence diagrams of pre-copulatory courtship transitions in allopatric and parapatric TGf x TGm pairings.	289
6.5	Sequence diagrams of pre-copulatory courtship transitions in allopatric and parapatric TGf x TSm pairings.	291
6.6	Sequence diagrams of pre-copulatory courtship transitions in allopatric and parapatric TSf x TGm pairings.	292
6.7	A radial dendrogram showing the relationship between courtship behaviours.	296
6.8	Behaviours significantly heterogeneous in total frequency between transition matrices.	297
6.9	Mean palpal insertion duration as a function of female/male genital size: conspecific pairings.	306
6.10	Mean palpal insertion duration as a function of female/male genital size: heterospecific pairings.	308

0.4 Acknowledgements

I wish to gratefully acknowledge the excellent and enthusiastic supervision, support, guidance, and friendship of Dr Geoff Oxford and Dr Jeremy Searle. There are far too many people deserving acknowledgement to name specifically. Thanks are due to the members of the British Arachnological Society who provided specimens and enthusiasm, to all members of the public who collected spiders, to the BBC and local newspapers in York, Dorset and Hampshire for publicity, and to everybody who kindly provided spider depositories during the surveys. Special thanks go to R. D. Jones, P. Merrett, M. Emerit (Montpellier), R. Leborgne (University of Nancy), J O'Connor (National Museum of Ireland), and G. Binford (University of Arizona). I am indebted to Dr Rosemary Gillespie and Dr George Roderick (University of Hawaii), Dr Marshal Hedin (University of Arizona), and Dr Carles Ribera (University of Barcelona), for their advice and help. In York, I thank the academic and support staff, in particular Dr Angela Douglas, Dr Terry Crawford, Dr Calvin Dytham, Dr John Sparrow, Dr Peter Hogarth, Prof. Peter Young, Prof. John Currey, John Byrne, Allan Mould, Kevin Brear, Paul Waites and Bryan Adamson. I am also grateful for the love and support of my colleagues and friends: Islam Gündüz, Sue Haynes, Alex Lewis, Maarit Jaarola, Cara Morgan, Vicki Jenkins, Ian Gibson, Lucinda Birkle, Ben Raymond, Sarah Bence, Sarah Turner, Ian Barnes, Ros Stockley, Colin Tosh, Jerry Kemp, Jane Gaskin, Andrea Verardi, Patricia Mirol, David Bilton, Chris Palmer and Thorun Helgason, among very many others. Unwaivering love and support came from Silvia Mascheretti - *ti ringrazio*, and from my family. I am grateful to Fritz Vollrath (University of Århus) for advice and support in completing this thesis. This research was supported by a Biotechnology and Biological Sciences Research Council Studentship.

0.5 Author's Declaration

I declare that the work presented in this thesis is the original sole effort of the author and has not previously been presented for examination in any form.

1 Introduction

"The spider taketh hold with her hands, and is in kings' palaces."
- *Proverbs of Solomon.*

"...I never expect to make a detailed study of House Spiders - for whom I have no affection."
- W. S. Bristowe, *The World of Spiders.*

"Almost all spiders are rather nice-looking. I'm not as flashy as some, but I'll do." (Charlotte).
- E. B. White, *Charlotte's Web.*

1.1 Thesis Structure

This thesis is about speciation and about spiders. The evolutionary interactions of closely related large house spiders of the *Tegenaria atrica* group are explored in relation to speciation theory, in particular to look for evidence of reinforcement. Speciation is a substantial topic central to the understanding of a major component of biodiversity: the evolution of species diversity. The study of speciation therefore sits at the heart of evolutionary and population genetics and draws upon all aspects of organismal biology including natural history, biogeography, behaviour, phylogenetics and systematics. This thesis reflects the scale and breadth of speciation problems and a number of very different approaches to understanding the species interactions of the *T. atrica* group (in particular *T. saeva* and *T. gigantea*) have been employed. Consequently, the thesis does not follow the 'traditional' format of a long general introduction, followed by a general methods chapter, with subsequent results chapters. Rather it starts with a short general introduction focusing on the relevant areas of speciation theory (namely species, hybrid zones, reinforcement and character displacement), and an introduction to the study organisms. Chapter 2 is short and describes the distribution of the *T. atrica* group (based upon the literature and extensive field surveys). This chapter might be viewed as an extension of the general introduction as it is the curious distributions of *T. saeva* and *T. gigantea* that have propelled this project. Each experimental

chapter that follows is largely self-contained, with an introduction (sometimes necessarily long), methods section, results and discussion. The thesis ends with a final discussion in which the results are placed into the context of current theory. A glossary of selected terms (mostly arachnological and phylogenetic) has been provided to help the reader.

1.2 Introductory Theory

1.2.1 Species and Hybrids

Any study of the processes connected with speciation must begin with the underlying assumption that species are real biological entities and not simply subjective human divisions of what is really a natural continuum among organisms (Coyne and Orr, 1998). The fact that independent observers consistently recognise distinct groups living in sympatry yet separated by genetic and phenotypic gaps, provides the strongest evidence for the reality of species. It suggests a species concept based upon mechanisms that prevent interbreeding, thus allowing these groups to maintain their genetic identity (Coyne and Orr, 1998). Hence Dobzhansky's (1935) and Mayr's (1942) biological species concept (BSC) is nowadays generally couched in terms of gene flow such that species are considered to be groups of populations that are reproductively isolated from other such groups by genetically based traits that prevent gene exchange (Coyne and Orr, 1998). These genetic traits are known as isolating mechanisms and are introduced below. The reality of species might seem obvious (obviousness is not always a good indicator of reality) but a universally useful definition of species has proved elusive. Ridley (1993) lists seven species concepts, and Mayden (1997) notes at least 22 in current use by biologists. Different species concepts may prevail because they are operational under different circumstances (for instance a palaeontologist would find a morphologically based species concept more practical than a worker studying bacterial phylogenetics) - but it can be argued that simply being practically

useful is not really a sufficient basis for a conceptual framework (Hull, 1997; Mishler and Donoghue, 1982). This is not the place to discuss the merits of different species concepts or the existence of species (excellent reviews and comment, both practical and philosophical, can be found in: Claridge *et al.*, 1997; Cracraft, 1989; Hull, 1978; 1997; Mayden, 1997; Minelli and Foddai, 1997; Mishler and Donoghue, 1982; Nelson, 1989; Ridley, 1993; Templeton, 1989, among others). The BSC is *loosely* adopted in this thesis, which is essentially a study of reproductive isolation, because it consistently proves to be the most useful (and because the study of isolation mechanisms by default implies the BSC!). As Coyne and Orr (1998) note: "every recent study on the 'genetics of speciation' is an analysis of reproductive isolation". The BSC is applied loosely in that it is accepted that some gene flow *can* occur between species without negating their specific status and that gene flow may be an important feature in the completion of reproductive isolation (see Chapter 7, section 7.3).

The term 'hybrid' (the product of hybridization) is defined according to Arnold (1997) (modified from the definition of Harrison (1993)): 'a natural hybrid derives from crosses in nature between individuals from two populations, or groups of populations, which are distinguishable on the basis of one or more heritable characters'. Several advantages of this definition have been outlined (Arnold, 1997; Harrison, 1993). First, the definition does not depend on the acceptance of any particular species concept (see above). Secondly, the populations from which the hybridizing individuals derive do not have to be assigned to particular taxonomic categories (for example, species or races). Thirdly, the relative fitness of hybrids or the adaptive norms of parental types need not be known. Fourthly, the definition is empirical - it can be tested whether the individuals involved in putative hybridization events come from populations that are diagnostically different in one or more heritable characters.

1.2.2 Reproductive Isolation

Reproductive isolating mechanisms are divided into those that act before fertilization (prezygotic isolating mechanisms) and those that act after fertilization (postzygotic isolating mechanisms).

In animals, prezygotic mechanisms include (Ridley, 1993; after Dobzhansky, 1970):
a) Ecological isolation (differences in habitat).
b) Temporal isolation of mating seasons (phenology).
c) Sexual isolation (lack of attraction between the sexes of the different species).
d) Mechanical isolation (physical copulatory incompatibility).
e) Gametic isolation (in animals with internal fertilization the gametes of one species may be inviable in the ducts of the other species).

Postzygotic mechanisms include:
a) Hybrid zygote inviability.
b) Hybrid sterility (F_1 hybrids of one or both sexes fail to produce functional gametes).
c) Hybrid breakdown (F_2 or backcross hybrids suffer reduced viability or fertility).

Postzygotic isolation occurs when hybrids are unfit (Coyne and Orr, 1998). It is a major assumption of this thesis - which is primarily an investigation into *prezygotic* isolation - that the hybrids produced in the system under study do indeed suffer unfitness. The validity, or otherwise, of this assumption, and relevant models of postzygotic isolation, will be discussed later (Chapter 7) in light of the data presented throughout this thesis. For now, it is sufficient to note that postzygotic isolation will generally occur as a direct result of genetic differences at the cellular and molecular levels, and that these differences fall into three categories: chromosomal rearrangements, ploidy level differences, and different alleles that do not function properly together when united in the

mixed genetic background of a hybrid (Coyne and Orr, 1998); with the latter being by far the most important in animal species (Coyne and Orr, 1998). Having said this it should also be noted that hybrids may be unfit for reasons other than direct genetic effects. Hybrids may be perfectly viable and fertile but unfit if they are ecologically or behaviourally maladapted. Such a case is found in the two species of wolf spider (Lycosidae) *Schizocosa ocreata* and *S. rovneri*. These species are behaviourally isolated; not only do females of both species tend to reject the courtship of heterospecific males, but F_1 hybrid males are rejected by females of both species and F_1 hybrid females reject males of both species (hybrids produced by 'forced' copulation of anaesthetized females) (Stratton, 1997; Stratton and Uetz, 1981; 1983; 1987; Uetz and Stratton, 1982. Also see Chapter 6). In another example, hybrids between the butterflies *Heliconius erato* and *H. himera* appear to suffer decreased fitness from being poorly adapted to either of the biotopes present where the species meet, and from enhanced frequency-dependent predation as a result of their intermediate warning coloration patterns (McMillan *et al.*, 1997).

1.2.3 Hybrid Zones

Hybrid zones are regions where two taxa meet, mate, and hybridize (Barton and Hewitt, 1985; Hewitt, 1990). More specifically, hybrid zones can be defined as regions where 'two populations of individuals that are distinguishable on the basis of one or more heritable characters overlap spatially and temporally and cross *to form viable and at least partially fertile offspring'* (Arnold, 1997, Harrison, 1990; 1993). The italicised section was an addition by Arnold (1997) to Harrison's (1990; 1993) definition in recognition of the fact that even when there is *very* strong postzygotic isolation it is unlikely that 100% of F_1 offspring will be *completely* sterile - rare events are important in evolution (Arnold, 1997). Hybrid zones are often very narrow (only a few hundred metres) yet may be several hundred kilometres in length. They have been described in a wide variety of organisms, with the number and

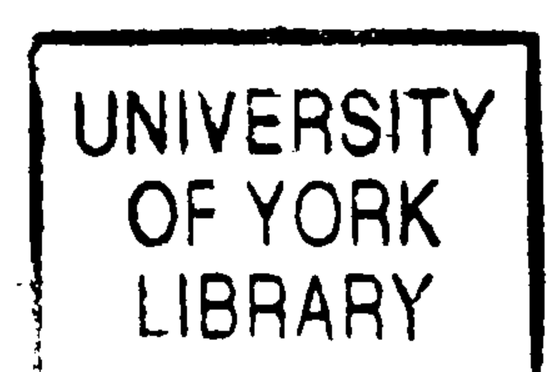
type of characters used to differentiate the two taxa, and which have been demonstrated to mix in the hybrid zone, varying greatly (Hewitt, 1990). For example, the *Heliconius* butterflies mentioned above not only form hybrid zones between species but also between races with differing colour patterns; these distinct Müllerian mimicry patterns are controlled by only a few major genes (Hewitt, 1990; McMillan *et al.*, 1997; Turner, 1982). In orthopteran hybrid zones many characters have been demonstrated to be involved - DNA and enzyme variation, chromosomes, morphology, physiology and behaviour (Harrison, 1986; Harrison and Rand, 1989; Hewitt, 1989; Hewitt, 1990). The wide range of genotypes found in hybrid zones can be used to understand the genetic differences and selective forces separating the taxa involved and help in the understanding of isolation; indeed some models of parapatric speciation (see below) involve the formation, movement, and modification of hybrid zones (Barton and Hewitt, 1985). In considering hybrid zones it is important to think in terms of genes because hybridizing genomes will tend to be broken down by segregation and recombination, and different genes will experience different patterns of selection and therefore have different fates. This mixing also means that in the middle of a hybrid zone there is typically little chance of finding a pure parental (= 'racial' or incipient species) genotype or F_1 genotype, and therefore to think in terms of parentals and hybrids and not genes can be misleading (Hewitt, 1989). Another way of viewing this would be that all characters differing between the taxa should be considered: so that the degree of hybridity is examined. If there is only one character (for instance a chromosomal rearrangement), then there will only be the two parentals and the F_1 hybrids.

Whether the differences between the hybridizing taxa arose *in situ* (i.e. in primary contact) or arose in allopatry, followed by secondary contact (possibly followed by the accumulation of further differences), is generally difficult to decide. Indeed over geological time both processes may have occurred in many cases and contributed to extant differences (Harrison, 1990; 1993; Hewitt, 1990). However, it seems likely that most hybrid zones formed

by secondary contact after the most recent glaciation, following range expansion from refugia - both in the temperate and probably the tropical regions (Endler, 1982; Haffer, 1982; Hewitt, 1990; 1996).

Hybrid zones have commonly been equated with clines (Barton and Hewitt, 1985). Clines are simply gradients in the frequency of certain characters. In the case of a hybrid zone, these would be gradients in the frequency of characters typical of each parental taxon, across the area of contact (Harrison, 1993). Hewitt (1989) notes that "a cline between two races fixed for different alleles at a locus could be the result of 1) heterozygote disadvantage, 2) differential environmental selection on the two homozygotes, 3) selective equality of homozygotes and heterozygotes, 4) frequency-dependent selection on homozygotes, 5) superiority of one homozygote over the other, or 6) superiority of heterozygotes over homozygotes in the particular environment of the zone". He also notes that "where several genes differ between the races the recombinants may be disadvantaged if there is coadaptive epistasis between alleles of the same race, which will produce clines of type 1". There seems to be little strong evidence for types 5 or 6 and although the others can explain some cases, it seems that most cases of hybrid zones in the literature are of type 1 (Barton and Hewitt, 1985; Hewitt, 1989; 1990). Such clines are maintained by a balance between dispersal and selection against hybrids and are not maintained by responses to local environmental conditions; they can therefore move. The clines tend to move so that their length is minimized (Figure 1.1) and are therefore known as 'tension zones' (Barton and Hewitt, 1985). From this point forward 'hybrid zone' will generally be taken to imply a tension zone.

An important quality of tension zones is that they tend to coalesce to form coincident and often concordant clines in different characters (Figure 1.2) (Barton and Hewitt, 1985; Hewitt, 1989). There are a number of explanations for the coincidence of clines (Hewitt, 1989): repeated range contractions and expansions into secondary contact, the trapping of hybrid zones by dispersal


UNIVERSITY
OF YORK
LIBRARY

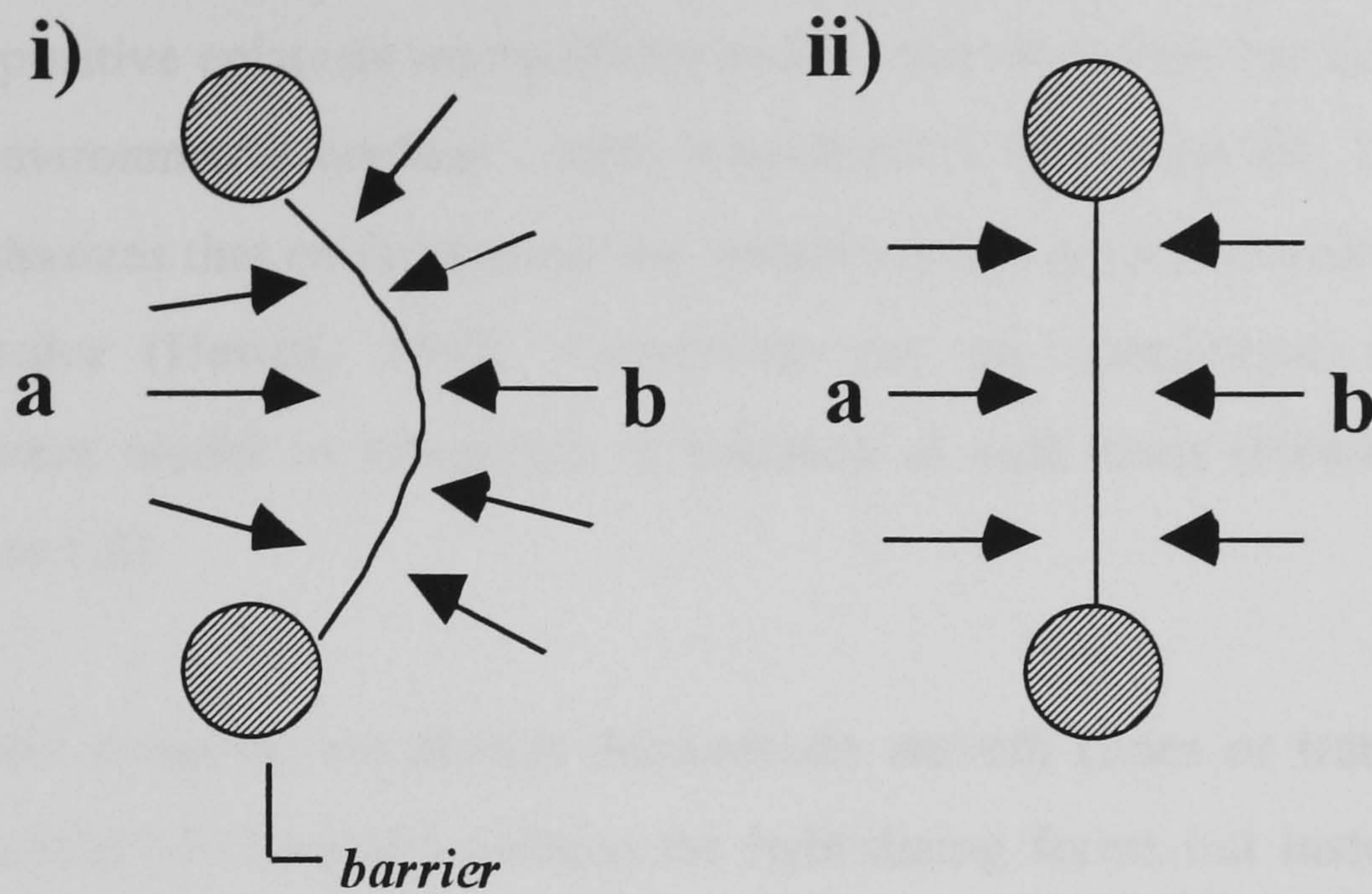


Figure 1.1. Tension zones move to minimize their length. Two populations, **a** and **b**, meet and form a tension zone between physical barriers. The tension zone initially bulges (i) but is smoothed by the greater flux of genes pushing in from the convex side of the bulge (ii) (Barton and Hewitt, 1985).

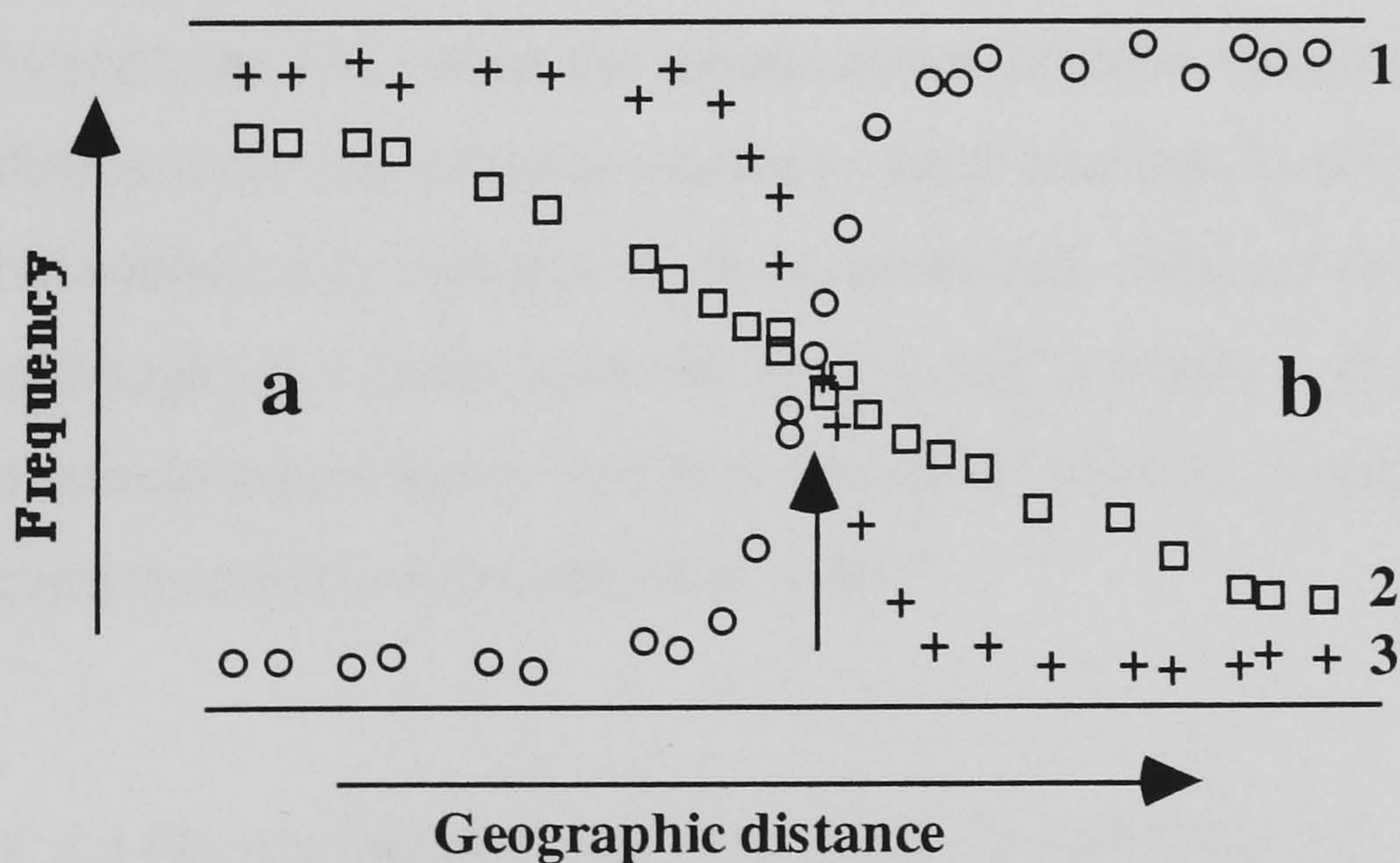


Figure 1.2. Coincident and concordant clines. Two populations, **a** and **b**, meet and form a tension zone. Clines are recorded in three characters, **1**, **2** and **3** (for example the frequency of a diagnostic enzyme allele, a mitochondrial DNA haplotype, and a morphological variable). The clines for all three characters meet at the same location (marked with an arrow) and are therefore *coincident*. However, only characters **1** and **3** are *concordant*; the cline for character **2** is not so steep and the hybrid zone with respect to this character is therefore wider than for characters **1** and **3**.

into density troughs, genes causing hybrid disadvantage tending to move together as a result of linkage disequilibria and migration into the 'hybrid sink', and positive epistasis among alleles within each race where the cline falls across an environmental gradient - with recombinants being less fit. This last point emphasizes that environmental and tension zones are not necessarily mutually exclusive (Hewitt, 1989). Coincident yet non-concordant clines suggest different modes or intensities of selection at each locus (Hewitt, 1989) (see Figure 1.2).

Hybrid zones do not always demonstrate smooth clines or transitions in the frequency of characters defining the hybridizing forms but instead consist of mosaics of genotype frequencies (Arnold, 1997; Harrison, 1986; 1990; Harrison and Rand, 1989). Mosaic hybrid zones are assumed to arise from the adaptation of the two parents to different environments, which are patchily distributed in the area of contact. An example is the crickets *Gryllus firmus* and *G. pennyslvanicus* which are adapted to different, patchily distributed soil types (Arnold, 1997; Harrison, 1986; 1990; Harrison and Rand, 1989). If the dispersal ability of the organism relative to the patch size is small, then the hybrid zone will reflect the mosaic nature of the environment, and if dispersal distances are much smaller relative to patch size then local interactions at patch boundaries may conform to more traditional clinal models of hybrid zones (although on a larger scale the zone is still a mosaic). However if dispersal distances exceed patch size then the hybrid zone may reduce to a traditional clinal model (Harrison and Rand, 1989).

1.2.4 Reinforcement and Character Displacement

Two primary questions concerning the species under consideration in this thesis are 1) to what extent are prezygotic isolation mechanisms present between *Tegenaria saeva* and *T. gigantea?* And 2) if present, what is their nature (behavioural, mechanical etc.)? Whilst attempting to answer these empirical and species-specific questions a third and more general question will

be addressed: is there evidence of reinforcement of prezygotic isolation between *T. saeva* and *T. gigantea*?

Reinforcement is the *process* whereby two incipient species, that have evolved some postzygotic isolation in allopatry, subsequently evolve increased reproductive isolation as a result of natural selection acting against maladaptive hybrids when they later become sympatric (or parapatric) (secondary contact). The process therefore acts to complete the speciation process (Butlin, 1989; Coyne and Orr, 1998; Hostert, 1997; Noor, 1995; Servedio and Kirkpatrick, 1997). Reinforcement is a controversial aspect of speciation theory that has gone in and out of fashion, but is currently undergoing a resurrection (Coyne and Orr, 1998). Reinforcement has been extensively discussed and reviewed recently in Butlin (1987; 1989), Coyne and Orr (1998), Harrison, (1990), Howard (1993), among others. Some important considerations need be addressed here. Butlin (1987; 1989) made the important distinction between reinforcement and "reproductive character displacement". In this context, character displacement is a process that occurs when speciation has already been completed and any F_1 hybrids are either completely inviable or completely sterile. Greater divergences in prezygotic isolation mechanisms (for example female mate discrimination) found in areas of species range-overlap are therefore simply the product of direct selection against individuals that waste their gametes in fruitless interspecific matings. Conversely, reinforcement is a process of speciation (the erection of barriers to gene flow). Reinforcement and character displacement do not form extremes of a continuum, for even if F_1 fitness is very low some gene flow will occur: some F_1 hybrids will on average survive to reproduce and generate F_2 or backcross progeny. Continued production of hybrids will lead to a variety of genotypes within the hybrid zone (Butlin, 1989). This distinction would seem to be important and is adopted here; however other authors (for example, Howard, 1993) disagree, viewing reinforcement as a process that *results* in character displacement. Reinforcement requires that the genes determining prezygotic isolation (positive assortative mating) and the genes determining hybrid fitness must

become nonrandomly associated (linkage disequilibrium); recombination during meiosis acts to break associations between alleles from the parental types and reduces linkage disequilibrium (the selection-recombination antagonism). This constitutes the major theoretical difficulty with the reinforcement mechanism (Hostert, 1997; Butlin, 1989). No such linkage is required for character displacement (following Butlin's (1987; 1989) definition).

The selection-recombination antagonism may be overcome or reduced under some circumstances. Recombination does not interfere with reinforcement if positive assortative mating results from fixation of the same allele in both populations, rather than the fixation of alternative alleles (Felsenstein, 1981a; Harrison, 1990). Also, if hybrid unfitness results from one or a few genes with major effects (which may often be the case (Coyne and Orr, 1998; Howard, 1993)), then 'pure' parental types (with respect to the postzygotic barrier) may continue to segregate at high frequency in the population. In other words, the post-mating barrier will not be destroyed by recombination and more time will be available for the completion of reproductive isolation (Harrison, 1990; Howard, 1993). Genomic events that can reduce recombination, such as chromosomal inversions, could also increase the likelihood of reinforcement (Trickett and Butlin, 1994).

Support for reinforcement has emerged from some recent models and empirical evidence. In an analysis of 171 pairs of *Drosophila* species, Coyne and Orr (1997; 1998) found that recently diverged pairs showed far more prezygotic isolation when sympatric than when allopatric, whereas levels of postzygotic isolation were similar in allopatry and sympatry. Howard (1993) has reviewed many possible examples of reinforcement after reanalysing the earlier literature. Noor (1995) has provided intriguing evidence of increased sexual isolation of female *Drosophila pseudoobscura* from their sibling species, *D. persimilis*, by selection against maladaptive hybridization in areas of sympatry. Saetre *et al.* (1997) have demonstrated reduced hybridization in areas of sympatry in two species of European flycatcher. In this latter study, increased prezygotic

isolation is caused by a divergence in male plumage in the area of overlap. This is presumably caused by sexual selection (Coyne and Orr, 1998; Saetre *et al.*, 1997) and is important because recent models, that allow sexual selection, have shown that reinforcement can occur even when there is only moderate postzygotic isolation (Coyne and Orr, 1998; Kelly and Noor, 1996; Liou and Price, 1994). The studies of Noor (1995), Saetre *et al.* (1997), and a recent laboratory study on *Drosophila melanogaster* by Hostert (1997) which supported the possibility of reinforcement when gene flow is low, are also important because they counter a common criticism of many previous studies of reinforcement. Most studies have either dealt with species not known to hybridize in the wild or to produce some fertile offspring, or have involved laboratory studies in which hybrids have been removed at each generation, thus reducing these studies to examples of Butlin's (1987; 1989) character displacement (Butlin, 1987; 1989; Hostert, 1997; Noor, 1995). In the examples mentioned above, gene flow was known to occur.

Finally, it should be mentioned that the geographical population structure of the hybridizing taxa can have a marked effect on the likelihood of reinforcement. For example, it has been suggested that mosaic hybrid zones are logical places to look for reinforcement because: 1) they are often broad with the centre of the zone far from large 'pure' populations of the parental taxa (and hence the swamping effect of immigrating alleles from outside the zone will be minimized); 2) the species often occupy different habitats or utilise different resources and therefore have populations that are likely to be regulated independently (minimizing the likelihood of the global extinction of one taxon); 3) the patchy nature of the hybrid zone allows multiple contacts between the taxa and many independent opportunities for reinforcement to occur (Harrison, 1990; Harrison and Rand, 1989; Howard, 1993). Other geographical effects on patterns of gene flow may also affect the likelihood of reinforcement. For instance a recent model by Servedio and Kirkpatrick (1997) suggests that reinforcement may occur under a wide range of parameters when migration is symmetrical, but becomes less likely as migration becomes more asymmetrical

(for example when a small 'island' population experiences immigration from the main population, but migration in the opposite direction is rare) (see Chapter 7).

Butlin (1989) posits that for a study to provide convincing evidence of reinforcement it must demonstrate:

"1) that gene flow occurs between the taxa, or did occur when they originally met,
2) that components of the mate recognition system have diverged in the area of contact and in the time since contact was established,
3) that this divergence is sufficient to alter the pattern of mating in a way that decreases the frequency of production of unfit hybrid genotypes, and, ideally,
4) that divergence is not a result of other selection pressures on the mate recognition system".

1.3 The System

1.3.1 The Tegenaria atrica *Group*

The Agelenidae, a large and diverse family of spiders, is most easily characterized by their tendency to weave webs consisting of a tubular retreat opening out onto an often extensive and untidy sheet over which the spider runs down its prey. Agelenids are commonly known as 'funnel-weavers' (Roberts, 1995). Maurer (1992) lists approximately 228 species constituting around 18 genera for the family Agelenidae within Europe (including European Russia and Turkey). These figures are approximate because as Roberts (1995) points out 'the taxonomy of this family is in a fluid state; several genera have been moved to and from different families in recent years'. Approximately 40% (around 97 species) of the species listed by Maurer (1992) belong to the genus *Tegenaria* and this ratio appears to hold when smaller regions are considered:

examination of Roberts (1995) reveals that 39% (11 out of a mere 28 species) of the agelenid fauna belong to *Tegenaria* in northern Europe (Britain, Ireland, northern France, Belgium, the Netherlands, Germany, Denmark, Norway, Sweden and Finland); and further, within just Britain and Ireland, Roberts (1993) lists 44% (7 out of 16 species) as belonging to this genus. The British agelenid genera and species are summarized in Table 1.1.

Table 1.1. British representatives of the Agelenidae (after Roberts, 1993).

Genus	No.	Species
Agelena Walckenaer, 1805	1	*A. labyrinthica* (Clerck, 1757)
Tetrix Sundevall, 1833	1	*T. denticulata* (Olivier, 1789)
Coelotes Blackwall, 1841	2	*C. atropos* (Walckenaer, 1825) *C. terrestris* (Wider, 1834)
Cicurina Menge, 1869	1	*C. cicur* (Fabricius, 1793)
Tetrilus Simon, 1886	2	*T. macrophthalmus* (Kulczynski, 1896) *T. arietinus* (Thorell, 1871)
Cryphoeca T. Thorell, 1869	1	*C. silvicola* (C. L. Koch, 1834)
Tuberta Simon, 1884	1	*T. maerens* (O. P.-Cambridge, 1863)
Tegenaria Latreille, 1804	7	*T. gigantea* Chamberlain and Ivie, 1935 *T. saeva* Blackwall, 1844 *T. atrica* C.L. Koch, 1843 *T. parietina* (Fourcroy, 1785) *T. agrestis* (Walckenaer, 1802) *T. domestica* (Clerck, 1757) *T. silvestris* L. Koch, 1872

The genera are given together with the describing author. The number of British species in each genus follows. The final column lists the species and the describing authors.

The taxonomical and nomenclatural turmoil of the Agelenidae at genus and family level is mirrored at species level within the genus *Tegenaria* and is exemplified by the *Tegenaria atrica* species group. The *T. atrica* group contains three species: *T. atrica* C. L. Koch, 1843; *T. saeva* Blackwall, 1844; and *T. gigantea* Chamberlain and Ivie, 1935. These three species form the basis of this thesis, in particular *T. saeva* and *T. gigantea*. The nomenclatural changes in the group have been summarized by Oxford and Chesney (1994), as follows:

T. atrica prior to 1974, was known as *T. larva* Simon, 1875 (Locket and Millidge, 1953; Locket *et al.*, 1974).

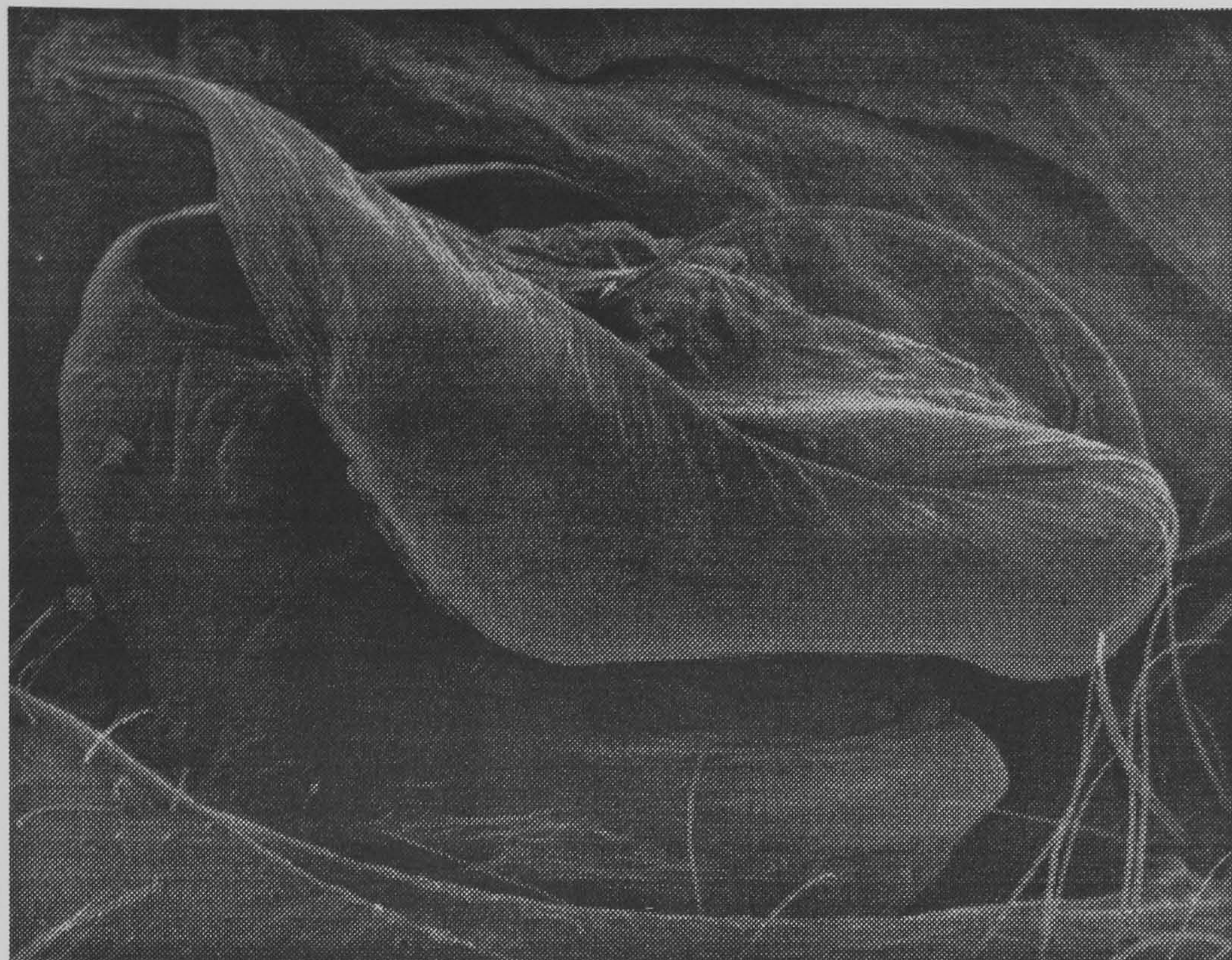
T. saeva , prior to 1974, was known as *T. atrica* (Locket *et al.*, 1974).

T. gigantea, prior to 1975, was undifferentiated from *T. saeva*. It was then described as *T. propinqua* Locket, 1975 (Locket, 1975). In 1976 *T. propinqua* was synonymized with *T. gigantea* Chamberlain and Ivie, 1935 (Crawford and Locket, 1976), and the name *T. propinqua* lapsed. Brignoli (1978) claimed that *T. gigantea* was synonymous with *T. duellica* Simon, 1875. Although adopted by some authors (for example Roberts, 1993; 1995) this identity has not been unequivocally demonstrated (Merrett *et al.*, 1985; Oxford and Chesney, 1994).

Maurer (1992) lists a fourth member of the *T. atrica* group, *T. aliquoi* Brignoli, 1971 but this species is apparently restricted to the island of Sicily and no specimens have been examined for this thesis. If *T. aliquoi* really is a member of the *T. atrica* group then its isolated and restricted distribution is interesting (see Chapter 2). The confusion surrounding the group makes interpretation of the older literature and records difficult if not impossible.

The members of the *T. atrica* group are commonly known as large house spiders and are familiar to most Europeans because of their tendency to associate with human dwellings. There are a number of other *Tegenaria* species also commonly known as house spiders. For example, the cardinal spider *Tegenaria parietina* (Fourcroy, 1785) is a very large species that is widely but rather locally distributed, tends to be associated with old buildings, and occurs outdoors only in southern Europe (Roberts, 1995). The small house spider, *T. domestica* Clerck, 1757, seems to be almost entirely synanthropic and has an almost global distribution (Bristowe, 1958). The members of the *T. atrica* group are only distinguishable by subtle differences in the adult genitalia (the female epigyne and the male palpal structures). These differences are treated in more detail in Chapter 3, but by way of introduction the scanning electron micrographs in Figure 1.3 illustrate the combined conductor + tegulum structure of the male palp in *T. saeva* and *T. gigantea*.

A)



B)

500 μm

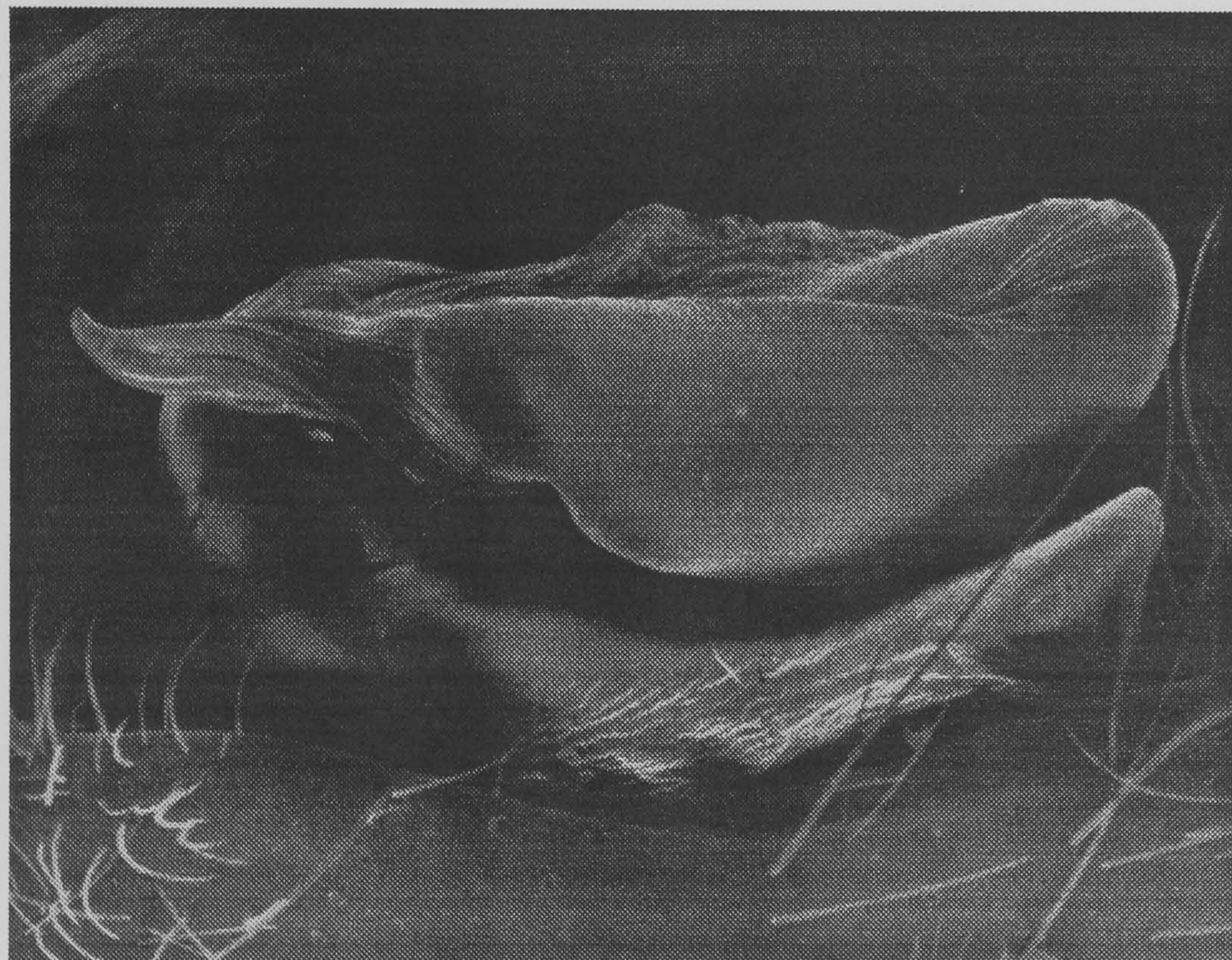


Figure 1.3. Scanning electron micrographs showing the combined tegulum + conductor structure of the palps. A) *T. saeva.* B) *T. gigantea.* Palps viewed laterally from the outside (ectal). Putative hybrids show an intermediate structure.

There have been no detailed ecological studies on members of the *T. atrica* group. However, there do not appear to be any obvious ecological differences between the species; they all spin identical webs and are found in similar habitats (at least in Britain with respect to *T. saeva* and *T. gigantea*) - typically within and around houses and outbuildings, in disused rabbit burrows, old mortared walls, old yew trees, and very commonly on structures thickly covered with ivy (*Hedera helix*) (characteristics which make old English churchyards particularly convenient and productive hunting grounds for the arachnologist). Indeed the species are found anywhere that provides sufficient humidity and is sufficiently complex in structure to facilitate the building of a sheltered tubular retreat. Further, Oxford and Smith (1987) in their survey of large house spiders in Yorkshire noted no significant differences between *T. saeva* and *T. gigantea* in the frequency of captures within houses (i.e. heated) or elsewhere. The apparently identical preferences of these species, together with their similar sizes, suggests that they should be in direct competition for suitable web-building sites when found syntopically. The phenologies of these species are also virtually identical. Adult males of all three species are encountered, from late summer through to late autumn, while in search of females. In Yorkshire, captures of both *T. saeva* and *T. gigantea* peaked at the end of September and beginning of October (Oxford and Smith, 1987). A similar peak was reported for *T. atrica* in northern Spain by Barrientos and Ribera (1988). Females of all the species may live for several years with mating occurring in the autumn, generally followed by a period of cohabitation until the male dies. The female over-winters and produces a series of egg-sacs in the spring (Jones, 1983; and personal observation).

1.3.2 Interspecific Hybridization in Spiders

Definitive reports of individual wild-caught spiders that appear intermediate in morphology to two co-occurring species are few in the literature. Cases include: the *Pardosa pullata* group (Lycosidae) in Europe (Den Hollander *et al.*, 1973: Locket and Millidge, 1951; Vlijm and Dijkstra, 1966); *Lycosa ammophila* and

L. ericeticola (Lycosidae) in Florida (Reiskind and Cushing, 1996); and *Meioneta mossica* and *M. saxatilis* (Linyphiidae) in Europe (Schikora, 1995). The latter example involved a detailed microscopical study and revealed intermediates in one of two locations of syntopic occurrence out of an otherwise non-overlapping distribution. The study by Reiskind and Cushing (1996) describes a hybrid zone. In addition, numerous examples exist of males attempting to court and copulate with females of other species - a fact that has been sometimes exploited to generate hybrids in the laboratory by anaesthetizing the females. These examples are discussed (and references cited) in terms of behaviour in Chapter 6.

The possibility of interspecific hybridization in the *T. atrica* group was first hinted at by Locket (1975) in his paper recognising that what was then known as *T. saeva* actually constituted the two species *T. saeva* and *T. gigantea* (then known as *T. propinqua*). In this paper he describes a specimen of '*T. saeva*' 'from the Senckenberg Museum (No. 8942/1) labelled "Monte Tibidabo b. Barcelona. 500m." that was 'in some ways intermediate between *T. saeva* and *T. atrica*'. Merrett (1980), in an examination of the *T. atrica* group, identified a number of males that were intermediate in palpal characters between *T. saeva* and *T. gigantea* from sites containing both putative parent species. He suggested that these might be hybrids. Similar conclusions were drawn by Oxford and Smith (1987) in a survey of large house spiders in Yorkshire. They report that 43 out of 643 (6.7%) males examined were found to be intermediate between *T. saeva* and *T. gigantea*, both with respect to palp morphology and relative palp dimensions. These specimens also originated from sites containing both putative parents. No females were apparently intermediate. Further evidence for hybridization was provided by Oxford and Plowman (1991) who applied linear discriminant function analysis to a large number of morphological variables measured on a sub-sample of the specimens collected by Oxford and Smith (1987). This analysis suggested that a number of males really were intermediate in morphology, but again no intermediate females were found (although the sample size was small). Experimental evidence that *T. saeva* and *T. gigantea* can hybridize in artificial laboratory crosses was provided by

Kennett and Dalingwater (1986). However, their report was only preliminary and the full extent of their success at producing hybrids is hard to assess. Unfortunately the work was never followed up with further publications and the viability and fertility of any hybrids resulting remains unknown.

1.4 Aims of the Thesis

The aims of the work presented here were as follows. Firstly, to map the distributions of *T. saeva* and *T. gigantea* in southern England, where it was believed that a relatively sharp and long-standing boundary between these species existed (see Chapter 2), and to re-examine the distributions of these species in the York area, where the species are believed to be fairly recent arrivals (again see Chapter 2). On the basis of this distributional information two basic questions were asked: 1) to what extent does hybridization and introgression (gene flow) occur at a relatively long-standing and sharp species boundary (southern England) compared with the situation in Yorkshire? In other words is there evidence of reinforcement? 2) What is the nature of hybridization and introgression between these two species; is it symmetrical or asymmetrical?

To begin to answer these two apparently basic questions, and therefore to begin to understand the evolutionary interactions between these two species, a number of very different approaches were taken: morphometrical, genetical (allozymes and mitochondrial DNA sequence analysis), and behavioural. Each of these approaches incorporated its own subset of specific questions which are introduced in the relevant chapters (though always bearing on the central question of 'reinforcement'). Additionally, mitochondrial DNA data were used to clarify the phylogenetic relationships within the *T. atrica* group. This knowledge, together with wide-scale distributional information is vital if one is to understand species interactions within the group both currently and historically.

2 Distributions

2.1 Introduction

2.1.1 Background

There are no published distribution maps for the *T. atrica* group at European level and, given the nomenclatural changes within the group, caution must be exercised in interpreting the older records and checklists which could be used to assemble such a map. Nonetheless, assembling a crude European distribution map is a simple but neglected task.

The distributions of the *T. atrica* group in Britain and Ireland have been recorded in more detail. The most recent distribution maps were published in Oxford and Chesney (1994) and have been reproduced in Figure 2.1. These maps were based on data from Merrett (1980; 1982; 1989) and used the pre-1974 county boundaries of Britain. A single specimen was therefore sufficient to render an entire county positive. For example, the apparent occurrence of *T. gigantea* in Cornwall relates to one such individual (G. S. Oxford, pers. comm.). These maps, however crude, clearly illustrate the occurrence in southern England of *T. saeva* in the west, and *T. gigantea* in the east. This east/west pattern becomes less clear further north. The scarcity of records for *T. atrica* in Britain is also evident from Figure 2.1. Oxford and Chesney (1994) assert that many of the British specimens were probably imports from Eire. *Tegenaria atrica* is well established in Eire and this has probably been true since at least the early 1800s. However, the species has only been recorded once in Northern Ireland and its status in the Province is unclear (Oxford and Chesney, 1994). Indeed, large house spiders have only been confirmed to occur in Northern Ireland, around Belfast, in the last twenty years. There appear to be established populations of both *T. saeva* and *T. gigantea* in this area. *T. saeva* has been recorded in Eire from Co. Dublin in 1985 and 1986, but there are no confirmed records for *T. gigantea* from Eire (Oxford and Chesney, 1994). Hence, the

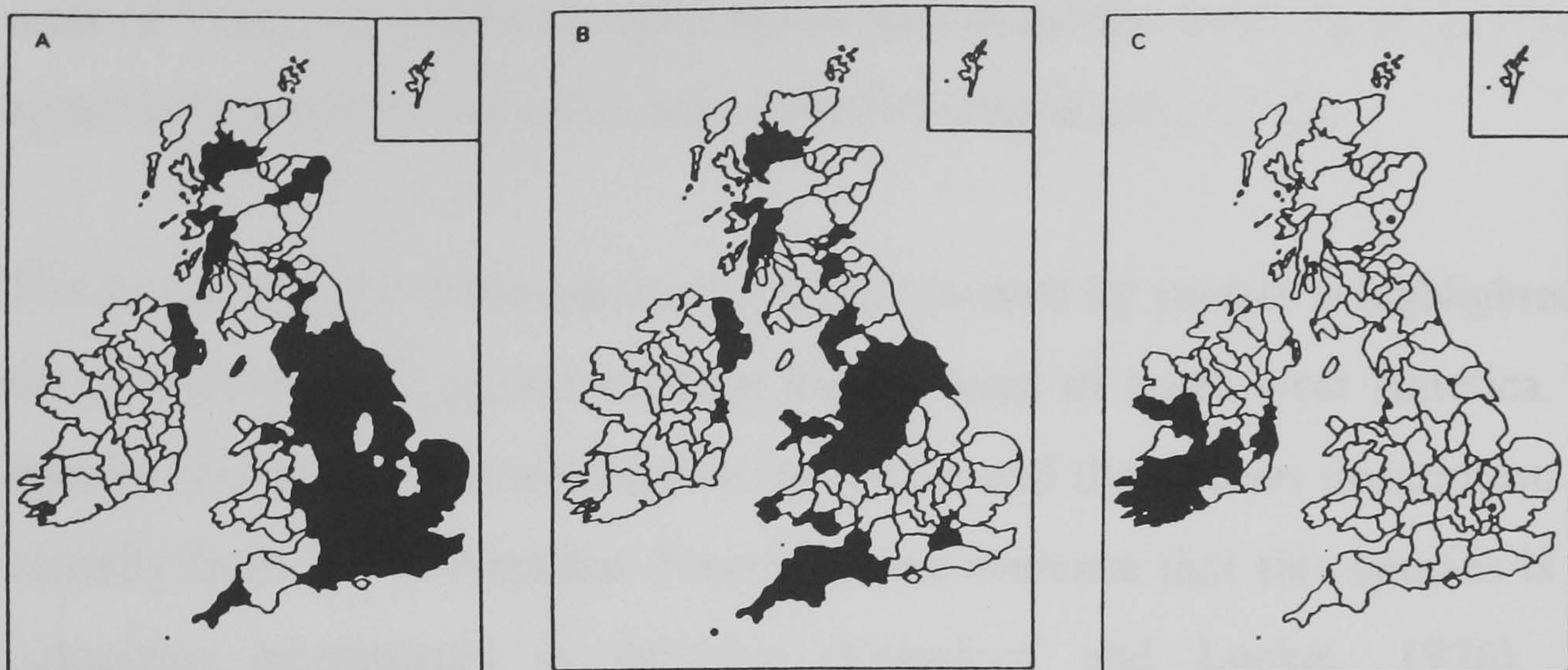


Figure 2.1. Distributions of A. *T. gigantea*; B. *T. saeva*; C. *T. atrica* in Britain and Ireland. Based upon data from Merrett (1980; 1982; 1989) and unpublished records. Records for Great Britain are denoted using pre-1974 county boundaries. In C, dots represent individual spiders. The shaded dot in A and B represents an old record that may represent either *T. saeva* or *T. gigantea.* (Taken from Oxford and Chesney, 1994).

curious situation exists whereby *T. saeva* and *T. gigantea* are frequently found in Britain, whereas *T. atrica* is rare and probably not established, yet in Ireland *T. atrica* is established in the south with only recently established populations of *T. saeva* and *T. gigantea*, mainly in the north. There are strong indications from county records that both *T. saeva* and *T. gigantea* have expanded their ranges in Britain in the recent past (Oxford and Chesney, 1994; Oxford and Plowman, 1991; Oxford and Smith, 1987; Parker, 1984; Smith, 1985). According to Smith (1985), large house spiders were first recorded from Yorkshire in the mid-1960s. Range expansion is also supported by anecdotal evidence; many older Yorkshire people, during the surveys described below, wrote that they did not recall encountering these large and memorable intruders when younger. The reasons for this range expansion are obscure. The distributions of *T. saeva* and *T. gigantea* in the York area were examined by Oxford and Smith (1987). They noted that *T. saeva* occurred in very high frequencies, relative to *T. gigantea*, within the city of York, its major suburbs

and adjacent villages (up to 88%). However, the frequency of *T. gigantea* was higher in the surrounding rural areas. The frequency of *T. saeva* was moderate (ca. 25%) in the conurbations of south Yorkshire and those to the immediate west of York, but picked up again as one moved further west. A proportion of apparently intermediate males was identified (see section 1.3.2).

The propensity for these species to be translocated by people is highlighted by the occurrence of *T. gigantea* on the Pacific coast of north-west America. The species was probably introduced in the 1920s and the species description was actually from North America. There is some evidence that this species is also colonizing northwards in America (Crawford and Locket, 1976). The introductions of *T. atrica* into Britain, the establishment of *T. saeva* and *T. gigantea* in Northern Ireland (described above), the introduction of all three species into Iceland (see below), the discovery of a member of the *T. atrica* group (apparently *T. atrica*) that was unloaded from a European ship in Perth, Australia (B. Y. Main, pers. comm.), and a specimen of *T. gigantea* from the Nilgiri Hills, India examined by Merrett (1980), all illustrate the ease of human-mediated transport for the *T. atrica* group.

2.1.2 Aims

This section simply aims to clarify the geographic distribution of the *T. atrica* group in Europe and the distributions of *T. saeva* and *T. gigantea* in southern England. The distribution of *T. saeva* and *T. gigantea* in the York area is also re-evaluated. A solid knowledge of distribution is a pre-requisite for a full understanding of the evolutionary history of these species.

2.2 Materials and Methods

2.2.1 European Distributional Information

Detailed distributional information for the *T. atrica* group (and for most spiders) at a continental level is lacking (although attempts to collate a clearer picture are being attempted (G. S. Oxford, pers. comm.)). Maurer's (1992) comprehensive checklist and review of the European Agelenidae included a country by country list of occurrence for the European species. The information for the *T. atrica* group was simply extracted from this review and used to a generate a crude, but informative European distribution map for these species. Additional information on Iceland, Yugoslavia, and Italy was taken from Agnarsson (1996), Nikoloc and Polenec (1981), and Pesarini (1994) respectively. The major European biogeographic regions were overlaid on this map following the schematic representation of Stanners and Bordeau (1995).

2.2.2 British Distributions

Samples of *T. saeva* and *T. gigantea* from the York area and from southern England were collected through surveys in which the public were encouraged to collect large house spiders, alive, and take them (together with information on the place and date of capture) to a network of collection centres. Extensive use of the media (local newspapers, radio and television) was made in order to advertise these surveys. The previously noted east/west distribution pattern in southern England, along with records from P. Merrett (G. S. Oxford, pers. comm.) suggested that the two species might meet in the county of Dorset. Hence, two surveys were organized in 1994, one in the York area and one in Dorset, and these were timed to coincide with the mating season which commences in September. At this time of year the adult males, searching for females, are a familiar sight in northern European houses and are therefore easily collected. A similar survey was undertaken in the autumn of 1995 in Dorset *and* Hampshire. This was necessary because it became clear that the

contact zone for the two species was further to the east of Dorset than had originally been anticipated.

Coincident with retrieving the specimens from the collection centres for the public surveys, field collections were made in southern England. Specimens were obtained from suitable locations by 'fishing' with blowfly larvae (*Calliphora vomitoria*) which were obtained from angling shops as 'gentles' (maggots). A blowfly larva was placed upon the web and, with patience, the incessant wriggling would usually eventually draw a *Tegenaria* from its inaccessible retreat. The spider was allowed to bite the larva (premature attempts at capture invariably fail - *Tegenaria* are fast) and then captured by grabbing with the hand from behind or where possible (and generally more successfully) by cupping a clear plastic container over the animal, with retreat being made impossible by the surface of the web. This approach has been described in Oxford and Croucher (1997). In addition, a number of members of the British Arachnological Society kindly provided specimens. Adult specimens were preserved in 95% ethanol (some specimens had their abdomens removed first - see Chapter 4). Most specimens collected were adult; any juveniles (unless very small) were maintained until adult following the protocol described in Chapter 6. Specimens, immersed in 95% ethanol, were identified from the appearance of the male palp or female epigyne under a Reichert dissecting microscope.

The distribution data for southern England presented here also include specimens from the collections for the behavioural experiments described in Chapter 6.

The collection localities, grid references, and identifications of the specimens have been retained by the author and by Dr. G. S. Oxford (University of York). Records have also been deposited with the national Spider Recording Scheme of the British Arachnological Society and it is hoped that they will be included in the forthcoming Biological Records Centre (Institute of Terrestrial Ecology, Monks Wood) spider distribution atlas.

2.3 Results

2.3.1 European Distributions

Figure 2.2 shows the country by country distribution of the *T. atrica* group within Europe, based upon information in Agnarsson (1996), Maurer (1992), Nikoloc and Polenec (1981) and Pesarini (1994). Although this is a crude distribution map it clearly illustrates that *T. atrica* holds a far greater range than either *T. saeva* or *T. gigantea*. The range of *T. atrica* appears to extend across most of Europe, with the exception of the south-eastern Mediterranean and southern Italy - all Italien records for *T. atrica* are from northern Italy (Pesarini, 1994). However, it is also clear that *T. atrica* is primarily a species of eastern and central Europe where it is the only representative of the *T. atrica* group. *Tegenaria saeva* and *T. gigantea* occupy a mainly Atlantic distribution, occurring from Portugal through Spain and France, with *T. saeva*, but not *T. gigantea*, being recorded from Belgium and The Netherlands. *Tegenaria atrica* apparently occurs throughout the ranges of *T. saeva* and *T. gigantea*, and there is clearly a need for a more detailed knowledge of the distributions of the *T. atrica* group in these areas. The presence of all three species in Britain and Ireland has been described above (and see section 2.3.2, below). All three species have also been recorded in Iceland although *T. gigantea* has been found only once, indoors. *Tegenaria saeva* and *T. atrica* are not only commonly imported but well established in warehouses in ports (Rekyavík). There is some morphological evidence that these two species are hybridizing in Iceland (Agnarsson, 1996; G. S. Oxford, pers. comm.).

Figure 2.2 also illustrates the restricted distribution of *T. aliquoi*, a putative fourth member of the *T. atrica* group. This species is restricted to the island Sicily and is currently totally isolated from other members of the *T. atrica* group. None of the species is recorded from neighbouring mainland southern Italy and no information exists for the north African coast.

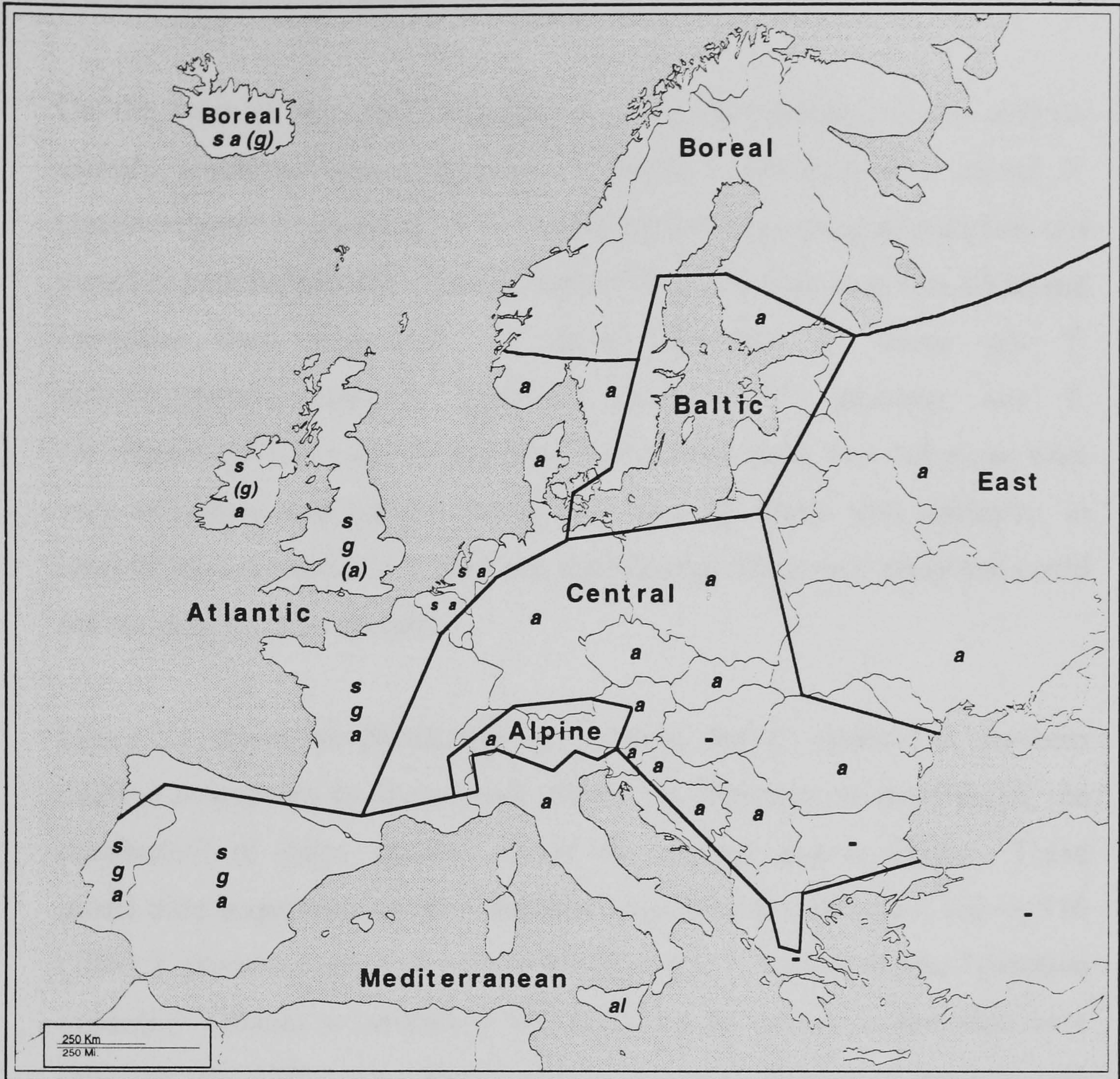


Figure 2.2. European Distribution of the *T. atrica* group. Country by country occurrence of the *T. atrica* group plotted from the checklist of Agnarsson (1996), Maurer (1992), Nikoloc and Polenec (1981) and Pesarini (1994). s = *T. saeva*, g = *T. gigantea*, a = *T. atrica*, al = *T. aliquoi*, - = apparently absent. No information is available for countries without symbols. The location of a symbol does not imply any regional location within that country. Biogeographic zones follow the schematic representation of Stanners and Bordeau (1995). *Tegenaria atrica* has been placed in parentheses in Britain because it occurs very infrequently. *Tegenaria gigantea* has been placed in parentheses in Iceland because has only been found once. The checklists do not differentiate between the four countries of Yugoslavia therefore the records for *T. atrica* in Yugoslavia are ambiguous in location. For clarity, the countries have not been indicated; those with checklists are as follows: Portugal, Spain, France, Ireland, Great Britain, Iceland, Belgium, The Netherlands, Germany, Switzerland, Austria, Italy, Czech Rep., Slovakia, Poland, Hungary, Yugoslavia (Slovenia, Croatia, Bosnia, Serbia), Romania, Bulgaria, Greece, Turkey, Ukraine, European Russia, Norway, Sweden, and Finland.

2.3.2 British Distributions

On the basis of the visual assessment of the morphology of the external genitalia, specimens were categorized as follows: *T. saeva* ('good' *T. saeva*), *T. gigantea* ('good' *T. gigantea*), or *T. saeva/gigantea?* (appearing intermediate and therefore putative hybrids). For the material from the York area two additional categories were recognized: *T. saeva?* (between *T. saeva* and *T. saeva/gigantea?*) and *T. gigantea?* (between *T. gigantea* and *T. saeva/gigantea?*). It must be emphasised that these latter two categories were very subjective and simply reflect the range of forms and difficulty in identification encountered in the York area material. These two categories could not be recognized for females.

Figure 2.3 shows the distribution of *T. saeva* and *T. gigantea* in southern England as recorded by the surveys. Figure 2.4 illustrates, in more detail, the distribution of these species around the contact area in Dorset. These distribution maps represent 837 identified specimens (512 (61%) *T. saeva*, 316 (38%) *T. gigantea*, and 8 (4 male, 4 female) (1%) *T. saeva/gigantea?* putative hybrids (= difficult to indentify)). 529 (63%) of the specimens identified were male and 299 (37%) were female (the proportion of females captured was greatly boosted by the field surveys). Most specimens were easily identifiable.

The separate ranges of these two species in southern England were very clear from the distribution maps. *T. saeva* occupies the west and *T. gigantea* the east and the two species meet in eastern Dorset. The boundary between the two species runs northwards from the conurbations around Wimborne, Poole and Bournemouth. Although the boundary was quite marked there was considerable heterogeneity in the detailed distributions of the two species (for example, the 'pocket' of *T. saeva* on the coast to the east of Bournemouth in Figure 2.4). This may partly reflect the complex and patchy nature of the conurbations and movement of people in this increasingly developed area of the south-coast.

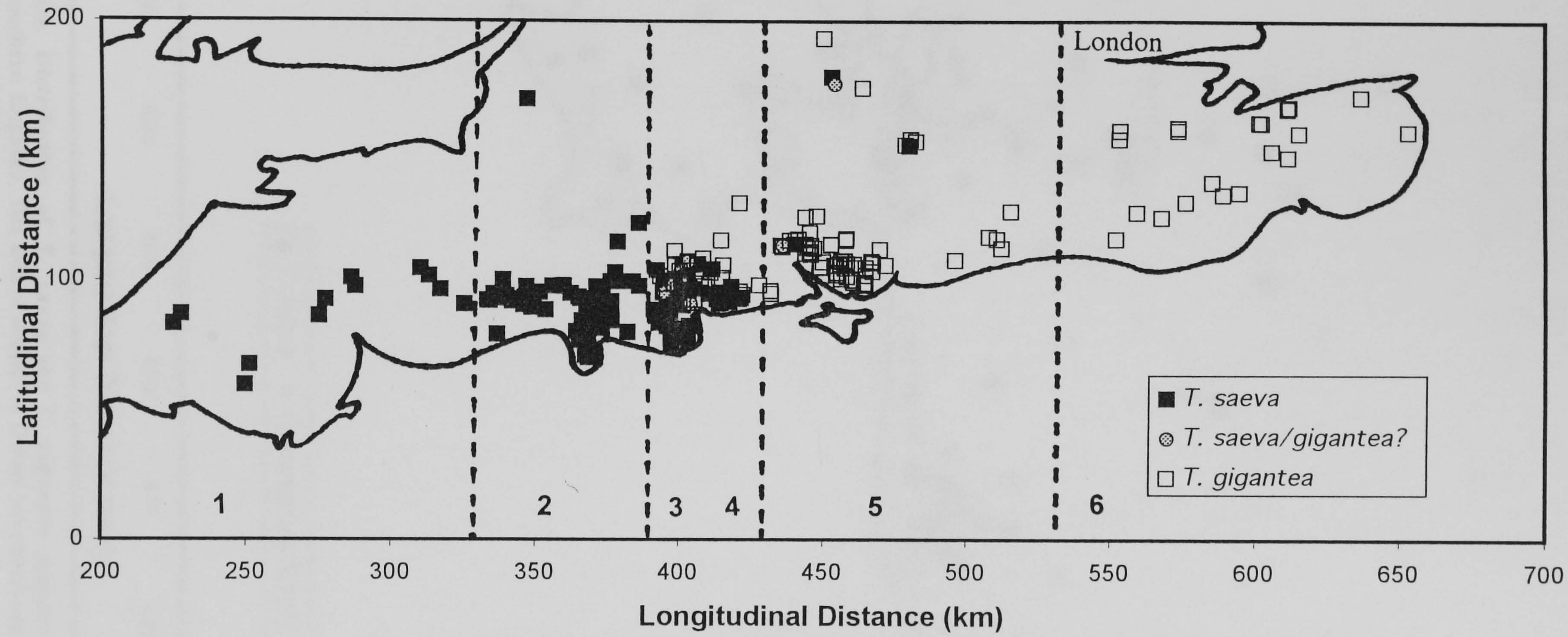


Figure 2.3. Distribution of *T. saeva* and *T. gigantea* samples in southern England. The division of southern England into six sampling zones is illustrated (see section 2.4). Several specimens were collected from many of the sampling sites.

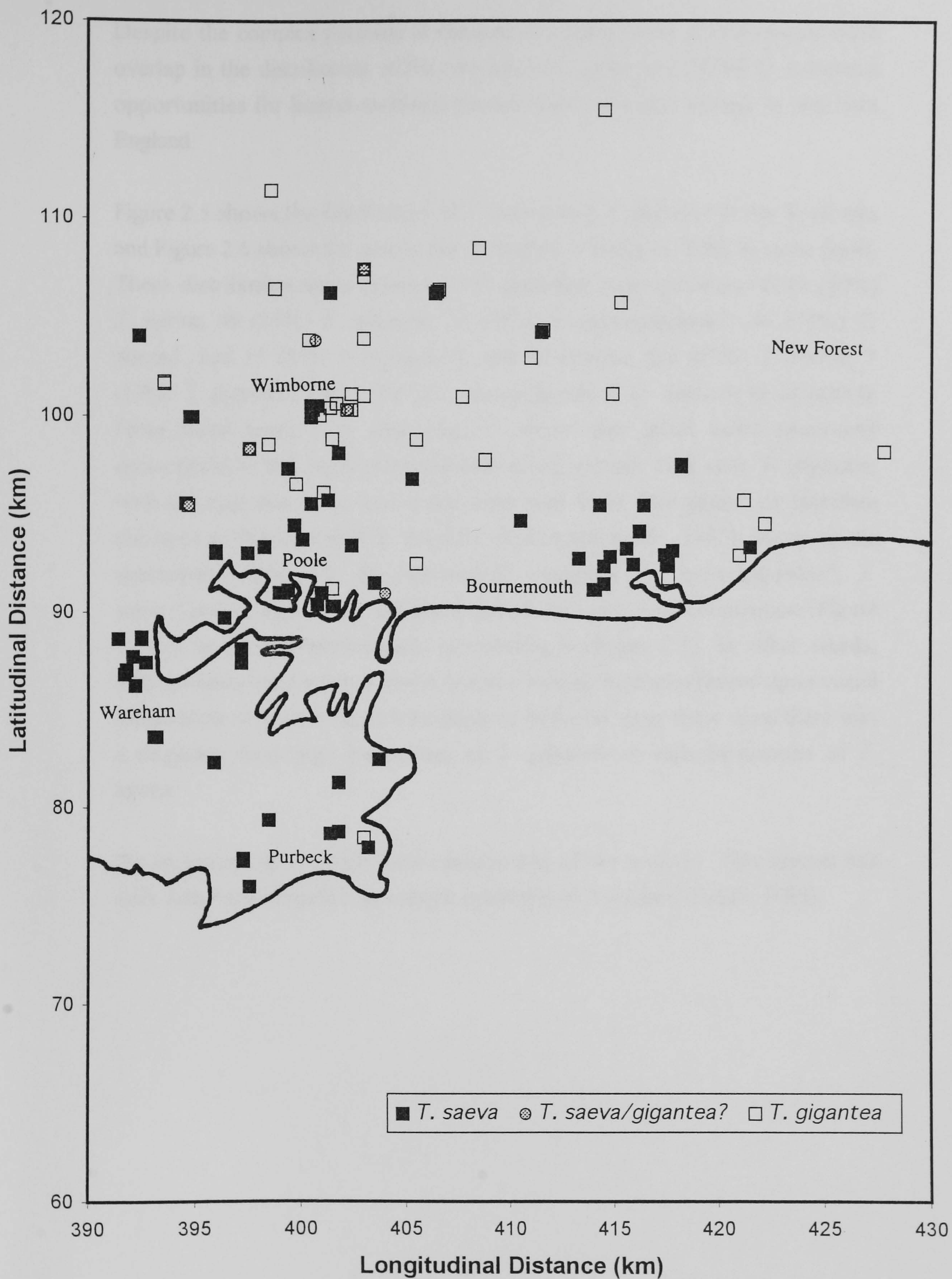


Figure 2.4. Distribution of *T. saeva* and *T. gigantea* samples at the contact zone in southern England. The area of contact in Dorset (corresponding to zones 3 and 4) is shown in more detail.

Despite the complex patterns in the area of contact there is surprisingly little overlap in the distribution of the two species - given the probably numerous opportunities for human-mediated translocation of house spiders in southern England.

Figure 2.5 shows the distribution of *T. saeva* and *T. gigantea* in the York area and Figure 2.6 shows the area in the immediate vicinity of York in more detail. These distribution maps represent 337 identified male specimens (196 (58%) *T. saeva*, 49 (15%) *T. gigantea*, 31 (9%) *T. saeva/gigantea?*, 46 (14%) *T. saeva?*, and 15 (4%) *T. gigantea?*), and 36 females (24 (67%) *T. saeva*, 7 (19%) *T. gigantea*, and 5 (14%) *T. saeva/gigantea?* (= difficult to identify)). From these maps it is clear that *T. saeva* was much more commonly encountered in the major conurbations in and around York than *T. gigantea*, with the opposite being true in the more rural areas. The pattern is therefore the same as that reported in detail by Oxford and Smith (1987). Many of the specimens falling into the intermediate categories (*T. saeva/gigantea?*, *T. saeva?* and *T. gigantea?*) originated within the main York conurbation (Figure 2.6) or in the area immediately surrounding it (Figure 2.5). In other words, intermediates were most common in areas were both species (based upon visual assessment of morphology) were likely to be found, or in areas were there was a transition from high frequencies of *T. gigantea* to high frequencies of *T. saeva*.

No specimens of *T. atrica* were taken in any of the surveys. This species has only once been recorded, as a single specimen, in Yorkshire (Smith, 1985).

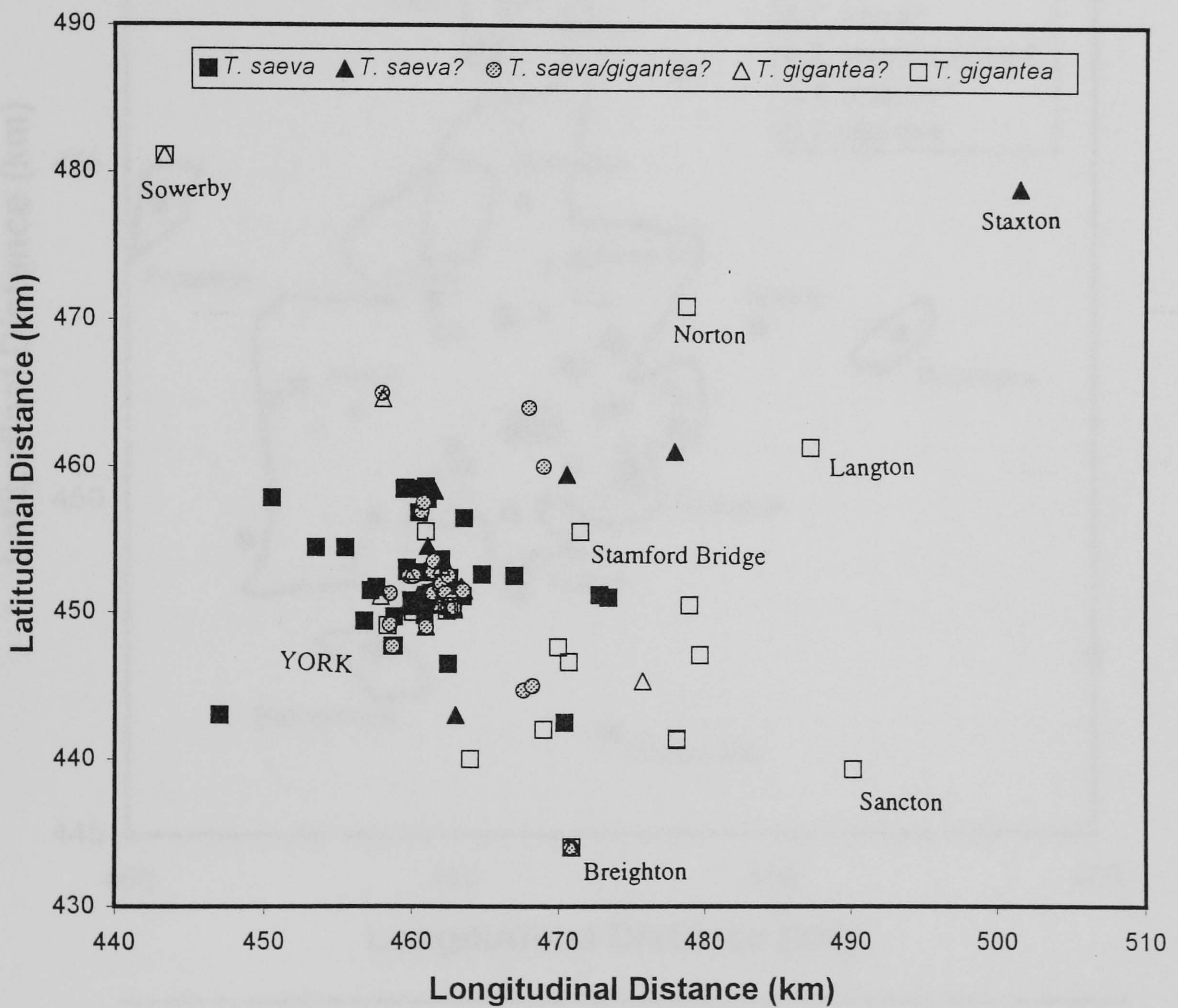


Figure 2.5. Distribution of *T. saeva* and *T. gigantea* samples in the York area. Specimens grouped according to categories of visual identification based on palp/epigyne morphology. Many sample sites represent several specimens.

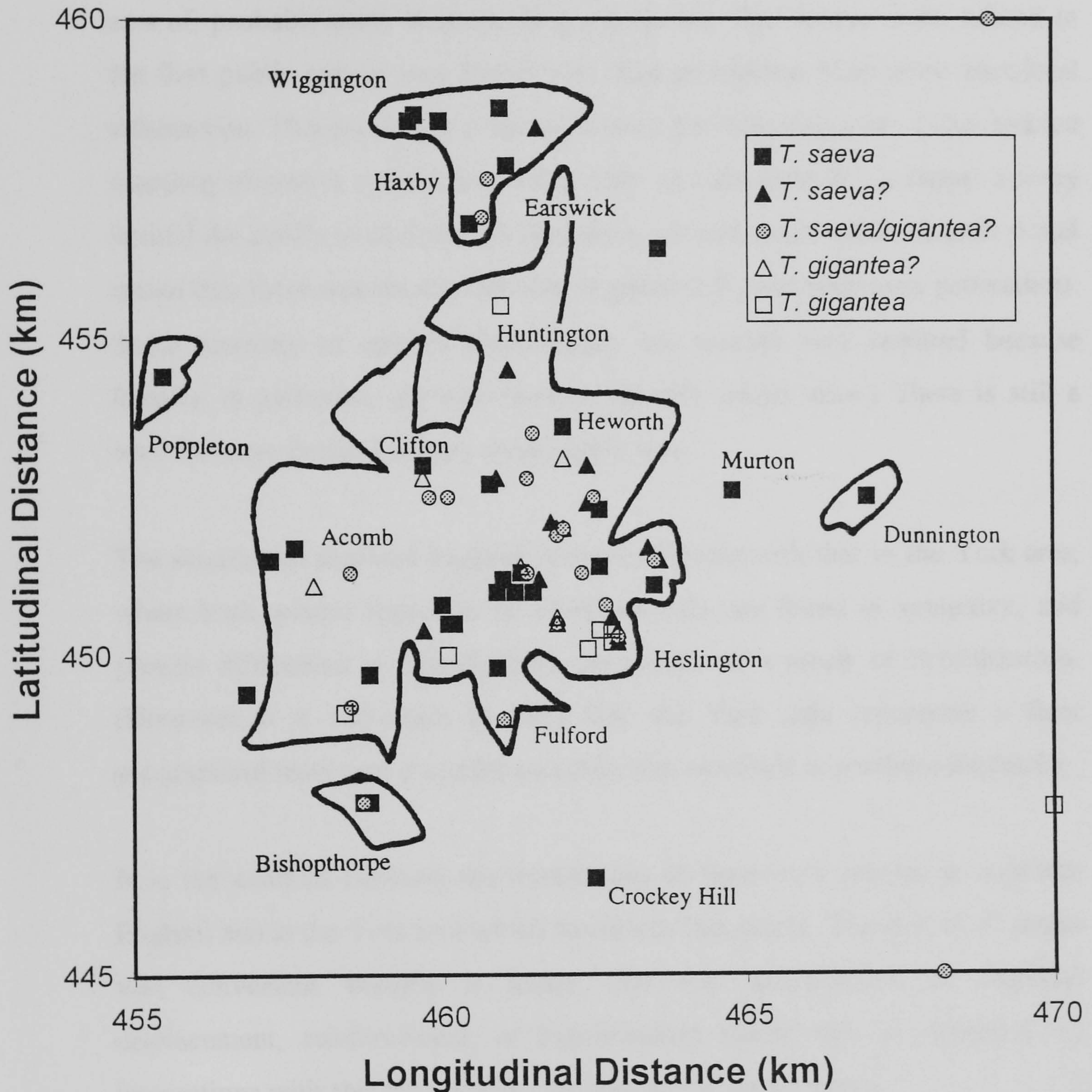


Figure 2.6. Distribution of *T. saeva* and *T. gigantea* samples in the York area (York detail). The sample sites in York City and its immediate surroundings are shown in more detail. Specimens grouped according to categories of visual identification based on palp/epigyne morphology. Many sample sites represent several specimens. In no cases were both *T. saeva* and *T. gigantea* recorded from the same location (house).

2.4 Discussion

The distribution maps for Britain, presented here, reveal an interesting situation. In southern England, *T. saeva* occupies the west and *T. gigantea* the east, and the two species meet, with apparently little hybridization, in eastern Dorset. Hence there are largely allopatric populations of both species and an area of, probably quite long-standing, parapatry. The contact area defined in the first public survey was further east than anticipated from prior anecdotal information. This prompted a second survey the following year. (The detailed mapping of such a contact area takes time and the need for a repeat survey limited the ability to perform the laboratory crosses described in chapter 6 and meant that there was insufficient time to generate F_2 and backcross generations. Some certainty of species identification by locality was required because females, in particular, are very hard to identify whilst alive.) There is still a need for more detailed survey work in this area.

The situation in southern England contrasts sharply with that in the York area, where both species appear to be recent arrivals, are found in sympatry, and present difficulties in identification, apparently as a result of hybridization. (However, it is important to note that the York data represents a finer geographical scale over a smaller area than that examined in southern England.)

It is the contrast between the interactions of these two species in southern England and in the York area which has driven this thesis. The lack of *T. atrica* was convenient because it meant that any assessments of character displacement, reinforcement, or hybridization would not be confused by interactions with this third species (at least in the recent past).

On the basis of the distributions, southern England was divided into six zones. The contact area fell conveniently onto the 40 km x 40 km area covered by Ordnance Survey map sheet 195. This area was therefore used to define zone 3 and zone 4, with zone 3 corresponding to specimens of *T. saeva*, and zone 4

corresponding to specimens of *T. gigantea*, from the contact area. The land to the west of this block (entirely occupied by *T. saeva*) was then divided into zone 1 and zone 2, such that zone 1 was 'the most allopatric'. The land to the east was similarly divided into zones 5 and 6 (entirely occupied by *T. gigantea*). The allocation of zones 1, 2, 5, and 6 was somewhat arbitrary, but an attempt was made to have each zone including a similar density of samples across a similar, but not identical, longitudinal distance. Zone 1 essentially corresponds to Devon and eastern Cornwall, zone 2 is most of Dorset, zone 5 is most of Hampshire and West Sussex, and zone 6 is East Sussex and Kent. No such geographical segregation was possible for the material from the York area; instead this material was simply divided according to the five identification categories outlined in section 2.3.2: *T. saeva*, *T. saeva?*, *T. saeva/gigantea?*, *T. gigantea?* and *T. gigantea.* These divisions will be important in the chapters that follow.

The European distribution map reveals the extensive geographic range of *T. atrica* compared to that of *T. saeva* and *T. gigantea*, which are mainly confined to the far west of Europe. The origins of the British distributions are not clear, however it seems almost certain that *T. saeva* and *T. gigantea* would not have occupied their current locations, or even been present in Britain, more than about 9000-10,000 years ago when the climate warmed after the last glaciation. If one, or both, of these species colonized Britain without human help then they must have done so before Britain was isolated from France around 8000 years ago (Hewitt, 1990). They certainly did not colonize Ireland in this time. Members of the *T. atrica* group are not known to use wind dispersal (ballooning), indeed no member of the Agelenidae is known to balloon (S. Toft, pers. comm.). They would therefore be unlikely to have blown across the English Channel. Further, most species of spider that *do* balloon rarely disperse from the egg-sac by more than a few metres (Wise, 1993). The European distributions suggest that *T. saeva* and *T. gigantea* may have possibly expanded from Iberian refugia whilst *T. atrica* was expanding from somewhere in the east. The land-bridge to Britain may have been closed before *T. atrica*

arrived in northern France. Without further evidence, which might be gained from mitochondrial DNA surveys, only speculation is possible. The establishment of *T. atrica* in Eire must surely result from the movement of people. Whether or not the populations of *T. saeva* and/or *T. gigantea* in Britain result from natural or human-mediated colonization cannot be known, however, *T. atrica* has certainly failed to establish itself in Britain. This may result from direct competitive exclusion with the already established *T. saeva* and *T. gigantea* populations or from a tendency to hybridize with the already established species. Similar processes could explain the surprisingly discrete distributions of *T. saeva* and *T. gigantea* in southern England. Recent human impact should have blurred the parapatric boundary, but this doesn't seem to have occurred, suggesting that groups of *T. saeva* have difficulties in establishing themselves when translocated to *T. gigantea* territory, and similarly for groups of *T. gigantea* on translocation to *T. saeva* country. It would be interesting to know if the distributions of the three species fall into discrete patches in the Iberian Peninsula and in France, or whether they are found in sympatry.

T. aliquoi warrants further investigation, both at a morphological and molecular level. If this species really belongs to the *T. atrica* group then its currently discrete distribution could shed light on the past distributions of the group; did ancestral forms occupy refugia in southern Italy during previous glaciations?

3 Morphometrical Analyses of *T. saeva* and *T. gigantea.*

3.1 Introduction

3.1.1 The Mechanical Isolation Hypothesis

As stated in Chapter 1, the members of the *T. atrica* group are visibly distinguished only by differences in the structures of the male palp and the female epigyne. An examination of any arachnological guide will reveal that, as for many other invertebrates, genitalic differences are the usual, and often the only, criteria used to distinguish closely related species. The traditional explanation for these differences in the sclerotized external genitalia of spiders has been that they function as a prezygotic mechanical isolation system by allowing only conspecifics - with the correct specifically-shaped genitalia - to copulate effectively. This mechanical isolation or "lock-and-key" hypothesis was originally proposed by Dufour in 1844 (Mayr, 1970). Clearly, if the differences between the genitalia of two closely related species have evolved to limit costly interspecific matings, and that these two species are occasionally likely to engage in interspecific matings, then two inescapable predictions follow: in areas of long-standing sympatry or parapatry compared to areas of allopatry for each species 1) the genitalic differences between the species should be greater, and 2) the genitalia should exhibit less variation (Ware and Opell, 1989). In other words, if the "lock-and-key" hypothesis is correct then character displacement or reinforcement of mechanical isolation should be observed (assuming that interspecific mating does result in either infertile or relatively inviable hybrids).

Although there can be no doubt that structural differences in the genitalia do act as mechanical barriers to interspecific mating between many closely related species (Mayr, 1970; and see Chapter 6), this does not explain how these differences have evolved. The notion that they have evolved to facilitate mechanical isolation has been under attack since 1905 when Jordon (Mayr, 1970) noted that 48 out of 698 species of Sphingidae (hawkmoths) examined did not show any differences in genitalia and that many species also showed considerable geographic variation in their genitalia. In more recent times Eberhard (Eberhard, 1985; Ware and Opell, 1989) has raised a number of objections to the "lock-and-key" hypothesis. Such objections include the fact that species-specific sclerites in the males of some species contact parts of the female genitalia that do not differ between species, and that the male genitalia are frequently far more elaborate than would be required for mechanical isolation. As Ware and Opell (1989) rightly point out, courtship and copulation are costly for a female. Therefore, identification of a potential suitor's specific status early on in courtship, before any attempts at copulation, should be under strong selection. Suggested alternative explanations for species-specific genitalia (Eberhard, 1985; Ware and Opell, 1989) include: 1) complex pleiotropic by-products of alleles selected in other contexts (also suggested by Mayr, 1970), 2) mechanical conflicts of interest between males and females, 3) isolation by genitalic stimuli, and 4) sexual selection by (cryptic) female choice (Eberhard's main tenet).

3.1.2 Hybridization, Introgression, and Morphometrics

The possibility that the members of the *T. atrica* group might hybridize in the wild was first recognised on the basis of individual specimens that appeared intermediate in genital morphology (see section 1.3.2). Indeed, to date, the only evidence for naturally occurring hybrids of any spider species is morphological. Further, the arachnological literature reveals only four studies

that have examined variation in spider genitalia within a population: Coyle (1985); Huff and Coyle (1992); Reiskind and Cushing (1996); and Schikora (1995). Only the studies by Huff and Coyle (1992) and Schikora (1995) have examined genitalic variation across a species range. Schikora (1995) examined the morphological variation in the genitalia throughout the known distribution of the closely related species *Meioneta mossica* and *M. saxatilis* (Linyphiidae). In one out of two cases of syntopic occurrence, intermediate males and females were noted. Reiskind and Cushing (1996) reported a hybrid zone between *Lycosa ammophila* and *L. ericeticola* (Lycosidae) in which morphometrical measurements on males revealed clear evidence of hybridization, and possibly introgression between the species.

Several morphological examinations of the *T. atrica* group have been made previously. Merrett (1980) examined the variation in the group and showed that plots of the combined tegulum + conductor length against the carapace length could distinguish the three species. *Tegenaria atrica* was shown to have the shortest tegulum + conductor length relative to carapace length, with *T. gigantea* the next longest, and *T. saeva* the longest. As previously mentioned, Merrett (1980) also noted a number of males that appeared intermediate between *T. saeva* and *T. gigantea.* Oxford and Smith (1987) applied Merrett's (1980) approach to material from the York area and compared their results to those of Merrett (1980). They found that the separation of the two species was much poorer than Merrett (1980) had observed. Oxford and Smith's (1987) data indicated much more variation than Merrett's (1980), and interestingly they also showed that *T. saeva* from York was more variable than *T. gigantea.* They found that 6.7% of the males examined appeared intermediate in morphology between *T. gigantea* and *T. saeva*, though no females were identified as being intermediate. This observation was taken further by Oxford and Plowman (1991) who performed a linear discriminant function analysis on material from the York area. In this analysis all females were assigned unambiguously to either *T. saeva* or *T. gigantea*. Nine out of eighteen males that were thought to

be intermediate from visual inspection were reclassified as *T. gigantea*, whereas the remaining nine fell between *T. saeva* and *T. gigantea*, suggesting strongly that they were indeed of hybrid origin. Barrientos and Ribera (1988) examined the variation in the *T. atrica* group in the Iberian Peninsula from a taxonomical perspective. They took a descriptive approach and concluded that the characters used to distinguish the three species were not reliable and were critical of the taxonomy; tending to regard the *T. atrica* group more as a single 'morphospecies' yet not wanting to synonymize them.

3.1.3 Fluctuating Asymmetry

Bilaterally symmetrical organisms may exhibit deviations from bilateral symmetry in three ways, each of which is characterized by a different combination of mean and variance of the distribution of right-minus-left (R-L) differences (Palmer and Strobeck, 1986). *Directionally asymmetrical* traits exhibit normally distributed R-L differences about a mean that is significantly greater or less than zero and represent a consistent bias within a species towards greater development on one side of the body than on the other, for example the asymmetry of flatfish (Palmer and Strobeck, 1986). *Antisymmetrical traits* exhibit a broad-peaked or bimodal distribution of R-L differences about a mean of zero. For example, the oversized signalling claw of fiddler crabs is always larger than the other claw but may occur on either side with equal frequency (Palmer and Strobeck, 1986). *Fluctuating asymmetry* (FA) refers to small random deviations in R-L differences which therefore show a normal distribution with a mean of zero (Palmer and Strobeck, 1986). FA is frequently taken as an indicator of developmental stability, which is the stable development of phenotype under given environmental conditions (Møller, 1997). Developmental instability has been shown to be caused by a range of environmental factors (for example, climatic perturbations, malnutrition, pollution and parasitism), and by a range of genetic factors (for

example, inbreeding, hybridization and mutation) (Møller, 1997). However, the relationship between FA and environmental and genetic factors is *very* controversial. For example, Fowler and Whitlock (1994) found that FA was not increased by moderate inbreeding in *Drosophila melanogaster*, and Windig (1998) suggests that variation in FA in peacock butterflies (*Inachis io*) has a purely random cause and is unlikely to be influenced by sex, environment or the quality of genes.

Part of the appeal of FA as a measure of developmental stability derives from the a priori knowledge that the ideal situation is perfect symmetry. Further, the degree to which an individual might show deviations from symmetry in *either* direction often has a heritable basis (Palmer and Strobeck, 1986). Studies of FA generally examine morphological traits but there is no reason why such deviations might not exist in physiology or immunology (Møller, 1997). Developmentally unstable individuals may experience reduced fitness directly as a result of the factors causing the instability. They may also experience reduced fitness from increased FA because, for example, mate preferences for symmetrical individuals may be very common (Møller, 1997). Symmetrical individuals generally have faster growth, higher fecundity, and better survival than asymmetrical individuals which implies that developmentally unstable individuals have a lower metabolic efficiency, expending more resources on controlling their development than less unstable individuals (Møller, 1997). Reviews on the statistics and relationships of FA to fitness and developmental stability can be found in Møller (1997) and Palmer and Strobeck (1986; 1992) among others. (The statistics commonly employed in FA studies have been strongly criticised by Sullivan *et al.* (1993)). Two observations are of particular interest here. First, repeated observations have shown significant negative associations between protein heterozygosity and FA: more heterozygous individuals exhibit a smaller degree of FA. It seems that increased heterozygosity provides some 'buffering' against environmentally induced perturbations during development (Palmer and Strobeck, 1986). Secondly,

interspecific hybrids often show increased FA relative to their parents, even though they are more heterozygous. The 'buffering' or canalizing role that heterozygosity plays within species appears to be outweighed by the disruption of 'coadapted gene complexes' from the parental species (Palmer and Strobeck, 1986).

3.1.4 Aims

The aims of this Chapter are to examine to the patterns of morphological variation in samples of *T. saeva* and *T. gigantea* from southern England and from the York area and to ask: 1) to what extent do hybridization and introgression appear to have been occurring in these areas? 2) Is there any evidence in support of the mechanical isolation hypothesis (reinforcement or character displacement)? 3) Can any differences in the degree of fluctuating asymmetry be detected between samples with little exposure to hybridization (allopatric) compared to those that may be experiencing the effects of more hybridization?

3.2 Materials and Methods

3.2.1 Samples and Measurements

The specimens used in these analyses were collected either in the 1994 and 1995 distributional surveys as outlined in Chapter 2, or in the 1996 field collections for the behavioural experiments described in Chapter 6. Sample sizes are given for each analysis, as appropriate, in the Results (section 3.3). Specimens were divided into nine 'samples' or data sets based upon geographical location and initial visual identification. Specimens from the York area were simply grouped according to the visual identification classes outlined in Chapter 2: York *T. saeva*; York *T. saeva?*; York *T. saeva/gigantea?* (putative hybrids); York *T. gigantea?*; and York *T. gigantea*. No females were classified as York *T. saeva?* or York *T. gigantea?*. Specimens from southern England were taken from zones 1, 3, 4, and 6 (as defined in Chapter 2): zone 1 = *T. saeva* from Devon and Cornwall; zone 3 = parapatric *T. saeva* from the contact zone in Dorset (corresponding to sites within the 40 km x 40 km area covered by Ordnance Survey map sheet 195); zone 4 = parapatric *T. gigantea* (as zone 3); zone 6 = *T. gigantea* from Kent and East Sussex. The samples of *T. saeva* from Devon and Cornwall, and of *T. gigantea* from Kent and East Sussex, were regarded as consisting of individuals from relatively pure (allopatric) populations. This assertion was supported by the relative ease with which all individuals from these regions could be allocated to species through visual inspection. The 'purity' of the samples from other regions was not so certain, therefore these allopatric samples were treated as 'types' against which all others could be compared or classified. In addition to the above sample groups, a small number of laboratory bred F_1 hybrids (see Chapter 6) were included as appropriate, as were a very small number of indeterminate females from southern England (all parapatric). Males and females were treated separately.

The characters measured on each individual are shown, together with their abbreviations, in Figure 3.1 (males) and Figure 3.2 (females). Twenty characters were measured on each male. Eighteen of these measurements were taken from the palps and two were measures of the prosoma. All palpal measurements were taken on both the left and the right palps (unless one palp was missing), and the average of each paired measure taken. Similarly, the measure of prosoma length (PROL) consisted of a left and right measurement, which were averaged. Therefore, in total, up to 39 measurements were taken on each male.

Fifteen variables were measured on each female. Eleven of these measurements were taken from the epigyne, and four were measures of the prosoma (prosoma length, prosoma width, and two pattern characters on the sternum). (The pattern characters were not measured for the males because they were frequently poorly defined in the specimens examined.) Six of the epigyne measurements existed in both left and right forms; the average measure was calculated. Including these paired measures, and the paired measure PROL, 22 measurements in total were taken on each female.

The characters measured were chosen from an initially larger set incorporating those recommended by Oxford and Plowman (1991), and many additional variables. Many potentially useful characters were quickly eliminated when they proved difficult or ambiguous to measure (and would therefore not be replicable), or when initial analyses suggested that they were highly correlated with body size (PROL). An emphasis was placed on retaining characters associated with the palps and epigynes as these would be most likely 1) to distinguish the species, and 2) to be involved in any mechanical isolation and thus be the characters most likely to experience character displacement or reinforcement.

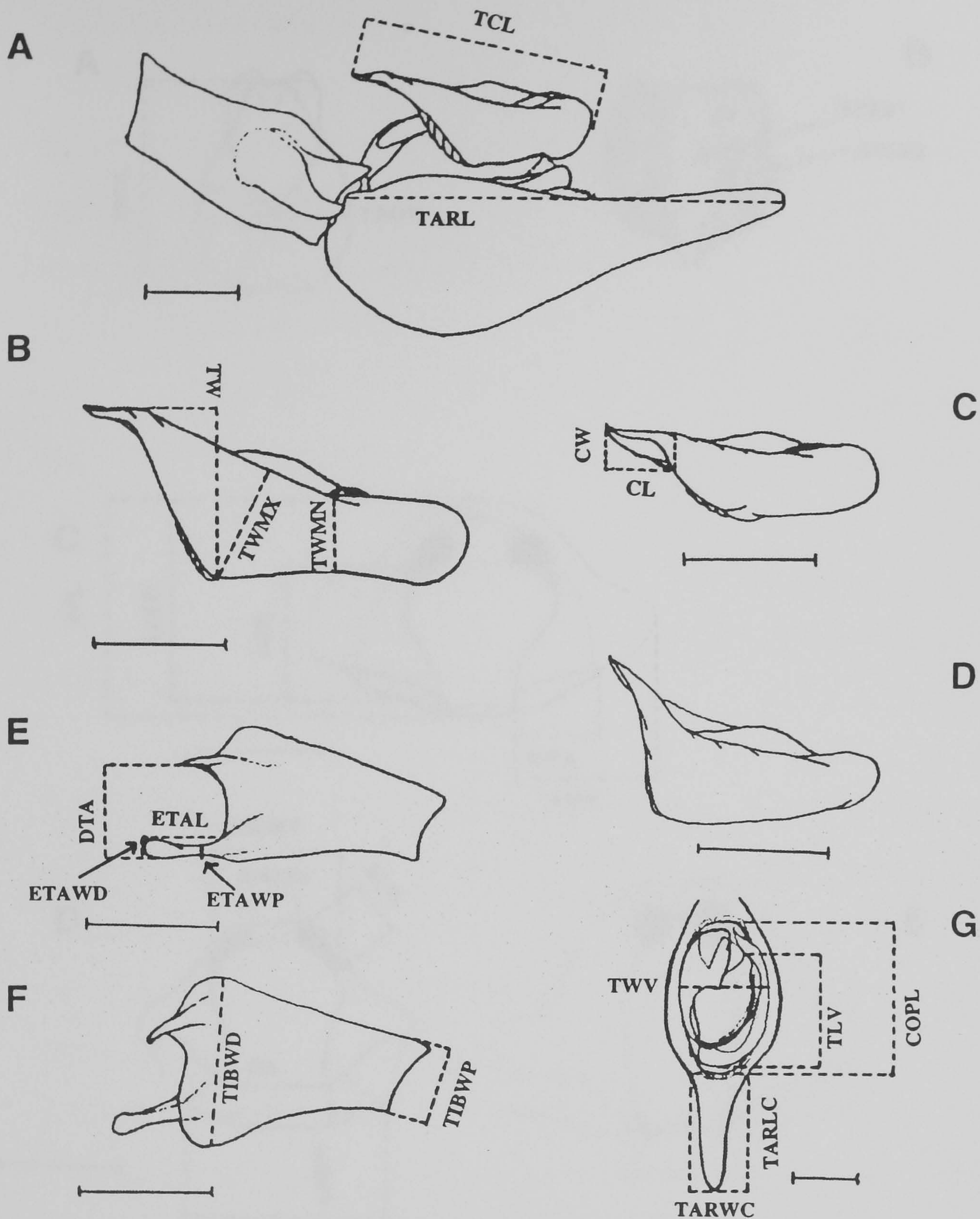


Figure 3.1. Male measurements used in the morphometrical analyses.
Scale bar = 0.5 mm. **TCL** = maximum tegulum + conductor length; **TW** = tegulum width; **TWMX** = maximum tegulum width; **TWMN** = minimum tegulum width; **CL** = conductor length; **CW** = maximum conductor width; **TARL** = tarsus length; **ETAL** = ectal tibial apophysis length; **ETAWP** = ectal tibial apophysis *minimum* width (proximal); **ETAWD** = ectal tibial apophysis *maximum* width (distal); **DTA** = distance between tips of tibial apophyses; **TIBWD** = maximum distal tibia width; **TIBWP** = maximum proximal tibia width; **TLV** = tegulum length (ventral); **TWV** = tegulum width (ventral); **TARLC** = tarsus length from (distal) end of cymbium; **TARWC** = tarsus width at (distal) end of cymbium; **COPL** = maximum cymbial operculum length. Diagrams are of the male left palp. **A**, **B**, **C**, **D** are ectal views (from the outside) with ventral side uppermost, **E** is a ventroectal view, **F** is also ventroectal but slightly more ventral, **G** is a ventral view. All diagrams are of *T. saeva* except **C** and **D** which illustrate the tegulum and conductor shape in *T. gigantea* and *T. atrica*. Measurements on the prosoma are shown in Figure 3.2.

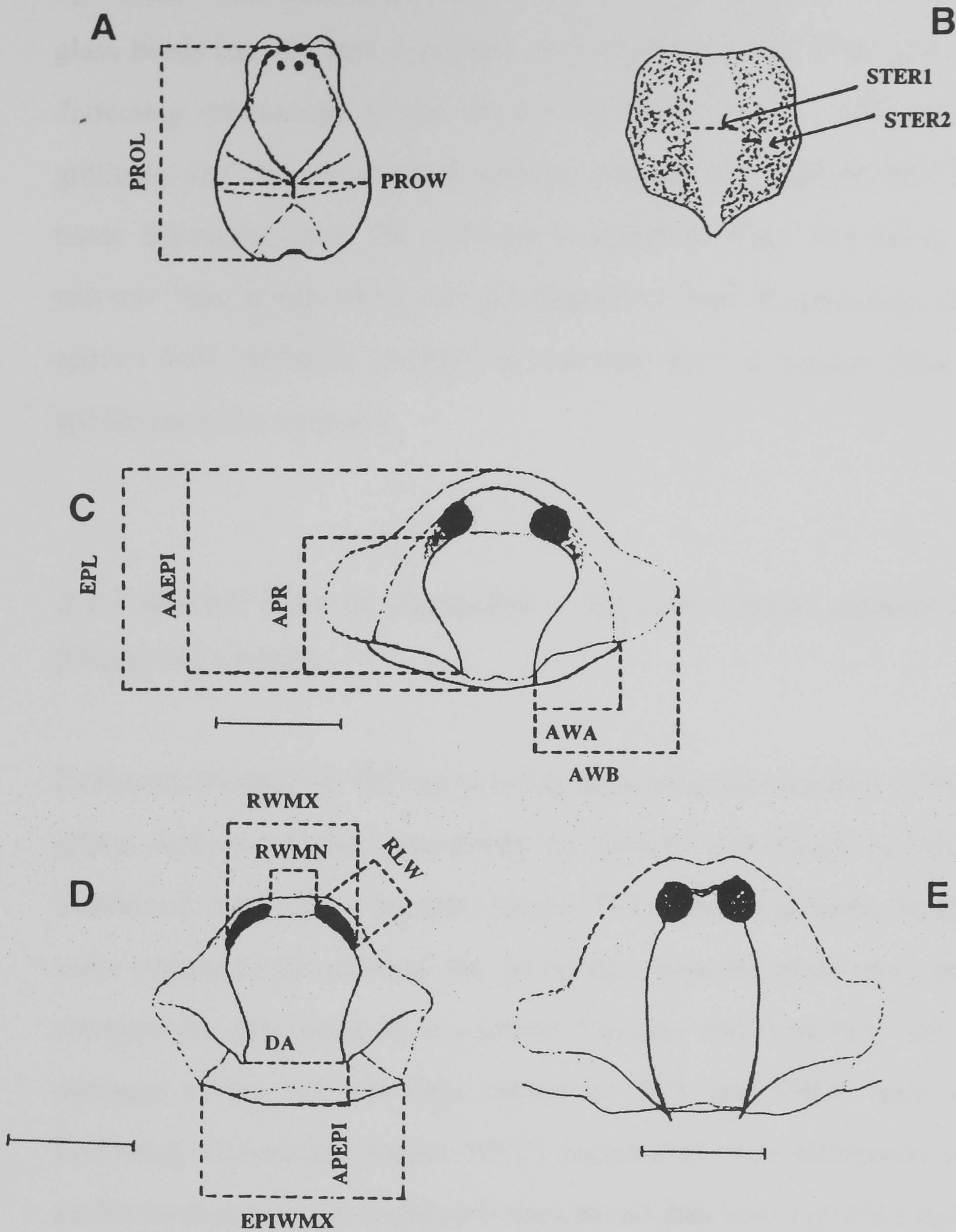


Figure 3.2. Female measurements used in the morphometrical analyses.
Scale bar = 0.5 mm (not shown for **A** and **B**). **PROL** = prosoma (carapace) length (left side measurement shown); **PROW** = prosoma width. Both **PROL** and **PROW** were also measured in males. **STER1** = sternum pattern width 1; **STER2** = sternum pattern width 2; **EPL** = maximum epigyne length; **AAEPI** = tip of apophysis to anterior limit of epigyne; **RWMX** = maximum width between receptacles; **RWMN** = minimum width between receptacles; **DA** = distance between tips of apophyses; **RLW** = receptacle long width; **AWA** = apophysis width 'A'; **AWB** = apophysis width 'B'; **APR** = tip of apophysis to posterior edge of receptacle; **APEPI** = tip of apophysis to posterior limit of epigyne; **EPIWMX** = maximum width of epigyne at posterior limit. Diagram **A** is a dorsal view of the prosoma, **B** is a ventral view of the sternum, **C**, **D**, and **E** are ventral views of the epigynes of *T. saeva*, *T. gigantea* and *T. atrica* (for illustration).

Specimens were held in position in either a watch-glass or a Petri-dish using glass beads (ca. 0.4 mm diameter). Measurements were taken with a Reichert dissecting microscope to the nearest half-division of a calibrated eye-piece graticule, and converted to millimetres. Magnification ranged from 12.4 to 60 times depending upon the character in question. Care was taken to prevent recorder bias in measuring the specimens (i.e. runs of specimens of the same species from the same 'sample') by selecting them at random from a pool of specimens to be measured.

3.2.2 Merrett Plots of Conductor + Tegulum Length against Carapace (Prosoma) Length

Following Merrett's (1980) approach to separating the members of the *T. atrica* group, and as applied statistically by Oxford and Smith (1987), plots of 'combined' conductor + tegulum length (TCL) against prosoma length (PROL) were compared for each of the groupings. Separate plots were prepared to compare the specimens from southern England and from the York area. The averages of the left and right values for TCL and PROL were employed. Following Oxford and Smith (1987), regressions were performed on the data set for each group. The regression lines for all data sets were then examined for equality of slope and elevation in an analysis of covariance (ANCOVA). For regression lines with non-significantly different slopes, ANCOVA tests the elevation (in effect the intercept) of the lines by testing for homogeneity among the group means of the dependent variable Y (TCL), after adjusting the means for differences in the independent variable X, or covariate (PROL) (Sokal and Rohlf, 1995). The significance of the differences among the adjusted TCL means was tested using the GT2-method for unplanned multiple comparisons, employing the Gabriel approximation to generate 95% comparison intervals (Sokal and Rohlf, 1995).

3.2.3 Measurement Error

Following the approach of Lynch and Hayden (1995), an estimate of possible measurement error was made by selecting six specimens of each sex at random (three from the allopatric *T. saeva* specimens and three from the allopatric *T. gigantea* specimens). Each of these specimens was measured eight times, on separate occasions, and a mean obtained for each individual and each character. This gave 48 replicates with six individual means for 20 measurements in males and 15 measurements in females (only right side measurements were taken). The 'within-individual' coefficient of variation for each mean was taken as a measure of measurement error (%ME). Following Lynch and Hayden (1995), Haldane's (1955, in Lynch and Hayden, 1995) correction for small sample sizes was used to calculate this coefficient:

$$\%ME = 100\,(1+1/4n)\sigma/x$$

where n is the sample size (8), σ is the standard deviation and x is the mean of the eight replicates per individual. This estimate for each individual was then averaged to give a mean %ME for each character.

3.2.4 Univariate Analysis

For the allopatric *T. saeva* and the allopatric *T. gigantea* samples, the means for each of the variables were compared, in turn, by separate one-way analyses of variance (ANOVA).

3.2.5 Multivariate Analysis

Those morphological variables measured, that discriminated most efficiently between the allopatric samples of *T. saeva* and *T. gigantea*, were identified through a combination of multiple-group principal components analysis (MGPCA) (Thorpe, 1988) and canonical discriminant analysis (canonical variates analysis - CVA). Principal components analysis (PCA) is an ordination technique commonly used to reduce the number of variables required to describe the variation within a group of samples. PCA aims to find, through transformation of the coordinate system, a set of new orthogonal (= uncorrelated) axes, each of which takes up in decreasing order as much of the remaining variance as possible. Linear combinations of the original variables are transformed to fit these new axes by multiplying each variable by eigenvectors (principal component coefficients) derived from a correlation or covariance matrix of the original data, thus the original number of variables may be reduced to a few components only (Airoldi and Flury, 1988; James and McCulloch, 1990; Thorpe, 1988). PCA is a one-group method and cannot be applied directly to a multiple group situation by simply pooling the data without the within-group and the between-group variation confounding each other (Airoldi and Flury, 1988; James and McCulloch, 1990; Thorpe, 1988). MGPCA attempts to escape this problem by performing an PCA on the pooled within-group covariance matrix of the groups under consideration. This procedure results in new orthogonal axes with zero within-group intercorrelation, and which therefore define independent patterns of variation (Lynch and Haden, 1995; Prenter *et al*, 1995). Further, the MGPCA typically results in a first principal component (PC) on which all measures are positively loaded and which may be interpreted as 'size' variation, and subsequent PCs with both positive and negative loadings which may be interpreted as 'shape' variation (Lynch and Haden, 1995; Prenter *et al*, 1995; Thorpe, 1988). As the PCs are uncorrelated they provide ideal input for canonical variate/discriminant function analyses (Lynch and Haden, 1995; Thorpe, 1988).

All values for each variable were first multiplied by 1000 and then log-transformed to make the variance independent of the mean. Multiplication by 1000 avoided negative logarithms. The values for the allopatric *T. saeva* samples and the allopatric *T. gigantea* samples were used to generate a pooled within-groups covariance matrix using the canonical discriminant feature of SPSS (Norusis/SPSS, 1994). An MGPCA was then performed on this matrix using the SAS (version 6.07) package running on the University of York UNIX system. The resulting PCs were then applied to the log-transformed data. For the allopatric *T. saeva* and allopatric *T. gigantea* samples, the individual PC scores were examined for within- and between-group variation using univariate ANOVA. The PC scores for the allopatric *T. saeva* and allopatric *T. gigantea* sample groups were then used as a 'training set' in a canonical discriminant analysis using SPSS (Norusis/SPSS, 1994). The resulting function was used to ordinate and classify all other samples. The analyses were performed separately on males and females.

3.2.6 Fluctuating Asymmetry Analysis

All paired character measurements (those occurring on both the left and right of the animal) were examined for evidence of fluctuating asymmetry (FA). The extent of FA in each of the sample groups (the morphologically defined groups from the York area and zones 1 to 6 from southern England) was compared. For each individual the asymmetry of each character was calculated by subtracting the size of the left character from that of the right (R-L).

It was important to distinguish asymmetry from measurement error. The ideal approach to this is to take each measurement twice and to compare measurement error to the other non-directional asymmetry in a two-way, mixed model analysis of variance (Palmer and Strobeck, 1986; Rettig *et al.*, 1997). However, repeated measurements for each variable on each individual would have been prohibitively time-consuming and therefore the estimate of

percentage measurement error (%ME) for each variable, as outlined above, was used as a conservative measurement error correction. For each individual, the mean size for each character was calculated ((L+R)/2) and multiplied by %ME/100 to give an estimate of possible error, scaled to size, for that particular individual and character. This correction was subtracted from the absolute asymmetry value (|R-L|) and any resulting negative values reset to zero (negative values were taken to indicate that the asymmetry (R-L) was within the limits of error). The corrected values of |R-L| were then reallocated their former signs.

Having attempted to correct for measurement error, it was then necessary to eliminate variables where the asymmetry might include or represent directional asymmetry or antisymmetry. To check for directional asymmetry the signed differences for each character (R-L) in each grouping, were subjected to one-sample *t*-tests (Rettig *et al*., 1997; Swaddle *et al*., 1994). The *P*-values were adjusted using a sequential Bonferroni (Dunn-Sidák) correction (Rettig *et al.*, 1997; Sokal and Rohlf, 1995). Antisymmetry may be checked for by examining the skewness and kurtosis of R-L values in each sample, with the expectation that in a normally distributed population these will equal zero; however, this is not recommended for sample sizes less than 30 (Rettig, *et al.*, 1997). Instead, departures from normality were assessed using a Kolmogorov-Smirnov test (Palmer and Strobeck, 1996) with the *P*-values again adjusted using a sequential Bonferroni (Dunn-Sidák) correction (Rettig *et al.*, 1997; Sokal and Rohlf, 1995).

Palmer and Strobeck's (1986) FA index 6 was employed as the measure of FA:

$$\text{FA} = \text{var}\ \frac{R-L}{(L+R)/2}$$

This measure is the most powerful measure for detecting differences between groups when the mean and FA scale together (Fowler and Whitlock, 1994; Palmer and Strobeck, 1986). Although inspection of plots of asymmetry (R-L)

against size ((L+R)/2) suggested that there was probably little size-scaling, the above index was preferred over Palmer and Strobeck's (1986) index 5 (suggested as best in such cases) as it is more readily applied in a multiple-group comparison. FA index 5 simply returns a variance and, in a two-group situation, the most powerful test *is* simply an *F*-test of the two variances (Palmer and Strobeck, 1986). However, in a multiple group situation a value for each individual (using FA index 6) in each group is required in order to compare the group variances of these values. These variances - the differences in the extent of FA between the sample groupings - were tested using a Levene's test, which is an homogeneity of variance test (an analysis of variance on the absolute deviations from the group means), as recommended by Palmer and Strobeck (1986).

3.3 Results

3.3.1 Merrett Plots of Conductor + Tegulum Length against Carapace (Prosoma) Length

Plots of tegulum + conductor against prosoma length for the material from southern England and from the York area are shown in Figure 3.3. The separation of *T. saeva* and *T. gigantea* in southern England is very clear, with the allopatric and parapatric data sets within each species appearing similar. The three known F_1 hybrids fall roughly between the two groupings. The resolution of the species from the York area data is much poorer. Although the 'good' specimens of *T. saeva* and *T. gigantea* separate reasonably well, they do not do so to the same extent as the specimens from southern England. Again, the unidentified, putative *T. saeva/gigantea* hybrids, clearly fall between the 'good' species clouds. The *T. saeva?* data set appears to fall largely between the 'good' *T. saeva* and the putative hybrids, and the *T. gigantea?* data set similarly appears to fall largely between the 'good' *T. gigantea* and the putative hybrids. It could be argued that it is not surprising that the visually designated taxonomic categories show a pattern concordant with their original assessment (for example, that the visually intermediate York *T. saeva/gigantea?* samples show intermediate values of tegulum + conductor length against prosoma length). However, it should be stressed that the argument is not (totally) circular: the original visual assessment was based largely on an assessment of shape, whereas in the analyses presented in this chapter the relative sizes of the structures are being analysed. Examination of the regression lines for all the data sets (Figure 3.4) reveals the patterns more clearly. The regression equations, residual mean squares (RMS) and sample sizes for each data set are presented in Table 3.1. The RMS from these regressions were compared using *F*-tests (Table 3.2). These results suggest that the data for the parapatric *T. saeva* samples are more variable than the data for all other samples from

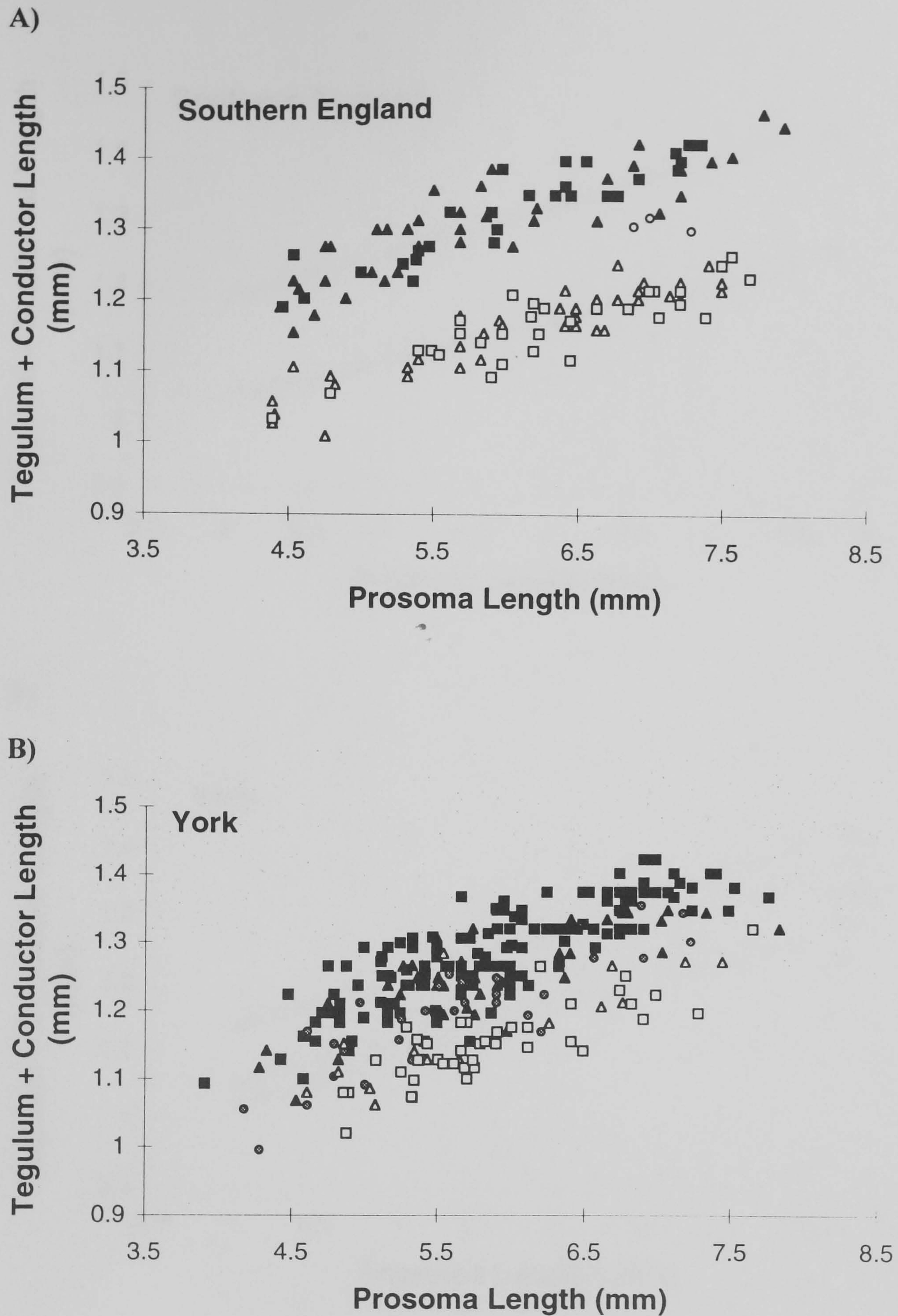
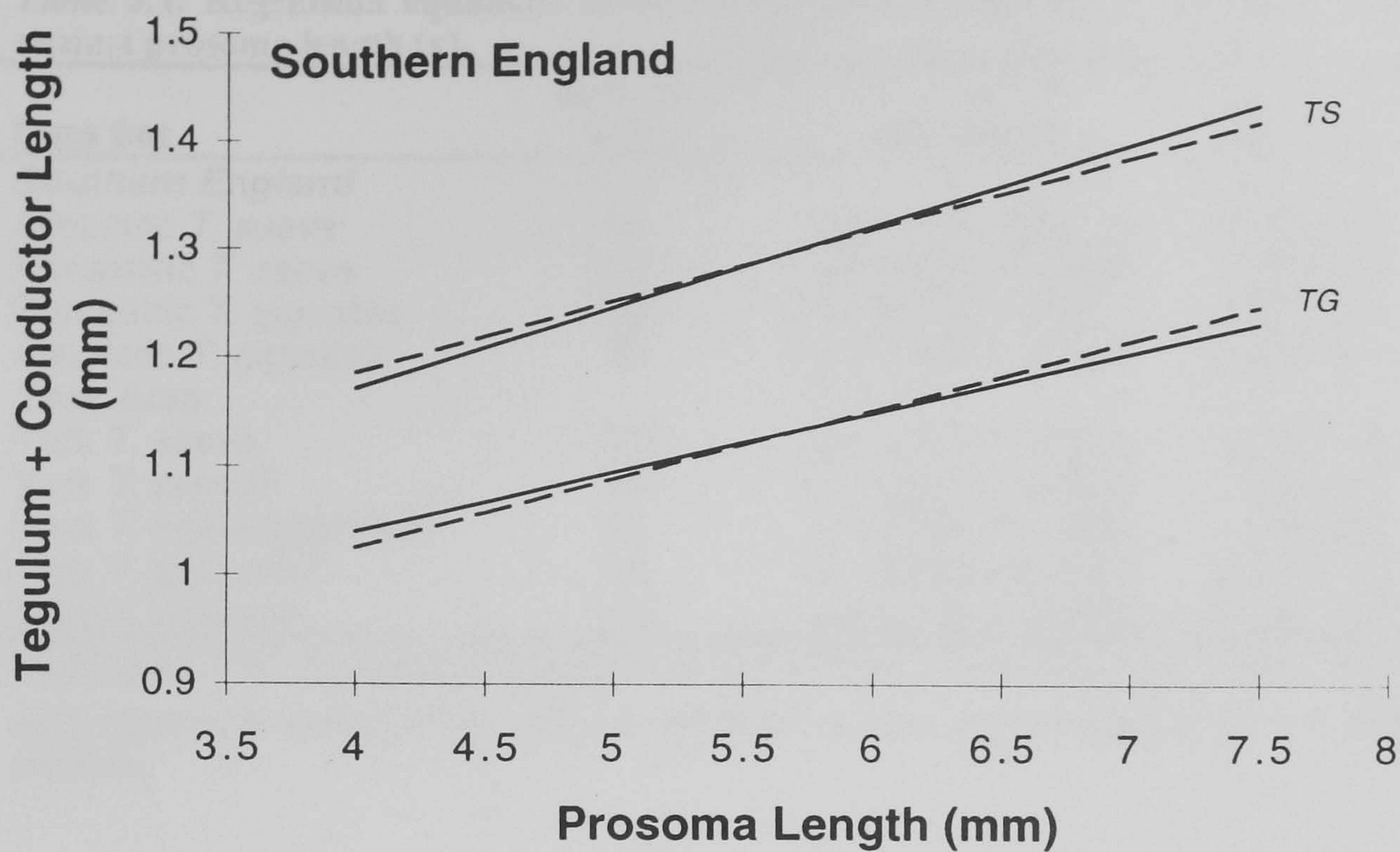


Figure 3.3. Scatterplots of conductor + tegulum length against prosoma length. A) Data for Southern England. Solid squares = allopatric *T. saeva;* solid triangles = parapatric *T. saeva;* open squares = allopatric *T. gigantea;* open triangles = parapatric *T. gigantea*; open circles = known F1 hybrids. B) Data for the York area. Solid squares = *T. saeva;* solid triangles = *T. saeva?;* open squares = *T. gigantea;* open triangles = *T. gigantea?;* shaded circles = putative *T. saeva/gigantea* hybrids (unidentified).

A)



B)

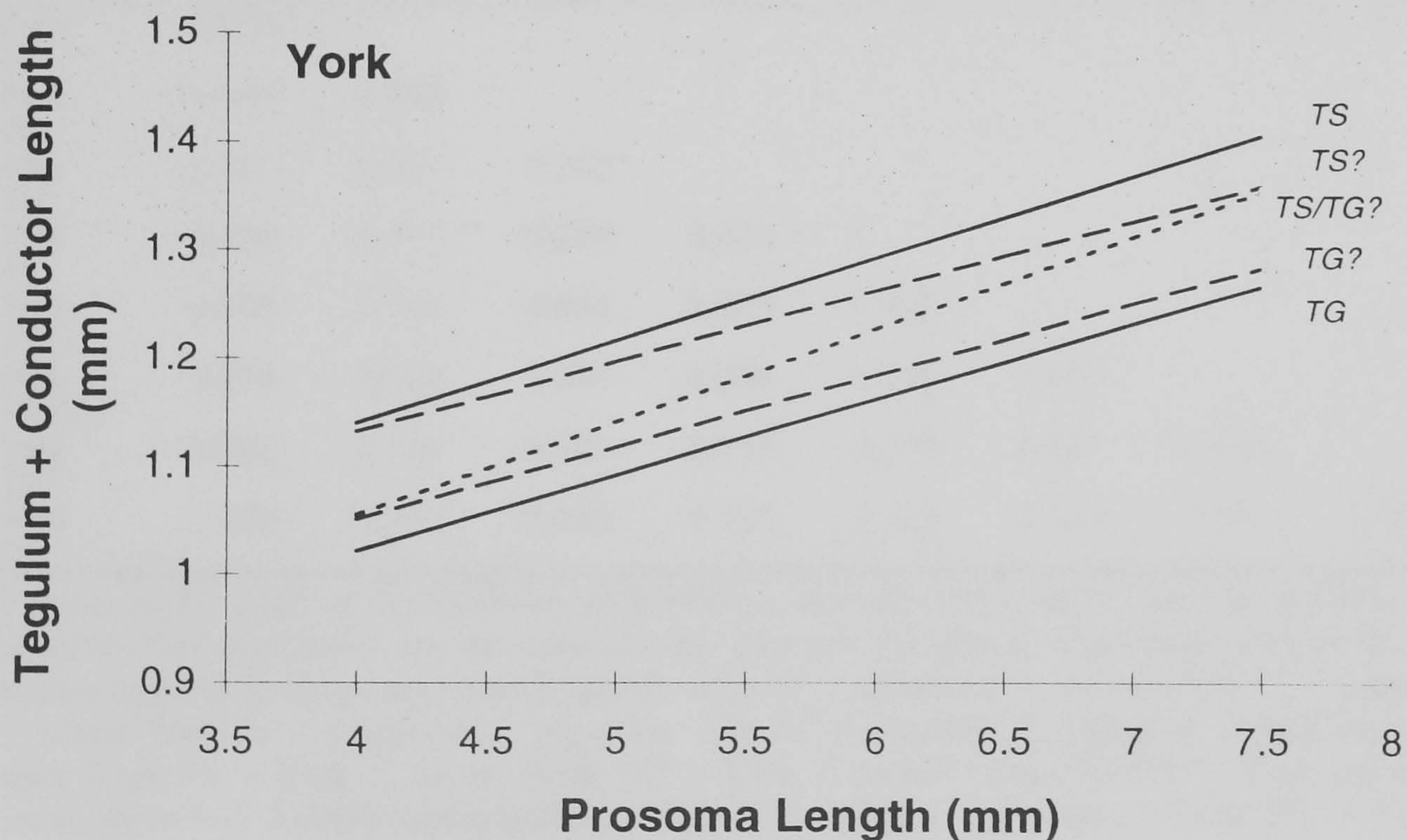


Figure 3.4. Plots of regression lines of conductor + tegulum length against prosoma length. Individual data points shown in figure 3.3. Regressions were performed for all data sets excluding the known hybrids from Southern England. A) Data for Southern England. *TS* = *T. saeva;* *TG* = *T. gigantea;* solid lines = allopatric samples; broken lines = parapatric samples. B) Data for the York area. *TS* (solid line) = *T. saeva;* *TS?* (broken line) = *T. saeva?;* *TG* (solid line) = *T. gigantea;* *TG?* (broken line) = *T. gigantea?;* *TS/TG?* (finely broken line) = putative *T. saeva/gigantea* hybrids (unidentified). Sample sizes are given in Table 3.1.

Table 3.1. Regression equations derived from plots of tegulum + conductor length (y) against prosoma length (x).

Data Set	Sample Size	Equation	RMS x10³ (*df*)
Southern England			
Allopatric *T. saeva*	30	y = 0.878 + 0.073x	0.669 (28)
Parapatric *T. saeva*	40	y = 0.928 + 0.064x	1.283 (38)
Parapatric *T. gigantea*	39	y = 0.775 + 0.062x	0.580 (37)
Allopatric *T. gigantea*	30	y = 0.823 + 0.054x	0.766 (28)
York Area			
York *T. saeva*	145	y = 0.835 + 0.076x	1.659 (143)
York *T. saeva?*	43	y = 0.871 + 0.065x	1.692 (41)
York *T. saeva/gigantea?*	31	y = 0.713 + 0.086x	1.752 (29)
York *T. gigantea?*	15	y = 0.785 + 0.066x	2.155 (13)
York *T. gigantea*	40	y = 0.742 + 0.070x	1.163 (38)

In all cases *P* for the regression coefficient was < 0.001. The residual mean squares (error) for each regression (multiplied by 10^3) are given in the final column along with their degrees of freedom.

Table 3.2. Comparison of residual mean squares from regression for each of the tegulum + conductor length against prosoma length data sets.

	Allo *TS*	Para *TS*	Para *TG*	Allo *TG*	York *TS*	York *TS?*	York *TS/TG?*	York *TG?*
Para *TS*	**0.038**							
Para *TG*	0.338	**0.009**						
Allo *TG*	0.361	0.079	0.212					
York *TS*	**0.003**	0.179	**<0.001**	**0.010**				
York *TS?*	**0.006**	0.196	**0.001**	**0.015**	0.450			
York *TS/TG?*	**0.006**	0.182	**0.001**	**0.016**	0.400	0.452		
York *TG?*	**0.005**	0.106	**0.001**	**0.011**	0.220	0.267	0.309	
York *TG*	0.066	0.382	**0.018**	0.127	0.101	0.123	0.117	0.070

Figures in the body of the table are probabilities, derived from *F*-tests, that the residual mean squares from regression are the same for the data sets compared. Significant probabilities are shown in bold. Data for southern England: Allo *TS* = allopatric *T. saeva;* Para *TS* = parapatric *T. saeva;* Para *TG* = parapatric *T. gigantea;* Allo *TG* = allopatric *T. gigantea.* Data for the York area: York *TS* = York *T. saeva;* York *TS?* = York *T. saeva?;* York *TS/TG?* = York putative *T. saeva/gigantea?* hybrids (unidentified); York *TG?* = York *T. gigantea?;* York *TG* = York *T. gigantea*. The residual mean square values are given in Table 3.1.

southern England (although the comparison with the allopatric *T. gigantea* is not quite significant). Also, with the exception of the 'good' York *T. gigantea*, all data sets from the York area are more variable than those from southern England (with the exception of the parapatric *T. saeva* samples which did not differ from the York data sets). Overall there was no difference in variability among the York data sets.

The ANCOVA performed on all the data sets showed that the regression lines fitted to each data set did not differ in slope ($F_{(8, 385)}$ = 1.86; n.s.). However, the elevations of the regression lines (the means of tegulum + conductor length adjusted for the covariation in prosoma length) were highly significantly different ($F_{(8, 403)}$ = 153.72, p << 0.001).The adjusted means for tegulum + conductor length, along with their standard errors, estimated 95% comparison limits, and the unadjusted means, are given in Table 3.3. The adjusted means, along with their estimated 95% comparison intervals, and bars highlighting the homogenous groupings, are shown in Figure 3.5. This figure very powerfully demonstrates the pattern of variation among the data sets with respect to this important character. The adjusted means follow the pattern expected from the morphological groupings, with the allopatric *T. gigantea* and allopatric *T. saeva* from southern England showing the greatest differences in this measurement, and with the York *T. saeva/gigantea?* falling approximately half way between these values. Allopatric and parapatric *T. saeva* from southern England were not significantly different, however the York *T. saeva* showed a significantly smaller mean than *T. saeva* samples from southern England. The mean for the York *T. saeva?* was significantly smaller again than the York *T. saeva*. Similarly, the York *T. saeva/gigantea?* had a significantly smaller mean than the York *T. saeva?*. The *T. gigantea* and *T. gigantea?* data showed a similar pattern (increasing in adjusted mean TCL length from allopatric *T. gigantea* to York *T. gigantea?*). The York *T. gigantea* did have a slightly greater adjusted mean than the allopatric and parapatric *T. gigantea* from southern England, however this difference was not statistically significant (*cf. T. saeva*). Interestingly, the York *T. gigantea* and the York *T. gigantea?* were

not significantly different and the York *T. gigantea?* and the York *T. saeva/gigantea?* also showed overlapping comparison intervals. The overlap between these data sets could reflect a greater degree of introgression affecting *T. gigantea* than *T. saeva* in York. Introgression could explain the relatively large amount of variability in the York *T. gigantea?* data (as indicated by the RMS values in Tables 3.1 and 3.2 and in the 95% comparison limits in Table 3.3), however, this could also reflect the smaller sample size and ambiguities in this morphological grouping. Overall, these results suggest that there may be some introgression in both directions in York such that *T. saeva* and *T. gigantea* are becoming 'pulled' morphologically towards each other.

Table 3.3. Tegulum + conductor length means and standard errors adjusted by ANCOVA on prosoma length, with 95% comparison limits.

Data Set	Adj. Mean	S. E.	95% Limits	Mean (Unadj.)
Southern England				
Allopatric *T. saeva*	1.311	0.007	± 0.015	1.325
Parapatric *T. saeva*	1.309	0.006	± 0.013	1.302
Parapatric *T. gigantea*	1.142	0.006	± 0.013	1.156
Allopatric *T. gigantea*	1.136	0.007	± 0.015	1.164
York Area				
York *T. saeva*	1.285	0.003	± 0.007	1.282
York *T. saeva?*	1.256	0.006	± 0.013	1.250
York *T. saeva/gigantea?*	1.214	0.007	± 0.015	1.193
York *T. gigantea?*	1.178	0.010	± 0.022	1.172
York *T. gigantea*	1.155	0.006	± 0.013	1.153

Shown for each data set: the adjusted (least squares) mean; the standard error of the adjusted mean (S.E.); 95% comparison limits generated from GT-2 multiple comparison tests using the Gabriel approximation (Sokal and Rohlf, 1995); and the unadjusted mean of tegulum + conductor length. Values in mm.

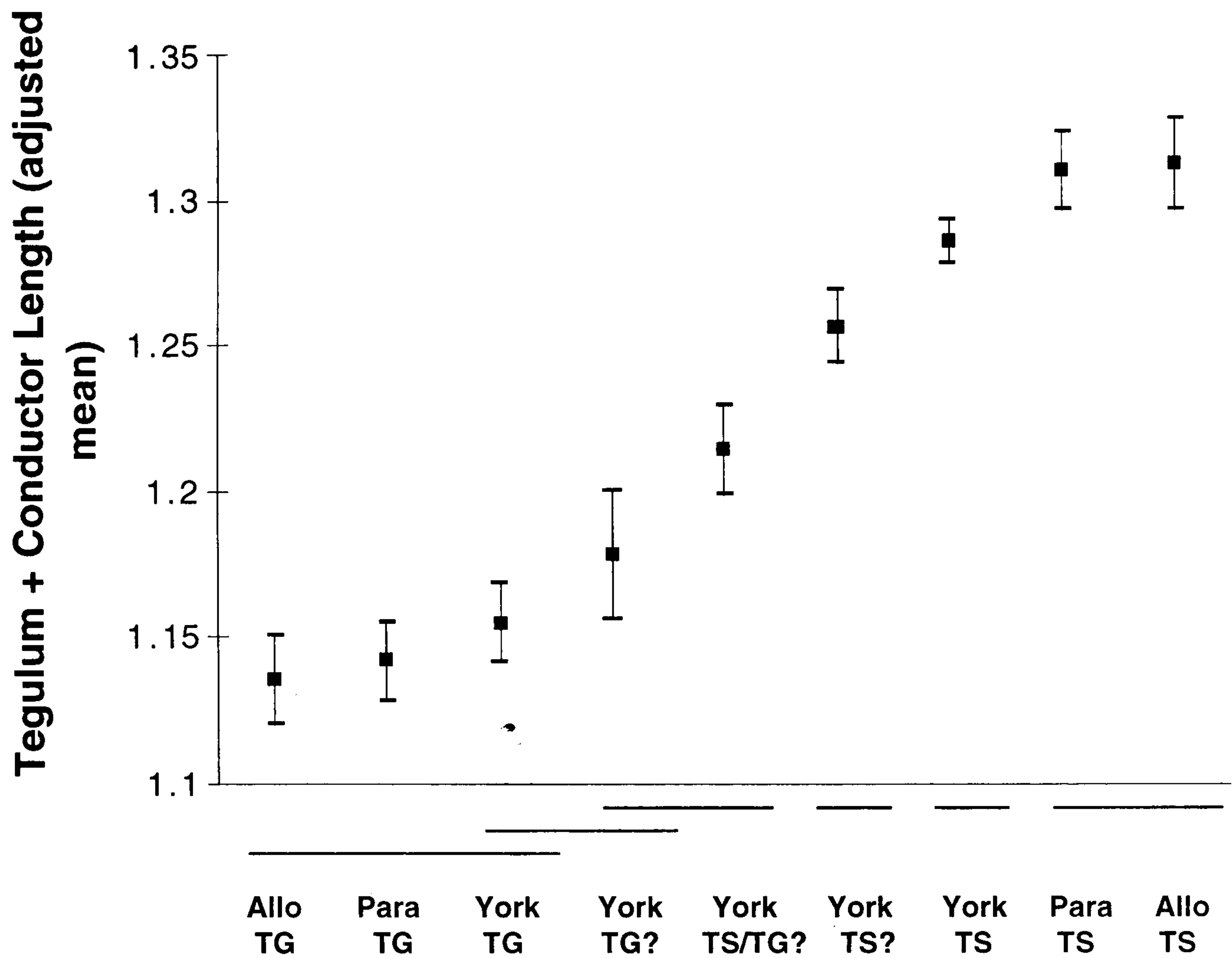


Figure 3.5. Plot of tegulum + conductor length means, adjusted for prosoma length by ANCOVA, with 95% comparison intervals. The adjusted means and intervals are plotted in order of increasing magnitude. Means whose intervals do not overlap are significantly different - homogeneous groupings are shown by bars below the chart. Data for Southern England: Allo TS = allopatric *T. saeva;* Para TS = parapatric *T saeva;* Allo TG = allopatric *T. gigantea;* Para TG = parapatric *T. gigantea*. Data for the York area: York TS = York *T. saeva*; York TS? = York *T. saeva?*; York TS/TG? = York putative *T. saeva/gigantea?* hybrids (unidentified); York TG? = York *T. gigantea?*; York TG = York *T. gigantea.*

3.3.2 Univariate Analysis and Measurement Error

Estimates of percentage measurement error (%ME) were in general low (Table 3.4 and Table 3.5), but increased markedly with decrease in mean character size (males: mean %ME = 3.04%, minimum = 0.22 (PROW), maximum = 10.19 (CL); females: mean %ME = 4.63%, minimum = 0.11 (PROL), maximum = 16.43 (APEPI)). Univariate one-way ANOVAs on the allopatric *T. saeva* and allopatric *T. gigantea* data sets from southern England provided an indication as to which variables might demonstrate real differences between the species (Table 3.4 and Table 3.5). The measurements on the prosoma ('size'), PROL and PROW, showed little between-group variation (BGV) and taken individually cannot readily separate the species (emphasising the similarity between the two species). It is interesting to note however, that males were more variable in size than females (*F*-test on PROL (allopatric *T. saeva* and allopatric *T. gigantea* measurements pooled), male variance = 0.699, female variance = 0.286: $F_{(59,73)} = 2.44$; $P < 0.001$). Examination of Table 3.4 and Table 3.5 shows that a large number of the other variables do show appreciable BGV, often despite quite large %ME, for example 0.94 (94%) of the variation in CW (males) is BGV despite an %ME of nearly 10%. For the males 12 out of 20 (60%) of the variables measured showed significant BGV, and for the females, although BGV was rarely as high as in males, 12 out of 15 variables (80%) showed significant BGV. Overall, a high degree of reliance could be placed on these variables to demonstrate real differences between *T. saeva* and *T. gigantea*.

Table 3.4. Summary statistics for all variables measured on males.

	%ME	Species	Min	Max	Mean	SD	WGV	BGV	*F*
PROL	0.40	*T. saeva* (n = 30)	4.46	7.34	6.11	0.86	0.98	0.02	0.94 n.s.
		T. gigantea (n = 30)	3.29	7.69	6.32	0.81	-	-	-
PROW	0.22	*T. saeva* (n = 30)	3.10	5.04	4.24	0.60	0.95	0.05	3.09 n.s.
		T. gigantea (n = 30)	3.02	5.40	4.50	0.56	-	-	-
TCL	0.45	*T. saeva* (n = 32)	1.14	1.42	1.32	0.07	0.38	0.62	104.63 ****
		T. gigantea (n = 33)	1.03	1.26	1.16	0.05	-	-	-
TW	5.19	*T. saeva* (n = 32)	0.50	0.65	0.59	0.04	0.07	0.93	796.08 ****
		T. gigantea (n = 33)	0.22	0.39	0.33	0.04	-	-	-
TWMX	5.39	*T. saeva* (n = 32)	0.34	0.52	0.39	0.04	0.30	0.70	147.44 ****
		T. gigantea (n = 33)	0.19	0.35	0.28	0.03	-	-	-
TWMN	2.91	*T. saeva* (n = 32)	0.22	0.32	0.27	0.03	0.96	0.04	2.30 n.s.
		T. gigantea (n = 33)	0.23	0.31	0.28	0.02	-	-	-
CL	10.19	*T. saeva* (n = 32)	0.11	0.22	0.17	0.03	0.19	0.81	271.34 ****
		T. gigantea (n = 33)	0.21	0.32	0.27	0.02	-	-	-
CW	9.69	*T. saeva* (n = 32)	0.03	0.06	0.04	0.01	0.06	0.94	921.19 ****
		T. gigantea (n = 33)	0.09	0.13	0.11	0.01	-	-	-
TARL	0.61	*T. saeva* (n = 32)	1.78	2.90	2.45	0.29	0.91	0.09	6.03 *
		T. gigantea (n = 33)	1.94	3.02	2.62	0.25	-	-	-
ETAL	2.60	*T. saeva* (n = 32)	0.20	0.31	0.26	0.03	0.95	0.05	3.53 n.s.
		T. gigantea (n = 33)	0.22	0.31	0.27	0.02	-	-	-
ETAWP	7.61	*T. saeva* (n = 32)	0.08	0.15	0.10	0.02	0.47	0.53	71.44 ****
		T. gigantea (n = 33)	0.05	0.10	0.07	0.01	-	-	-
ETAWD	4.77	*T. saeva* (n = 32)	0.11	0.17	0.14	0.01	0.38	0.62	101.80 ****
		T. gigantea (n = 33)	0.08	0.13	0.10	0.01	-	-	-
DTA	1.66	*T. saeva* (n = 32)	0.43	0.57	0.47	0.03	0.78	0.22	17.79 ****
		T. gigantea (n = 33)	0.44	0.59	0.51	0.03	-	-	-
TIBWD	0.78	*T. saeva* (n = 32)	0.67	0.96	0.85	0.07	0.65	0.35	34.08 ****
		T. gigantea (n = 33)	0.59	0.85	0.75	0.06	-	-	-
TIBWP	1.32	*T. saeva* (n = 32)	0.32	0.53	0.44	0.06	1.00	0.00	0.01 n.s.
		T. gigantea (n = 33)	0.32	0.53	0.44	0.05	-	-	-

continued ->

Table 3.4 continued.

	%ME	Species	Min	Max	Mean	SD	WGV	BGV	*F*
TLV	0.83	*T. saeva* (n = 32)	0.73	1.15	0.89	0.08	0.93	0.07	4.50 *
		T. gigantea (n = 33)	0.74	0.95	0.86	0.05	-	-	-
TWV	1.60	*T. saeva* (n = 32)	0.83	1.16	1.02	0.08	0.96	0.04	2.91 n.s.
		T. gigantea (n = 33)	0.84	1.08	0.99	0.06	-	-	-
TARLC	1.61	*T. saeva* (n = 32)	0.60	1.32	1.02	0.17	0.78	0.22	17.38 ****
		T. gigantea (n = 33)	0.76	1.48	1.19	0.16	-	-	-
TARWC	2.20	*T. saeva* (n = 32)	0.44	0.73	0.58	0.08	0.99	0.01	0.71 n.s.
		T. gigantea (n = 33)	0.47	0.71	0.59	0.06	-	-	-
COPL	0.73	*T. saeva* (n = 32)	1.03	1.36	1.21	0.09	1.00	0.00	0.11 n.s.
		T. gigantea (n= 33)	1.00	1.33	1.20	0.08	-	-	-

Values are given for the allopatric *T. saeva* and the allopatric *T. gigantea* samples from southern England. For each variable the percentage measurement error (%ME), minimum, maximum, and mean (with standard deviation) in millimetres are shown. Also provided are the comparison *F*-statistics and associated probabilities from separate univariate one-way ANOVAs on the allopatric *T. saeva* and allopatric *T. gigantea* data. The proportion of within-group variation (WGV) and proportion of between-group variation (BGV) attributable to each of these variables from the ANOVAs are also shown. (* $0.05 > P > 0.01$; ** $0.01 > P > 0.001$; *** $0.001 > P > 0.0001$; **** $P < 0.0001$; n.s., not significant). See Figure 3.1 and Figure 3.2 for abbreviations.

Table 3.5. Summary statistics for all variables measured on females.

	%ME	Species	Min	Max	Mean	SD	WGV	BGV	*F*
PROL	0.11	*T. saeva* (n = 34)	5.26	7.40	6.25	0.60	0.97	0.03	2.05 n.s.
		T. gigantea (n = 40)	5.62	7.56	6.43	0.47	-	-	-
PROW	0.30	*T. saeva* (n = 34)	3.56	5.18	4.30	0.42	0.95	0.05	4.06 *
		T. gigantea (n = 40)	3.89	5.18	4.47	0.32	-	-	-
STER1	3.00	*T. saeva* (n = 34)	0.36	0.86	0.64	0.13	0.78	0.22	20.21 ****
		T. gigantea (n = 40)	0.54	1.01	0.77	0.11	-	-	-
STER2	7.00	*T. saeva* (n = 34)	0.22	0.43	0.33	0.05	0.99	0.01	0.41 n.s.
		T. gigantea (n = 40)	0.18	0.50	0.33	0.07	-	-	-
EPIL	1.57	*T. saeva* (n = 34)	0.74	1.01	0.90	0.07	0.36	0.64	127.65 ****
		T. gigantea (n = 40)	0.65	0.86	0.73	0.05	-	-	-
AAEPI	2.70	*T. saeva* (n = 34)	0.70	1.29	0.84	0.10	0.32	0.68	152.76 ****
		T. gigantea (n = 40)	0.53	0.71	0.61	0.05	-	-	-
RWMX	1.68	*T. saeva* (n = 34)	0.48	0.72	0.60	0.05	0.53	0.47	63.93 ****
		T. gigantea (n = 40)	0.38	0.58	0.50	0.05	-	-	-
RWMN	6.30	*T. saeva* (n = 34)	0.17	0.53	0.34	0.08	0.36	0.64	126.10 ****
		T. gigantea (n = 40)	0.12	0.34	0.19	0.04	-	-	-
DA	1.03	*T. saeva* (n = 34)	0.24	0.36	0.30	0.03	0.64	0.36	40.10 ****
		T. gigantea (n = 40)	0.26	0.46	0.35	0.04	-	-	-
RLW	6.46	*T. saeva* (n = 34)	0.10	0.24	0.14	0.03	0.34	0.66	137.20 ****
		T. gigantea (n = 40)	0.16	0.26	0.21	0.02	-	-	-
AWA	7.41	*T. saeva* (n = 34)	0.17	0.32	0.22	0.04	0.81	0.19	16.94 ***
		T. gigantea (n = 40)	0.00	0.25	0.19	0.04	-	-	-
AWB	7.41	*T. saeva* (n = 34)	0.48	0.74	0.57	0.07	0.83	0.17	14.40 ***
		T. gigantea (n = 40)	0.32	0.70	0.51	0.08	-	-	-
APR	5.40	*T. saeva* (n = 34)	0.48	0.72	0.59	0.06	0.29	0.71	176.35 ****
		T. gigantea (n = 40)	0.34	0.53	0.42	0.04	-	-	-
APEPI	16.43	*T. saeva* (n = 34)	0.00	0.24	0.08	0.04	0.81	0.19	16.94 ***
		T. gigantea (n = 40)	0.06	0.43	0.13	0.06	-	-	-
EPIWMX	2.66	*T. saeva* (n = 34)	0.67	1.01	0.83	0.08	1.00	0.00	0.04 n.s.
		T. gigantea (n = 40)	0.67	1.08	0.83	0.08	-	-	-

For explanation, see legend to Table 3.4.

3.3.3 Multivariate Analysis

The principal components (PCs) derived from the MGPCA on the log-transformed data for the allopatric *T. saeva* and allopatric *T. gigantea* samples from southern England are shown in Table 3.6 (Males) and Table 3.7 (Females). The PCs represent orthogonal (uncorrelated) axes with the first PC accounting for most of the variation in the data and subsequent PCs accounting for increasingly small proportions of the variation. Clearly, for both males and females, the variation in the data could be represented quite efficiently by the first few PCs. In both sexes, the first 3 PCs accounted for 70% or more of the overall variation. However, data reduction was not the object of the exercise; MGPCA was used merely to generate uncorrelated variables as input for a discriminant function, and not to reduce the number of variables required to describe the variation. Those PCs that accounted for the majority of the variation in the data were not necessarily those that would separate the species most efficiently. An indication of which PCs would be most likely to contribute most to the separation of the two species was obtained by first applying the PC coefficients to the log-transformed data for each individual and then comparing the values for the allopatric *T. saeva* and allopatric *T. gigantea* for each PC in separate univariate one-way ANOVAs. From this it is clear that although PC1 for males (Table 3.6) accounted for 51% of the total variation in the data, taken alone it contributed virtually zero between-group variation (BGV). Examination of Table 3.6 and Table 3.7 shows that many of the later PCs, which contributed only a very small fraction to the overall variation actually exhibited high levels of BGV within this fraction. Hence, the overall variation described by each PC is not a good predictor of discriminatory power - many PCs apparently contributing little variation may be heavily weighted in the discriminant function. It is worth noting that for the males (Table 3.6), PC1 showed a positive loading for all variables and can therefore

Table 3.6. Multiple group principal component coefficients (PCs) for the allopatric *T. saeva* and allopatric *T. gigantea* data sets from southern England: MALES.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
PROL	+0.36	-0.12	-0.02	+0.07	-0.10	-0.09	-0.16	-0.29	-0.20	+0.17
PROW	+0.36	-0.11	+0.02	+0.07	-0.14	-0.16	-0.10	-0.26	-0.22	+0.08
TCL	+0.14	-0.07	+0.76	-0.62	+0.15	+0.01	+0.01	-0.01	+0.01	-0.01
TW	+0.10	+0.16	+0.15	+0.30	+0.50	-0.18	-0.10	-0.08	+0.22	+0.23
TWMX	+0.08	+0.26	+0.31	+0.44	+0.33	+0.12	-0.11	+0.32	-0.24	-0.05
TWMN	+0.15	+0.15	+0.15	+0.22	+0.07	+0.01	+0.08	-0.13	+0.35	-0.07
CL	+0.04	+0.54	+0.14	-0.02	-0.57	-0.15	-0.45	+0.23	+0.18	+0.14
CW	+0.10	+0.60	+0.03	-0.02	-0.12	+0.27	+0.48	-0.43	-0.10	-0.25
TARL	+0.27	-0.08	-0.00	+0.04	-0.09	+0.09	+0.06	+0.23	-0.12	-0.02
ETAL	+0.13	+0.02	-0.01	-0.01	-0.11	-0.30	+0.65	+0.26	+0.16	+0.50
ETAWP	+0.28	+0.32	-0.42	-0.40	+0.38	-0.46	-0.07	+0.13	-0.00	-0.23
ETAWD	+0.22	+0.13	-0.29	-0.27	+0.21	+0.65	-0.19	+0.04	+0.04	+0.39
DTA	+0.13	-0.02	-0.02	+0.01	-0.02	+0.12	-0.02	-0.12	+0.36	+0.32
TIBWD	+0.20	-0.08	-0.01	+0.03	-0.05	+0.00	-0.06	-0.11	-0.05	+0.04
TIBWP	+0.32	-0.07	+0.01	+0.08	-0.08	-0.09	-0.01	-0.13	-0.15	-0.08
TLV	+0.16	+0.05	+0.03	+0.04	-0.01	-0.09	-0.03	+0.07	-0.11	-0.04
TWV	+0.16	-0.01	+0.07	+0.15	+0.10	-0.01	+0.03	-0.11	+0.05	-0.00
TARLC	+0.38	-0.11	-0.01	+0.06	-0.12	+0.25	+0.18	+0.53	-0.02	-0.27
TARWC	+0.25	-0.22	-0.03	+0.05	-0.07	+0.06	-0.08	-0.06	+0.64	-0.39
COPL	+0.17	-0.06	+0.01	+0.02	-0.08	-0.02	-0.02	-0.02	-0.14	+0.17
V	0.51	0.12	0.08	0.08	0.06	0.03	0.02	0.02	0.02	0.01
WGV	1.00	0.33	0.95	0.98	0.05	0.39	0.11	0.24	0.93	0.11
BGV	0.00	0.67	0.05	0.02	0.95	0.61	0.89	0.76	0.07	0.89
F	0.18	130.43	2.99	1.41	1098.05	99.68	493.51	201.60	4.81	512.49
P	n.s.	****	n.s.	n.s.	****	****	****	****	*	****

Table 3.6 continued.

	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19	PC20
PROL	+0.16	-0.13	+0.05	+0.38	-0.31	+0.06	-0.13	-0.25	-0.46	+0.27
PROW	-0.04	+0.07	-0.29	-0.21	-0.16	-0.15	-0.01	-0.20	+0.35	-0.57
TCL	-0.01	+0.00	-0.01	+0.01	+0.00	-0.00	-0.01	-0.00	-0.01	+0.01
TW	-0.29	-0.52	+0.14	-0.22	+0.01	+0.14	+0.01	-0.13	+0.05	+0.01
TWMX	-0.05	+0.36	-0.23	+0.35	-0.13	+0.03	+0.10	+0.07	+0.01	-0.04
TWMN	+0.65	+0.23	-0.13	-0.45	-0.04	+0.06	-0.08	-0.02	-0.12	+0.13
CL	-0.11	-0.04	+0.03	-0.05	-0.00	-0.03	-0.09	+0.01	-0.02	+0.02
CW	-0.12	-0.16	+0.02	+0.10	-0.04	+0.02	+0.08	+0.01	+0.02	-0.02
TARL	+0.06	-0.23	-0.02	-0.16	-0.20	-0.44	+0.41	+0.02	+0.32	+0.50
ETAL	-0.15	+0.23	+0.03	+0.05	-0.17	+0.06	-0.08	-0.03	-0.01	+0.01
ETAWP	+0.10	+0.03	-0.14	+0.09	+0.02	-0.06	+0.02	+0..09	-0.03	+0.00
ETAWD	-0.09	+0.22	+0.10	-0.17	-0.11	+0.04	-0.07	-0.07	+0.00	-0.02
DTA	+0.24	-0.14	-0.27	+0.48	+0.50	-0.12	+0.23	-0.01	+0.14	-0.08
TIBWD	+0.06	-0.07	-0.04	+0.11	-0.05	+0.39	-0.42	+0.54	+0.49	+0.24
TIBWP	-0.37	+0.34	-0.05	-0.22	+0.59	+0.17	+0.07	-0.14	-0.09	+0.34
TLV	+0.30	+0.13	+0.77	+0.14	+0.12	+0.17	+0.20	-0.18	+0.28	-0.16
TWV	-0.08	+0.10	+0.32	+0.03	+0.17	-0.69	-0.41	+0.32	-0.14	-0.08
TARLC	+0.07	-0.38	-0.07	-0.03	+0.18	+0.13	-0.28	-0.12	-0.18	-0.23
TARWC	-0.32	+0.18	+0.09	+0.17	-0.30	+0.07	+0.19	+0.05	+0.01	-0.06
COPL	+0.03	-0.09	+0.07	-0.18	+0.00	+0.16	+0.48	+0.63	-0.39	-0.25
V	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00
WGV	0.85	0.44	0.84	0.44	0.90	0.14	0.52	0.62	0.47	0.72
BGV	0.15	0.56	0.16	0.56	0.10	0.86	0.48	0.38	0.53	0.28
F	10.70	80.81	11.87	79.95	6.89	374.98	58.60	38.74	71.17	24.51
P	**	****	**	****	*	****	****	****	****	****

PCs derived from the pooled within-groups covariance matrix of the allopatric *T. saeva* and the allopatric *T. gigantea* samples from southern England. V is the proportion of the total variance accounted for by each PC. Also given are the comparison F-statistics and associated probabilities from separate univariate one-way ANOVAs on the allopatric *T. saeva* and allopatric *T. gigantea* data, and the proportion of within-group variation (WGV) and proportion of between-group variation (BGV) attributable to each of these variables. (* $0.05 > P > 0.01$; ** $0.01 > P > 0.001$; *** $0.001 > P > 0.0001$; **** $P < 0.0001$; n.s., not significant).

Table 3.7. Multiple group principal component coefficients for the allopatric *T. saeva* and allopatric *T. gigantea* data sets from southern England: FEMALES.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
PROL	+0.06	+0.18	+0.20	-0.03	-0.02	-0.03	-0.04	+0.20	+0.25
PROW	+0.07	+0.17	+0.17	-0.03	-0.01	+0.00	-0.03	+0.25	+0.19
STER1	+0.09	+0.36	+0.44	-0.46	-0.00	-0.21	-0.47	-0.39	-0.16
STER2	+0.08	+0.27	+0.14	+0.87	+0.05	-0.15	-0.27	-0.19	+0.02
EPL	+0.08	+0.05	+0.17	+0.04	-0.07	+0.13	+0.25	-0.21	+0.03
AAEPI	-0.02	+0.07	+0.16	+0.08	-0.12	+0.25	+0.36	-0.39	-0.22
RWMX	+0.03	-0.10	+0.19	-0.02	+0.34	-0.11	+0.21	-0.05	+0.08
RWMN	+0.04	-0.74	+0.52	+0.08	+0.24	-0.04	-0.07	-0.08	-0.03
DA	+0.09	+0.12	+0.19	-0.05	+0.15	+0.20	+0.08	+0.02	+0.78
RLW	+0.05	+0.37	+0.10	-0.02	+0.68	+0.01	+0.35	+0.22	-0.34
AWA	+0.04	+0.01	+0.21	+0.02	-0.31	-0.74	+0.31	+0.33	-0.03
AWB	-0.00	+0.04	+0.38	+0.08	-0.26	+0.48	-0.19	+0.55	-0.29
APR	-0.05	+0.08	+0.17	+0.00	-0.36	+0.09	+0.41	-0.23	+0.04
APEPI	+0.98	-0.08	-0.15	-0.02	-0.07	+0.04	+0.02	-0.00	-0.08
EPIWMX	+0.05	+0.08	+0.26	+0.02	-0.16	+0.09	+0.18	-0.01	-0.01
V	0.46	0.14	0.12	0.09	0.04	0.04	0.03	0.02	0.02
WGV	0.73	0.38	0.51	0.81	0.44	0.99	0.50	0.50	0.98
BGV	0.27	0.62	0.49	0.19	0.56	0.01	0.50	0.50	0.02
F	27.23	115.80	69.11	16.47	92.38	0.84	71.05	71.88	1.49
P	****	****	****	***	****	n.s.	****	****	n.s.

Table 3.7 continued.

	PC10	PC11	PC12	PC13	PC14	PC15
PROL	-0.23	+0.28	+0.22	+0.23	-0.10	-0.76
PROW	-0.23	+0.33	+0.17	+0.51	+0.18	+0.60
STER1	+0.06	-0.01	-0.08	-0.08	+0.01	+0.04
STER2	+0.04	-0.01	-0.07	+0.01	+0.02	+0.01
EPL	-0.05	+0.07	+0.17	-0.02	-0.87	+0.19
AAEPI	-0.08	+0.31	+0.46	-0.29	+0.40	-0.01
RWMX	+0.81	+0.10	+0.16	+0.29	+0.03	-0.06
RWMN	-0.29	+0.03	-0.11	+0.01	+0.03	-0.01
DA	+0.04	-0.12	-0.10	-0.46	+0.11	+0.10
RLW	-0.23	-0.06	-0.22	-0.08	-0.01	-0.01
AWA	+0.03	+0.02	+0.06	-0.31	+0.00	+0.07
AWB	+0.28	+0.06	-0.08	-0.19	-0.04	+0.01
APR	+0.02	+0.12	-0.70	+0.28	+0.10	-0.09
APEPI	+0.02	+0.01	-0.05	+0.00	+0.05	-0.03
EPIWMX	-0.08	-0.82	+0.27	+0.30	+0.11	-0.05
V	0.01	0.01	0.01	0.00	0.00	0.00
WGV	0.82	0.23	0.83	0.59	0.63	1.00
BGV	0.18	0.77	0.17	0.41	0.37	0.00
F	15.87	236.92	14.95	50.39	42.19	0.02
P	***	****	***	****	****	n.s.

For explanation, see legend to Table 3.6.

be interpreted as representing 'size' variation; in other words 51% of the variation in the data for males represented 'size' and the remaining 49% represented 'shape' (Lynch and Hayden, 1995; Prenter *et al*, 1995; Thorpe, 1988). For the females it was not possible to identify a PC that could be exclusively associated with 'size'; the reason for this is obscure but may partly reflect the fact that size (PROL and PROW) was less variable in females than in males (see section 3.3.2). (The three negative values in PC1 of the female data (Table 3.7) are all very small and so the presence of the negative sign

could be stochastic. Normally, however, the 'size' component not only has all positive values but the values are all of similar magnitude (Thorpe, 1988). This is not the case here.)

After converting the log-transformed data for each individual into PC scores these were then used as input for the generation of discriminant functions. The resulting discriminant function coefficients, generated using the allopatric *T. saeva* and allopatric *T. gigantea* from southern England as the 'training set', are given in Table 3.8 (for both the male and female analyses). The discriminant functions for both the males and the females provided very good separation of the allopatric *T. saeva* and allopatric *T. gigantea.* For the males the eigenvalue[1] was high (60.63). The canonical correlation between the discriminant score and the groups was also high (0.99), indicating that 98% of the variation was explained by group differences. Wilks' λ was small (Wilks' λ = 0.02, χ^2 = 218.42, *df* = 20, $P \ll 0.001$) indicating that the means of the groups were different. The discriminant function for females was not quite as efficient as for males but still good (eigenvalue = 11.56; canonical correlation = 0.96 (92% of variation due to group differences); Wilks' λ = 0.08, χ^2 = 163.20, *df* = 15, $P \ll$ 0.001). For both males and females, 100% of the training sets (the allopatric *T. saeva* and allopatric *T. gigantea*) were reclassified correctly by the discriminant functions.

The aim of the discriminant function was not to generate a classification system for unknown samples, or to provide a simple set of coefficients for other workers to classify specimens. The discriminant function has been used here as a means to ordinate and describe the multivariate variation in the data sets in relation to the 'good' allopatric specimens of *T. saeva* and *T. gigantea*

[1] The eigenvalue is the ratio of the between-groups variability (sum of squares) to the within-groups variability. The square of the canonical correlation is equivalent to the ratio of the between-groups sum of squares and the total sum of squares in a one-way ANOVA with the discriminant score as the dependent variable and group as the independent variable. Wilks' λ = the ratio of the within-groups sum of squares to the total sum of squares. (Norusis/SPSS, 1994)

Table 3.8. Discriminant function coefficients derived from the allopatric *T. saeva* and allopatric *T. gigantea* data sets from southern England after transformation to principal components.

Males			Females		
PC	Coefficient	Correlation	PC	Coefficient	Correlation
PC5	+20.52	+0.54	PC11	+20.75	+0.53
PC10	+4.13	+0.37	PC2	-3.43	-0.37
PC7	-14.47	-0.36	PC5	-5.14	-0.33
PC16	+9.38	+0.31	PC8	-6.83	-0.29
PC8	+2.89	+0.23	PC7	+4.95	+0.29
PC2	-6.13	-0.18	PC3	+2.76	+0.29
PC6	-2.77	-0.16	PC13	+14.56	+0.25
PC12	+3.84	+0.15	PC14	-16.03	-0.23
PC14	+1.92	-0.14	PC1	-1.00	-0.18
PC19	+12.67	+0.14	PC4	+1.56	+0.14
PC17	-24.46	-0.12	PC10	+6.24	+0.14
PC18	-14.77	+0.10	PC12	+6.56	+0.13
PC20	+19.22	+0.08	PC9	-0.61	-0.04
PC13	-9.22	+0.06	PC6	+0.92	+0.03
PC11	-7.63	-0.05	PC15	+0.96	+0.01
PC15	-14.44	-0.04	(Constant)	-81.58	
PC9	+7.61	+0.04			
PC3	+48.47	-0.03			
PC4	-42.97	+0.02			
PC1	+3.05	+0.01			
(Constant)	-73.74				

The unstandardized canonical discriminant function coefficients are given for each principal component (PC). The discriminant score for an individual was calculated by multiplying its score for each PC by the appropriate coefficient, summing, and adding the constant. The columns headed 'correlation' show the pooled within-groups correlations between the discriminant function and the original variables (the PCs) - these give some indication of the relative contributions of each of the variables (PCs) to the discriminant function. The discriminant coefficients have been ordered in decreasing size of correlation with the function.

from southern England. Interpretation of discriminant coefficients in terms of the original variables is always problematic as the coefficient for a particular variable is dependent upon the other variables in the function (Norusis/SPSS, 1994). Interpretation is of course even more obscure when the discriminant coefficients apply to PCs which are in turn also linear combinations of the original variables! Table 3.8 includes the pooled within-groups correlations between the discriminant function and the PC scores and provides some indication of the contribution of each PC to the discriminant function. Note that those PCs that exhibited no significant BGV in the univariate one-way ANOVAs (Table 3.6 and Table 3.7) showed little correlation with the discriminant function whereas those showing highly significant BGV tended to

correlate most highly with the discriminant function. An example of how one may attempt to interpret the discriminant coefficients in terms of the original variables is as follows: PC5 (males) was most highly, and positively, correlated with the discriminant function (Table 3.8). Positive discriminant scores are associated with *T. saeva* (see below, for example Figure 3.6A). The PC coefficients in Table 3.6 and Table 3.7 are the correlations between the original variables and the PCs (Norusis/SPSS, 1994). Examination of PC5 in Table 3.6 reveals that CL received the largest coefficient (-0.57) and TW the next largest (+0.50). Therefore individuals with large measurements of tegulum width (TW) and short conductors (CL) would tend to generate larger positive values of PC5 and therefore be associated with *T. saeva.* Conversely, individuals with a less wide tegulum and longer conductor would tend to produce smaller (more negative) values of PC5 and be associated with *T. gigantea.* Examination of the diagrams in Figure 3.1 shows that this makes intuitive sense. Such interpretation is fraught with difficulty but could have value in identifying a subset of measurements to use in deriving a simple linear discriminant function for identification purposes - similar to that proposed by Oxford and Plowman (1991). Further interpretation will not be attempted here.

Discriminant scores were generated for all individuals in each data set and the results, plotted as histograms, are shown in Figure 3.6 A and B (males) and Figure 3.7 A and B (females). Figure 3.6A shows the clear separation of the male allopatric *T. saeva* and allopatric *T. gigantea* from southern England. The parapatric *T. gigantea* exhibit similar discriminant scores to the allopatric *T. gigantea*. The parapatric *T. saeva* also exhibit similar scores to the allopatric *T. saeva*, however the parapatric distribution appears slightly displaced towards the left (towards *T. gigantea*) with a small number of individuals appearing intermediate. The known (laboratory generated) F_1 hybrid males all fall between the two species clusters (although a little closer to *T. gigantea*). As expected, the males from the York area (Figure 3.6B) show a less clear picture.

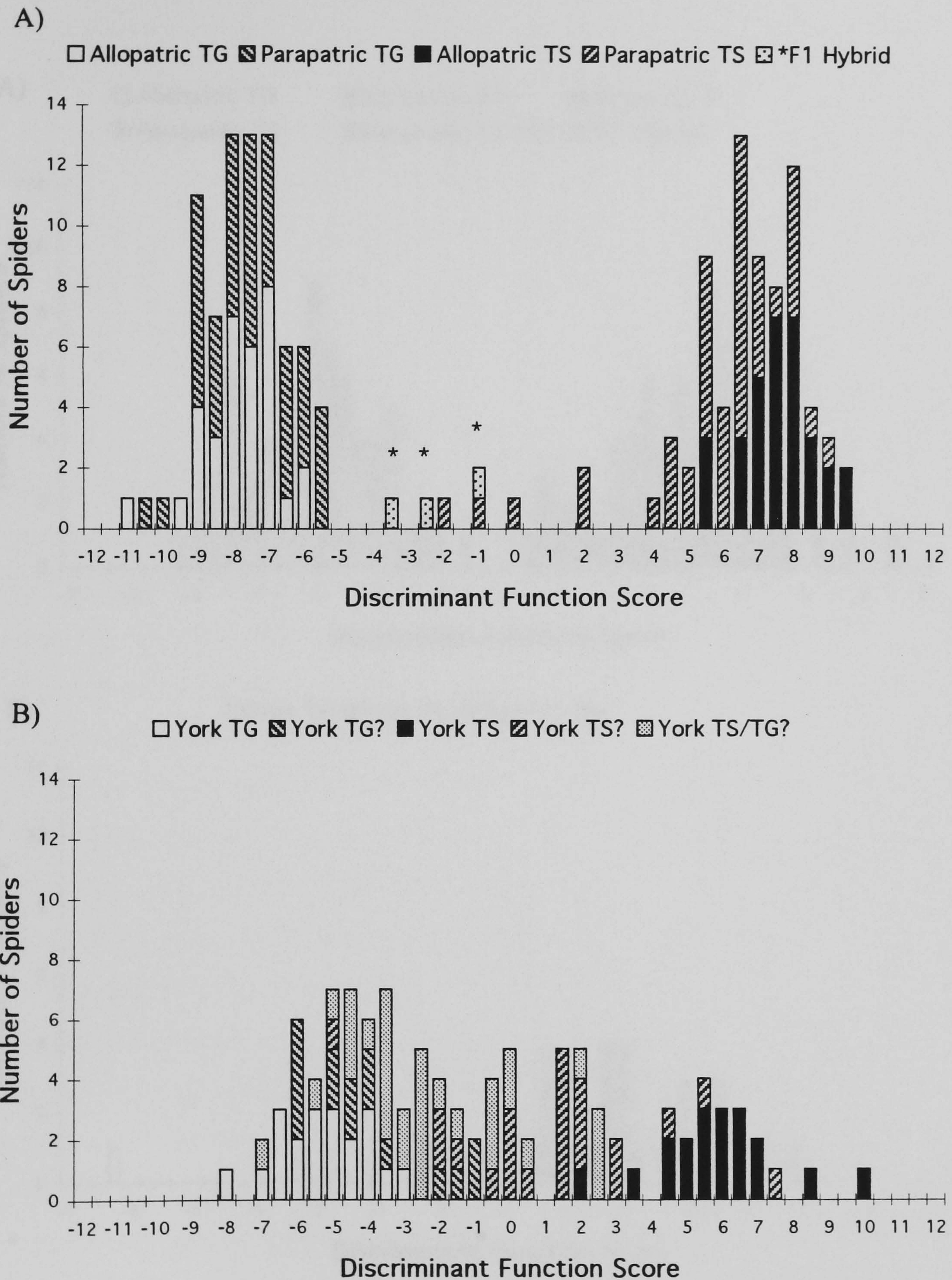


Figure 3.6. Histograms of the discriminant function scores for males.
A) Specimens from southern England. B) Specimens from the York area. A key is given above each chart.

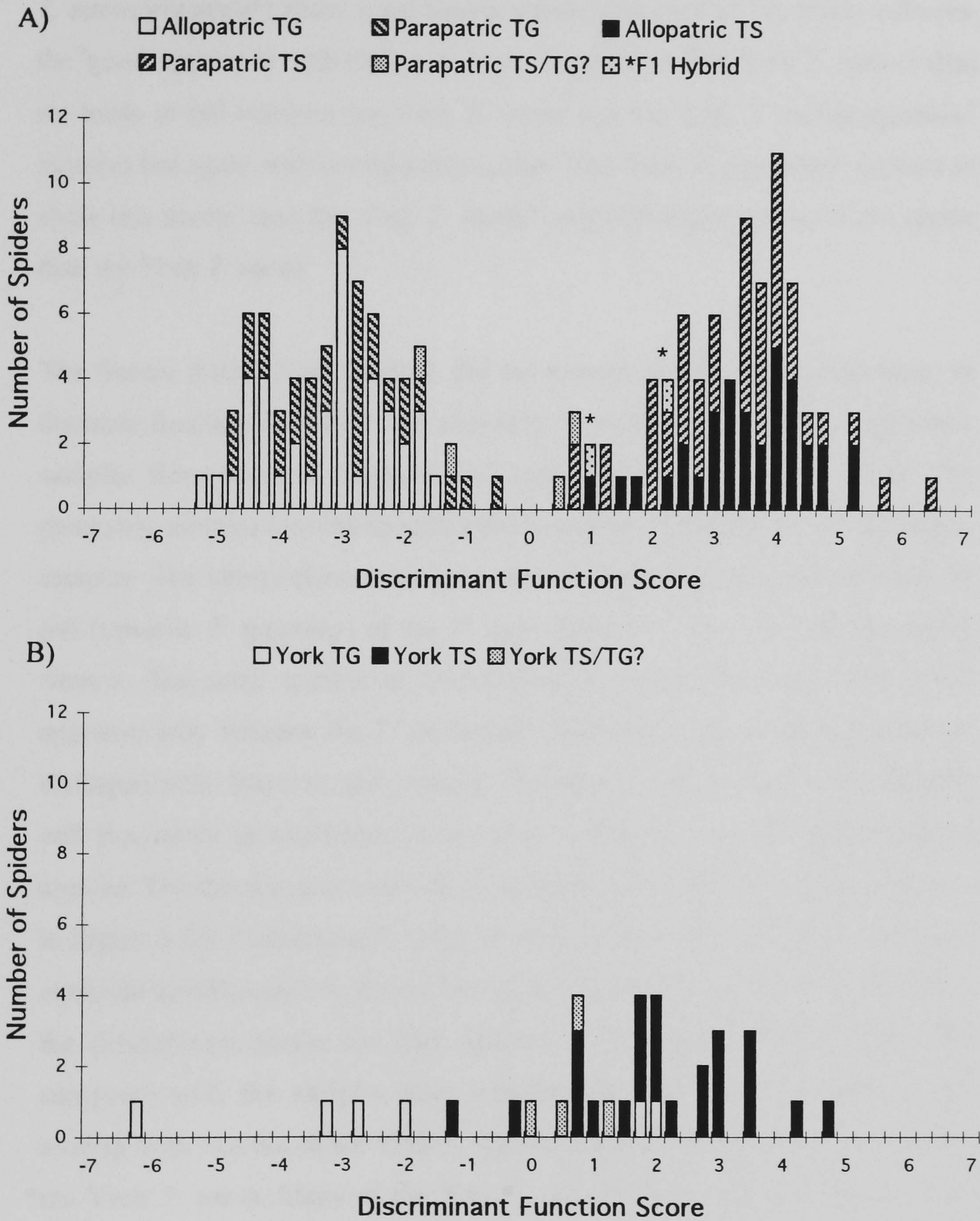


Figure 3.7. Histograms of the discriminant function scores for females.
A) Specimens from southern England. B) Specimens from the York area. A key is given above each chart.

The 'good' York *T. saeva* and the 'good' York *T. gigantea* both separate clearly but are noticeably displaced towards the centre compared with the samples from southern England in Figure 3.6A. The putative hybrid individuals (York *T. saeva/gigantea?*) show considerable scatter and tend to fall either between the 'good' species or with the 'good' York *T. gigantea*. The York *T. saeva?* data set tends to fall between the York *T. saeva* and the York *T. saeva/gigantea?* samples but again with considerable scatter. The York *T. gigantea?* appears to show less scatter than the York *T. saeva?* and falls slightly more to the centre than the York *T. saeva*.

The female discriminant function did not provide such a strong separation as the male function. However, the allopatric *T. saeva* and allopatric *T. gigantea* samples from southern England are clearly separated in Figure 3.7A. The parapatric samples for both species exhibit similar distributions to the allopatric samples. The known (laboratory generated) F_1 hybrid females fall towards the left (towards *T. gigantea*) of the *T. saeva* cluster but are not distinguishable from it. The small number of unidentified (*T. saeva/gigantea?*) females fall approximately between the *T. saeva* and *T. gigantea* clusters but again are not distinguishable from the 'pure' species. The results suggest that hybrid females will frequently be overlooked even when complex statistical approaches are applied. The discriminant scores for the females from the York area are shown in Figure 3.7B. Unfortunately relatively few females from the York area were available for this analysis, but the results do suggest that, as for the York males, the discriminant scores for both species are displaced towards each other compared with the samples from southern England. There is considerable overlap with two out of six York *T. gigantea* falling towards the centre among the York *T. saeva*. Many of the York *T. saeva* fall towards the centre, as do four out of the five unidentified York *T. saeva/gigantea?* putative hybrids (the fifth specimen being clearly classified as *T. saeva*).

Table 3.9. Descriptive statistics for the discriminant function scores of each data set.

	Males			Females		
Data Set	Sample Size	Mean	S.D.	Sample Size	Mean	S.D.
Southern England						
Allopatric *T. saeva*	32	7.78	1.04	34	3.63	1.00
Parapatric *T. saeva*	41	6.10	2.06	45	3.41	1.29
Parapatric *T. gigantea*	44	-7.28	1.26	31	-2.94	1.09
Allopatric *T. gigantea*	33	-7.55	0.96	40	-3.09	1.00
York Area						
York *T. saeva*	19	6.09	1.73	24	2.24	1.40
York *T. saeva?*	22	1.42	2.70	-	-	-
York *T. saeva/gigantea?*	31	-1.86	2.43	5	1.37	1.56
York *T. gigantea?*	15	-3.72	1.76	-	-	-
York *T. gigantea*	20	-4.97	1.28	6	-1.61	3.13

The sample size, mean, and standard deviation (S.D.) of the discriminant scores for each data set are provided.

The means, standard deviations, and sample sizes for the discriminant function scores for each data set are given in Table 3.9. The distributions of the discriminant scores for the allopatric and parapatric data sets for each species from southern England were examined for homogeneity of variance and equality of means[2] . The variances of the male allopatric and parapatric *T. saeva* data sets were significantly heterogeneous (Levene's Test: $F = 5.285$; $df = 1; 71$; $P = 0.025$) with the variance of the parapatric *T. saeva* being greater. The male allopatric and parapatric *T. gigantea* data sets showed no heterogeneity and neither did the equivalent two comparisons for the female data sets. The same data set pairs were tested for equality of means and again only the means of the male allopatric and parapatric *T. saeva* data sets were significantly different (Mann-Whitney *U*-test: $P < 0.001$) reflecting the displacement of the male parapatric *T. saeva* towards *T. gigantea*. Overall, the distribution of the mean discriminant scores for the data sets from southern England and the York area (males) were qualitatively similar to that seen for the adjusted tegulum + conductor length means (Section 3.3.1, Table 3.3), with the allopatric samples from southern England exhibiting the greatest differences in mean score, with the York samples more similar to each other,

[2] Non-parametric statistics have been presented here as the male parapatric *T. saeva* data set was left-skewed and non-normal (Shapiro-Wilks *W*-Test: $W = 0.863$; $P < 0.001$). Parametric statistics gave equivalent results.

and with *T. saeva* generally showing the greatest displacement (towards *T. gigantea*).

3.3.4 Fluctuating Asymmetry

In the characters examined, levels of asymmetry were low. After correction for measurement error, all female characters failed the tests for directional symmetry or antisymmetry and were rejected, and only four male palpal characters were retained in the analysis: ETAL (ectal tibial apophysis length), TIBWD (maximum distal tibia width), TIBWP (maximum proximal tibia width) and TLV (ventral tegulum length). The rejection of so many characters may be in part due to low levels of asymmetry, relatively high measurement errors for small characters (especially in the case of the females), small sample sizes, and the conservative nature of the tests for directional symmetry and antisymmetry. Differences in the extent of FA among the sample data sets were assessed by comparing the variances of the values of FA within each sample using a Levene's test. The results of these tests, along with the mean levels of FA for each data set and each variable, plus the standard deviations and Bonferroni 95% confidence intervals, are given in Table 3.10. Only TLV showed significant heterogeneity of variance and examination of the ordered plot in Figure 3.8 reveals that this significance was probably largely attributable to the values at the extremes of the plot (allopatric *T. gigantea* from southern England and York *T. saeva?*). The slightly higher values of FA in *T. saeva* generally, and in the York *T. gigantea?* and York *T. saeva?* samples, might be interpreted as a trend towards greater genetic instability in the samples which appear to experience most introgression; however there was no statistical support for this and overall there was no clear pattern in the data. In summary, there was little evidence of differences in FA between the sample data sets.

Table 3.10. Summary of fluctuating asymmetry in four variables measured on males and results of Levene's Tests for heterogeneity of variances between samples.

			R-L/((L+R)/2) (x10^{-2})			
Data Set	Variable	Sample Size	Mean	S.D.	Lower C.I.	Upper C.I.
Southern England						
Allopatric *T. saeva*	ETAL	30	-0.16	3.98	2.90	6.12
	TIBWD	30	-0.12	0.78	0.57	1.20
	TIBWP	30	-0.17	3.30	2.40	5.07
	TLV	29	-0.71	2.98	2.16	4.62
Parapatric *T. saeva*	ETAL	39	0.06	3.30	2.50	4.77
	TIBWD	39	-0.38	0.97	0.73	1.40
	TIBWP	39	-0.15	2.26	1.71	3.26
	TLV	39	-1.23	3.14	2.37	4.53
Parapatric *T. gigantea*	ETAL	40	0.36	3.76	2.85	5.40
	TIBWD	41	-0.41	1.35	1.02	1.92
	TIBWP	40	-0.10	2.42	1.83	3.47
	TLV	41	-0.12	1.54	1.17	2.20
Allopatric *T. gigantea*	ETAL	29	0.19	3.80	2.76	5.90
	TIBWD	29	-0.27	1.30	0.94	2.01
	TIBWP	29	-0.04	1.75	1.27	2.72
	TLV	29	0.04	1.35	0.98	2.09
York Area						
York *T. saeva*	ETAL	19	0.18	2.76	1.87	4.90
	TIBWD	19	-0.16	1.08	0.73	1.91
	TIBWP	19	-0.41	2.99	2.03	5.31
	TLV	19	-0.98	2.56	1.74	4.55
York *T. saeva?*	ETAL	22	0.28	2.65	1.84	4.47
	TIBWD	22	0.19	1.24	0.86	2.09
	TIBWP	22	0.20	2.12	1.48	3.58
	TLV	22	1.83	4.25	2.96	7.16
York *T. saeva/gigantea?*	ETAL	30	0.45	3.97	2.89	6.10
	TIBWD	30	0.19	1.49	1.08	2.29
	TIBWP	29	-0.04	1.98	1.44	3.07
	TLV	30	-0.14	2.23	1.63	3.43
York *T. gigantea?*	ETAL	15	0.00	3.10	2.02	6.08
	TIBWD	15	0.42	1.17	0.76	2.29
	TIBWP	14	-0.51	2.51	1.61	5.08
	TLV	15	-1.45	3.26	2.12	6.38
York *T. gigantea*	ETAL	20	-0.17	3.74	2.56	6.52
	TIBWD	20	-0.37	1.02	0.70	1.78
	TIBWP	20	-0.16	1.06	0.72	1.84
	TLV	20	0.39	1.56	1.07	2.72

Levene's Tests:	**Variable**	***F***	***d.f.***	***P***
	ETAL	0.36	8, 235	0.941
	TIBWD	0.81	8, 236	0.599
	TIBWP	0.77	8, 233	0.628
	TLV	3.42	8, 235	**0.001**

For each data set and each variable the sample size is given along with the mean value of individual asymmetry ((R-L)/((L+R)/2)) within each sample. The standard deviations (S.D.) and Bonferroni adjusted 95% confidence intervals of the standard deviations (C.I. - upper and lower values), as generated by the Levene's Test are also given. Mean, S.D. and C.I. are all x10^{-2}. Standard deviations are given in preference to variances for ease of plotting. The results of the Levene's tests are given below the main table. The single significant probability is shown in bold. ETAL = ectal tibial apophysis length; TIBWD = maximum distal tibia width; TIBWP = maximum proximal tibia width; TLV = ventral tegulum length.

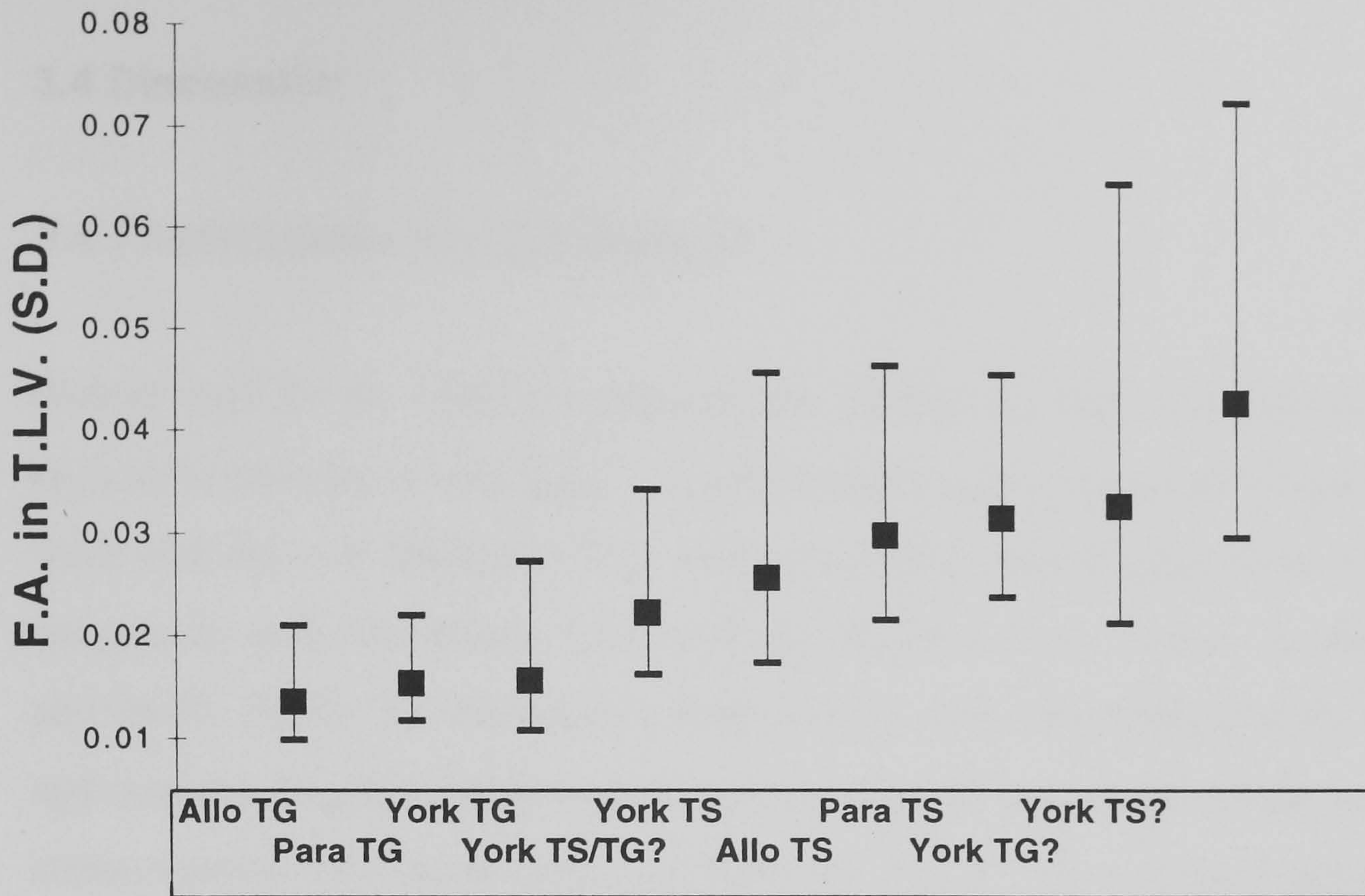


Figure 3.8. Plot of fluctuating asymmetry (FA) in male palpal ventral tegulum length (TLV) for all data sets. The standard deviation is plotted with 95% Bonferroni confidence intervals from the Levene's test output. Data for Southern England: Allo TS = allopatric *T. saeva;* Para TS = parapatric *T saeva;* Allo TG = allopatric *T. gigantea;* Para TG = parapatric *T. gigantea*. Data for the York area: York TS = York *T. saeva*; York TS? = York *T. saeva?*; York TS/TG? = York putative *T. saeva/gigantea?* hybrids (unidentified); York TG? = York *T. gigantea?*; York TG = York *T. gigantea.*

3.4 Discussion

3.4.1 Hybridization and Introgression

Oxford and Smith (1987) compared the results of their scatterplots and regression analyses of conductor + tegulum length against prosoma length from York with those of Merrett (1980). Merrett's (1980) samples came from a wide geographic area, but mostly from southern England (Pers. Comm. In Oxford and Smith, 1987). The comparison here of newly collected material from York and southern England yields similar conclusions, and has the advantage that all measurements were made by the same person. *Tegenaria saeva* and *T. gigantea* from southern England were clearly separated by this approach but there was considerable overlap in the York area material. The York material was also more variable (see Table 3.2) than that from southern England (with the exception of the parapatric *T. saeva* sample). Oxford and Smith's (1987) data were also more variable than that of Merrett (1980). However, there was no evidence in the present study that any of the sample groups from the York area were any more variable than each other; whereas Oxford and Smith (1987) found their *T. saeva* material from York to be more variable than their *T. gigantea* material (however, they did not recognize the *T. saeva?* and *T. gigantea?* as separate groupings). (The morphological groupings used for the York material, although somewhat subjective, appeared to be quite repeatable. Repeatability was tested by taking 10 individuals previously assigned to each class and then re-examining them, in an arbitrary order on a subsequent day, without reference to the previous identification; 92% of the individuals were reassigned to their original groups.) It is also interesting to note that the parapatric *T. saeva* from southern England showed more variability in conductor + tegulum length against prosoma length than the other samples from southern England.

A significant difference in variability for the relationship between conductor + tegulum length and prosoma length should result from hybridization and

introgression. Palpal morphology must have a genetic basis and an influx of genes from another species will provide new genetic variation and hence more variability in palpal phenotype. The influx of genes from another species could also increase the variance in any character as a result of effects on developmental stability (see section 3.4.3, below). Therefore, it seems likely that the increase in variation in the data for the York area samples and for the parapatric *T. saeva* from southern England stems from hybridization and introgression. It could of course be argued that *T. saeva* and *T. gigantea* might experience greater environmental stress of some form in the York area which could affect development; although it is hard to conceive what this stress might be or why the parapatric *T. saeva* from southern England should have been affected similarly. Another explanation could be that both *T. saeva* and *T. gigantea* experience a cline in the genetic variation affecting these characters - with more genetic variation as one moves north. Such a scenario also fails to explain the increased variability for the parapatric *T. saeva* data. Further, if anything, less genetic variability would be expected in the north compared with the south for species such as these that have been colonizing northwards. This is because new populations would be likely to have been established by few individuals ahead of the main population front, leading to a filtration of genotypes (Hewitt, 1996).

The means of tegulum + conductor length, adjusted for the variation in prosoma length (Figure 3.5), were also suggestive of hybridization and introgression. As expected, allopatric *T. gigantea* and *T. saeva* samples from southern England showed the maximum and minimum values respectively. The parapatric samples from southern England were not significantly different from the allopatric samples. This might appear to contradict the above assertion that the parapatric *T. saeva* samples show evidence of introgression; however a small number of individuals exhibiting hybrid characteristics could easily increase the variance of the sample data without significantly altering the mean. All the York samples had values that were intermediate to those of the samples from southern England. The values for the York samples were in the

order expected from the subjective visual groupings to which the samples had been assigned, and thus lent support to these groupings. In other words, the York *T. saeva/gigantea?* sample group fell in the middle of the plot and therefore *could* largely represent F_1 hybrids - at least on phenotypic grounds - and so on (but see section 3.4.3, below). More importantly, all the York samples including the 'parental types' - the 'good' *T. saeva* and the 'good' *T. gigantea* - showed values of adjusted tegulum + conductor length that were intermediate compared to those from southern England. This suggests that all the York samples show some evidence of introgression and convergence for this character.

The discriminant analyses, generated using the allopatric samples from southern England, provided a morphometrical score for each individual. The distributions of these discriminant scores, for the males (Figures 3.6 A and B), concurred with the data from the analyses of tegulum + conductor length. These two results were relatively independent because tegulum + conductor length (TCL) contributed (surprisingly) little to all the principal components (Table 3.6) except PC3 and PC4, with these two PCs showing little correlation with the discriminant function (Table 3.8). Allopatric males of the two species from southern England were cleanly separated, and there was no significant difference in the mean or variance of the discriminant scores for the allopatric and parapatric *T. gigantea* samples. However, the allopatric and parapatric *T. saeva* samples did differ, with the parapatric *T. saeva* showing an increased variance and a mean displaced towards *T. gigantea*. The parapatric *T. saeva* sample showed a clear tail towards the centre of the chart with a number of individuals showing increasing degrees of intermediacy. The end of this tail coincided with the discriminant scores of the known F_1 hybrids. It is therefore quite clear, on morphological grounds, that the parapatric *T. saeva* from southern England has experienced a degree of introgression of *T. gigantea* genes. The parapatric *T. gigantea* shows no evidence of introgression. The discriminant scores for the males from the York area samples, just as for the

adjusted tegulum + conductor means, were closer together compared with the samples from southern England indicating that most of the individuals had probably experienced some introgression.

The discriminant function scores for the females told a similar story (Figures 3.7 A and B), although comparisons with the York material were limited by the small sample sizes. The separation of the females was not so great as for the males and the mean and the variance of the parapatric female *T. saeva* from southern England were not significantly different from that of the allopatric *T. saeva*. However, careful examination of Figure 3.7A suggests that the parapatric *T. saeva* may have been slightly displaced towards *T. gigantea*. All previous studies on *T. saeva* and *T. gigantea* have failed to identify female hybrids. The inclusion of known F_1 hybrids, not only proves that they exist (Chapter 6), but shows that they are very hard to recognize on morphometrical grounds. The two known F_1 hybrid females grouped with the *T. saeva* samples from southern England, and the few females that had been unidentified (parapatric *T. saeva/gigantea?*) either fell with the *T. saeva* or with the *T. gigantea* samples (Figure 3.7 A). Females of these species are more difficult to identify than males, both on visual and morphometrical grounds, and it seems likely that females of hybrid origin will have been overlooked in all previous studies and in this one. This is probably partly because all these studies have, for convenience, focused on the external features of the epigyne - which are relatively simple. It might be profitable in future to include measurements from the internal features of the epigyne (after cleaning and clearing in bleach or clove oil).

The discriminant analyses had some advantages over the analyses of tegulum + conductor length in that they provided a composite morphometrical score based upon a range of characters, which were uncorrelated after the MGPCA. Although this statistical treatment makes it difficult to interpret which characters have contributed to the discriminant score, the lack of correlation in

the variables contributing to the discriminant scores does imply that the original characters have been transformed in such a way that those contributing most are likely to be under some degree of independent genetic control. Intermediate discriminant scores therefore derive from a set of largely independent characters and therefore suggest strongly that these individuals have arisen via hybridization (Oxford and Plowman, 1991). An individual may display an intermediate score because it has intermediate characters or because it has a mixture of 'good' features from both parental species. Which of these is true will depend on the nature of the genetic control of a character, and whether the individual is a first or subsequent generation hybrid (Oxford and Plowman, 1991). A discriminant score based upon a suite of characters is also less likely to overlook individuals with some hybrid features which may be otherwise missed by examining only one or two characters. It is interesting to note that the known male hybrids tended slightly towards *T. gigantea* in the discriminant analyses (Figure 3.6A), whereas the known female hybrids tended slightly towards *T. saeva* (Figure 3.7A). Although the sample sizes were very small, this may indicate that dominance, with respect to at least some of the key phenotypic characters used in species separation, is in different directions in the two sexes.

The previous multivariate analysis of *T. saeva* and *T. gigantea* in Yorkshire by Oxford and Plowman (1991) used specimens from the York area as the training set in a linear discriminant function analysis. The present analyses have the advantage in that relatively 'pure' samples, the allopatric *T. saeva* and the allopatric *T. gigantea* from southern England, were available to act as a reference to compare with the other data sets. The results suggest that very little of the material from the York area may be 'unsullied' by introgression.

3.4.2 Reinforcement/Character Displacement of Morphology?

The results of these analyses, as discussed above, provide clear evidence of hybridization and introgression between *T. saeva* and *T. gigantea. Tegenaria saeva* shows most evidence of having experienced introgression both in southern England and the York area, although all samples from the York area reveal some evidence of hybrid origins. There is no evidence that the parapatric samples from southern England show greater differences between the morphology of the species, or decreased variance in morphology, compared with the allopatric samples. Indeed, for the parapatric *T. saeva* the converse is true. The obvious difficulty with inspecting the data for evidence of reinforcement is that introgression acts to move the distribution of the data for one species towards the other, whilst reinforcement should act to move the distribution apart. The composite action of these two processes would complicate the distributions. However, it seems likely that if reinforcement was a significant force, then although the distribution of the data for the species in question would show a long tail towards the distribution of the other species due to limited introgression, the mean (or more properly the mode) should still show a displacement away from the other species. Obviously this was not the case. If reinforcement of mechanical isolation has been occurring to any degree then its effect has clearly been swamped by introgression.

3.4.3 Fluctuating Asymmetry

The genealogy of the individuals in each sample group was not known. The multivariate analysis and the analysis of combined conductor + tegulum length provided morphometrical evidence that there were relatively few intermediate individuals in southern England. Therefore, although some individuals from the parapatric samples (particularly of '*T. saeva*') showed evidence of hybrid origins, it was likely that many of the specimens in these samples were not of hybrid origin. Similarly, the York samples had been grouped according to the subjective visually identified groupings described previously. These groups were determined purely from a cursory human interpretation of phenotype and may not truly reflect genotype. Although it is tempting to consider the *T. saeva/gigantea?* grouping as largely consisting of F_1 hybrids, and the *T. saeva?* and *T. gigantea?* groupings as largely containing individuals derived from degrees of backcrossing, this is simplistic and can only be partly true. (It is certainly not accurate because the adjusted means of conductor + tegulum length, and in particular the distribution of the discriminant scores, suggested that very little of the York material should be of 'pure' parental type.) Nothing is known about the genetics (number of genes, dominance or linkage) determining the structure of the external genitalia in these species and therefore it is quite likely that F_1 hybrids, and in particular higher generation progeny, might exhibit an array of phenotypes. By way of example, in a hybrid zone of the tree frogs *Hyla cinerea* and *H. gratiosa*, parental types, F_1 and backcross progeny were easily identified using a battery of allozyme markers. However, a discriminant analysis which cleanly discriminated the parental species, failed to distinguish 18% of true F_1 hybrids from the parental species, 27% of backcrosses were not distinguished from F_1 hybrids, and 50% of *H. gratiosa* backcross progeny and 56% of *H. cinerea* backcross progeny were indistinguishable from the respective parental species (Avise, 1994).

The above discussion high-lights a weakness in the analysis of FA presented here. If FA is negatively correlated with developmental stability, and if developmental stability is *genetically* determined, then the lack of knowledge about the genetic make-up of the samples being compared is a serious limitation. As Palmer and Strobeck (1986) emphasize: 'FA as a measure of developmental stability is a very small signal easily lost in a tumultuous sea of entropic forces'. This means that if FA is to be used as a measure of developmental stability then the comparisons must be stringently defined. The presence of parental types in the parapatric samples from southern England would reduce the estimate of FA and could easily obscure the signal. In order to explore properly the effects of hybridization on developmental stability through FA analysis, in *T. gigantea* and *T. saeva*, controlled experiments are necessary. Large numbers of F_1 and backcross progeny should be generated in the laboratory from the allopatric parental species such that comparisons can be made between samples of known hybrid origin and control samples of the parental species that have been maintained under the same conditions (and with attempts to avoid inbreeding which might inflate FA scores in the controls). This would allow fair comparisons by eliminating uncertainty in terms of both genealogy and environmental effects.

Although the FA analysis presented here clearly suffered from limitations which contributed to non-conclusive results, what little trend there was in the data for TLV (Figure 3.8) might be suggestive of some of the conclusions that a more stringent approach could generate. Note that the York *T. gigantea?* and York *T. saeva?* samples, although not significantly different, did show a trend towards a greater degree of FA. In particular note that this was greater than for the York *T. saeva/gigantea?* samples. Although no statistical reliance can be placed on this trend, such a situation might be predicted on theoretical grounds. Let us assume that the *T. saeva/gigantea?* sample does largely contain F_1 hybrids, and that the *T. saeva?* and *T. gigantea?* samples largely consist of backcross progeny. F_1 hybrids would be expected to show greater degrees of

FA than the parental species due to disruption of 'coadapted gene complexes' from the parental species (Palmer and Strobeck, 1986). However, F_1 hybrids would also exhibit maximal heterozygosity which is known in general to have positive effects on developmental stability (Palmer and Strobeck, 1986). Backcross progeny would exhibit varying degrees of heterozygosity, less than F_1's, but would still experience the effects of the disrupted coadapted gene complexes - therefore they might be expected to display a greater degree of FA.

4 Allozyme Markers in *T. saeva* and *T. gigantea*

4.1 Introduction

4.1.1 Background

Until 1966 there were no accurate methods of estimating the amount of inherited variation in natural populations. Previous estimates were based on either morphological variation or the frequency of some particular class of variant such as lethal alleles or chromosomal polymorphisms (Avise, 1994). The situation changed with the introduction of starch gel electrophoresis as an efficient means of separating biomolecules. It was soon developed as a tool for detecting enzyme and protein variants in population samples by Harris (1966) in humans and by Hubby and Lewontin (1966) in *Drosophila pseudoobscura*.

Allozyme electrophoresis takes advantage of the fact that amino acids possess acidic (COO^-, primarily aspartic acid and glutamic acid) and basic (NH_4^+, primarily lysine, arginine, and histidine) groups on their side chains (Avise, 1994). Therefore each protein will carry a net positive, net negative or no net charge at a particular pH. The pH at which a protein has equal numbers of positive and negative charges or no net charge is its iso-electric point, which is characteristic for that protein structure (Richardson *et al.*, 1986). At a fixed point (the origin), the sample is applied to a medium (a gel) containing buffer. During electrophoresis the charged molecules migrate. The rate and direction of the migration depends, simply, upon: 1) the charge on the protein molecule - itself dependent on pH; 2) the ionic strength of the buffer - the higher the strength the slower the migration and the sharper the bands, but more heat is generated; 3) the size and shape of the protein; 4) the resistance to movement provided by the gel (some reciprocal function of pore size). Hence different

proteins will migrate to different locations and all of these variables can be modified to increase the resolution (Richardson *et al.*, 1986). The gel can be starch, acrylamide, agar or a cellulose acetate membrane on a rigid Mylar backing (Searle, 1983). Cellulose acetate plates have advantages over other methods in that they are less messy, non-toxic, simple to use, minimise the quantities of stain reagents required, and require very little sample volume; they are however expensive (Richardson *et al.*, 1986).

After electrophoresis, gels are stained so as to reveal the locations of specific enzymatic proteins. Enzyme-specific stains are applied in different ways according to the gel medium used. When using cellulose acetate plates, the stain reagents are usually mixed with molten agarose and applied to the membrane surface. The stain works in one of three ways: 1) the enzyme of interest converts a substrate directly into a visible product; 2) the substrate is converted into non-visible product that is made visible by other histochemicals; 3) the substrate is converted into a non-visible product that is converted by a linking enzyme into another product that can be made visible (Richardson *et al.*, 1986). The banding pattern produced depends upon the number of loci present and the quaternary structure of the enzyme (if there is a cytoplasmic and a mitochondrial locus (the enzyme is located in the mitochondrion but coded in the nucleus) for a particular enzyme then these structures may be different for the two loci). Homozygotes always produce one strong band. For a monomeric enzyme a heterozygote is revealed as two bands. For a dimeric enzyme a heterozygote is generally revealed as 3 bands; two weaker outer bands representing the homodimer (A_1A_1 and A_2A_2) and a central band staining with approximately twice the intensity and representing the heterodimer (A_1A_2). For a tetrameric enzyme five bands should be produced with the central one staining most intensely, and so on (see for example Richardson *et al.*, 1986; Hebert and Beaton, 1993; Murphy *et al.*, 1990). This co-dominance (the ability to distinguish homozygotes and heterozygotes) means that allele frequencies can be deduced simply by counting the number of copies of the allele in the sample. A number of further points need to be made. Firstly, the ability to

'count alleles' may complicated by the presence of activity variants (which may therefore not show co-dominant banding patterns), mobility variants, by bands that are artefacts of sample storage and treatment, by bands resulting from non-specific staining, and by the joint expression of more than one locus (which may make the assignment of alleles to a particular locus unclear). Secondly, allozymes represent the *alleles* of a particular enzymatic *locus* and differences between these alleles are recognized by the different positions to which they migrate on the gel. Variants of an enzyme that migrate to different locations during electrophoresis are known as electromorphs. However, because the isoelectric point of a protein is mainly determined by the charge on only a few amino acids, many allelic products that are different will produce the same electromorph and will therefore not be distinguishable. However, although some allelic variation will be missed in allozyme electrophoresis, observed differences are *real* and reflect *real* differences in genotype (Assmann and Weber, 1997). Finally the distinction between isozymes and allozymes should be made. Isozymes are simply variants of a particular enzyme, which may be genetically determined or result from some form of modification post-translation. Therefore isozymes are not always genetically determined and further, those that are genetically determined do not always reflect allelic variation at a particular locus, and are therefore not all allozymes (Murphy *et al.*, 1990).

Because allozymes reflect allelic genetic variation at specific loci they are suitable as genetic markers. Genetic markers like these are useful for the identification of individuals, the analysis of population structure analysis, the delineation of species boundaries and in phylogenetic reconstruction (Avise, 1994). Developments in molecular biology during the last three decades have led to an ever increasing range of new molecular (DNA) markers such as restriction fragment length polymorphisms (RFLPs), 'variable numbers of tandem repeats' (VNTRs - minisatellites or DNA fingerprints and microsatellites), randomly amplified polymorphic DNA (RADPs), and more recently amplified restriction fragment length polymorphisms (AFLPs). Rapid

techniques to read complete DNA sequence data - the ultimate markers - are also now widely used. These new markers have increased the range and resolution of population and phylogenetic analyses from interindividual and kin relationships (by generating large numbers of hypervariable markers) to deep phylogenetic nodes (slowly evolving sequences). However, allozymes often still have advantages over these DNA-based techniques in that they are relatively cheap, relatively quick (a very large number of individuals or loci can be typed in one day), usually require little background or preparative study and are generally well characterised (the nature of the band is known - not true of RAPDs for example). They are also co-dominant (see above), and are generally able, at least *en masse*, to be regarded as selectively neutral (Richardson *et al.*, 1986). Of course allozymes will not always be selectively neutral (and the adherence to the 'myth' of neutrality has been strongly criticised by Cain (1983)), for instance, if a species range extends from one climatic extreme to another one might expect to see a cline formed at some loci which possess alleles that have been selected for better performance under different environmental conditions (Avise, 1994). However, under most study conditions allozyme allele frequencies can be considered to result solely from the actions of drift and mutation (Avise, 1994; Richardson *et al.*, 1986). This point means that allozymes are particularly useful (although microsatellites may be better) for analysing population structure and constructing null models against which patterns of selection or gene flow can be examined for other genetic systems (Gillespie and Oxford, 1998; Mithen *et al.*, 1995). They should also be stable species markers.

Allozymes are well suited to studies of speciation and hybrid zones for a number of reasons. Firstly, because it is generally possible to screen a large number of systems, and hence a large and varied portion of the nuclear DNA, relatively quickly. The populations or species involved are, by definition, genetically differentiated and therefore it is likely that multiple markers will easily be uncovered for characterizing the hybrid zone. Secondly, because hybrid zones involve interactions between independently evolved genomes, the

effects of processes such as recombination and selection might be exaggerated - for example, leading to the appearance of new alleles. Thirdly, the combination with cytoplasmic markers (mtDNA) is a powerful tool in revealing asymmetries in hybridization and gene flow (Avise, 1994; Hewitt, 1988). A major advantage of nuclear markers, such as allozymes, is that they may make it possible to identify definitively parental types, and hybrids (F_1, backcross etc.) from natural populations, something that is not always possible from morphology (see Chapter 3).

Quite a large number of studies on spiders have employed allozyme electrophoresis. Most of these studies have focused simply on describing either the degree of genetic variability (heterozygosity) in and among populations of certain species, or the degree of genetic differentiation between related species (for example: Elliot *et al.*, 1982; Lubin and Crozier, 1985; Porter and Jakob, 1990; Smith, 1986; Steiner and Greenstone, 1992; Terranova and Roach, 1987a, 1987b, among others). Terranova and Roach (1987a) constructed an electrophoretic key to distinguish among species of the genus *Phidippus* (Salticidae). More recent studies have examined gene flow and biogeography. Ramirez and Fandino (1996) estimated rates of gene flow between populations of *Metepeira ventura* (Araneidae). Ramirez and Beckwitt (1995) used allozyme studies to reconstruct the phylogeny and historical biogeography of the genus *Lutica* (Zodariidae) in the Californian Channel Islands. Ramirez and Froehlig (1997) examined gene flow between populations of the trapdoor spider *Aptostichus simus* (Cyrtaucheniidae), and Johannesen *et al.* (1998) have examined relatedness and gene flow in *Eresus cinnaberinus* (Eresidae). Also, Gillespie and Oxford (1998) recently used allozyme data to generate a background model of population structure against which to investigate the role of selection on colour morph frequencies in the Hawaiian happy-face spider *Theridion grallator* (Theridiidae). Allozyme investigations of hybridization and gene flow between different spider species are lacking. There has been one recent paper involving *Tegenaria gigantea* and *T. saeva* by Oxford (1993).

However this was not a study on the population genetics of these species, but an investigation into the genetic control of esterase isozymes.

4.1.2 Aims

This chapter describes an allozyme survey of *T. gigantea* and *T. saeva* populations along a transect across the south coast of England and among morphological classes in the York area. The aim was to find species markers, ideally fixed alleles, that discriminate between the two species in allopatry, and to use these to elucidate any patterns of gene flow and introgression that exist between the two species in parapatry (southern England) and in sympatry (the York area). The problem addressed is a taxonomic one of trying to identify individuals, delineate species boundaries, and explore patterns in 'taxonomic space'. The sampling strategy was not designed to explore aspects of population structure. A limited sample of *T. atrica* (four specimens) was also included in the screening in order to see if there were any differences in electromorphs between this species and *T. saeva* and *T. gigantea* (the small *T. atrica* sample is not included in the analyses that follow).

4.2 Materials and methods

4.2.1 Samples

Adult spiders were killed by freezing at -20°C. Individuals were then placed in a plastic Petri dish on ice and the opithosoma removed with a clean razor blade. Care was taken to leave the epigynal region of females intact and attached to the prosoma for identification and morphometrical analyses as described in Chapter 3. The prosoma was placed into a previously prepared specimen vial containing 95% ethanol for preservation of both specimen and DNA. The removed intact opithosoma was placed in a labelled 1.5 ml microcentrifuge tube in ice and transferred to a -80°C freezer as soon as possible for storage.

In total 239 animals (83 males and 156 females) from Southern England and 80 animals (all males) from the York area were analysed. These specimens had been previously identified on the basis of morphological criteria.

Samples from southern England were partitioned according to the six sampling zones previously described in Chapter 2; thus forming a transect from west to east. Zone 1 and 2 consisted of allopatric *T. saeva*, with zone 1 being the most westerly. Zones 3 and 4 consisted of parapatric *T. saeva* and *T. gigantea* from the contact area in Dorset (corresponding to sites within the 40 km x 40 km area covered by Ordnance Survey map sheet 195). Zones 5 and 6 consisted of allopatric *T. gigantea*, with zone 6 being the most easterly. (These zones are illustrated again, along with the distribution of the sample sites in Southern England, in section 4.3.1, Figure 4.1.)

Samples from the York area were placed into the five morphologically defined groups as described in Chapter 2: *T. saeva* ('good' *T. saeva*), *T. gigantea* ('good' *T. gigantea*), *T. saeva/gigantea?* (appearing intermediate and therefore regarded as putative hybrids), *T. saeva?* (between *T. saeva ?* and *T. saeva/gigantea?*) and *T. gigantea?* (between *T. gigantea* and *T. saeva/gigantea?*).

4.2.2 Homogenization

Samples were homogenized as close as possible to the date of electrophoresis. Tissue homogenates do not retain their enzyme activity as well as frozen whole tissues as a result of the release of vesicle-bound proteases (lysozymes) during homogenization. Enzymes in homogenates are also more susceptible to freeze-thaw damage from ice-crystals than when protected within cellular structures (Richardson *et al.*, 1986). Each frozen opithosoma was weighed in a microcentrifuge tube and cold (4°C) homogenization buffer (see Appendix A.1) was then added, in a ratio of approximately 0.01g tissue: 20 µl buffer, and the tube placed in ice. Each sample was then ground with a clean pellet mixer (Scotlab) and the resulting suspension centrifuged for 20 minutes at 2550*g* to precipitate the cellular material and then returned to ice. The supernatant was pipetted into a clean 0.5 ml microcentrifuge tube, taking care not to remove too much of the cellular debris or fatty material on the surface. The supernatant was then divided into further 10-20 µl aliquots in 0.2 ml microcentrifuge tubes and frozen at -80°C; to obviate the need to freeze and thaw homogenates repeatedly.

In the initial screening for scorable enzyme systems, each opithosoma was sliced in half longitudinally just prior to weighing and each half placed in separate tubes. This permitted the a spider to be homogenized using both the standard homegenization buffer and one containing NADP. Some labile enzymes requiring NADP are stabilized if homogenized in NADP (Richardson *et al.*, 1986; Hebert and Beaton, 1993). However, once repeatedly scorable systems had been isolated this procedure was no longer necessary.

4.2.3 Allozyme electrophoresis

Specific details of the method for each enzyme and the reagents required (for the four polymorphic systems - see section 4.2.4 below) are listed in Appendix

A.1. Procedures approximately follow those outlined in Hebert and Beaton (1993). All operations were performed on ice as much as was possible to prevent enzyme deterioration.

Titan III cellulose acetate plates (94 mm x 76 mm) (Helena Laboratories Co., Beaumont, Texas, USA) were used for electrophoresis, and treated according to the manufacturer's instructions. Up to twelve sample homegenates (7-10 µl) were applied to the gel with a 'Super Z-12' applicator (Helena Laboratories). Samples were always applied along the long side of the plate. Details of the appropriate tank/gel buffer, position of application, number of applications and run time varied with the enzyme (Appendix A.1). Run times and voltages were determined from centrally loaded test-strips (small sections of cellulose acetate plate). These were used to optimise and screen the different enzyme systems. Freezer packs were placed above the chamber to keep the enzymes cool and to minimise evaporation and subsequent condensation of the buffer which could lead to smearing. All the enzymes of interest here migrated anodally, with the exception of *Got-1* and *Alcdh-1* (full enzyme names are given below).

Each enzyme was detected by a specific staining method. Stain solutions appropriate to the enzyme system (Appendix A.1) were prepared a few minutes before required and stored in a light-proof test tube. The stain was applied as an agarose overlay; 4 ml of molten 1.44% (w/v) agarose at 60°C was added to the stain solution immediately before application to the gel. (The staining procedure for amylase differed in that the *plate* was applied to a Petri-dish of solidified 1%/0.25% agar/starch - see Appendix A.1). The gel and overlay were kept in the dark at room temperature (ca. 21°C) in a sealed box with damp tissue to maintain high humidity, until sufficient staining was achieved (Appendix A.1). The overlay was then washed off and the plate was soaked in several changes of water for about 30 minutes to remove any unreacted stain reagents. The gel was scored whilst still wet and then air-dried and stored in the dark for reference.

For many enzyme systems it was possible to load two sets of samples per plate (24 samples). Weakly staining samples were rerun; often from a different application position (e.g. amylase tends to smear unless loaded centrally). 'Line-up' gels were also run to check the scoring of the different alleles from plate to plate (in other words, samples originally run on separate plates were re-run on a single new plate for comparison).

In scoring the gels, a number of assumptions were made about the patterns of enzyme variation recorded: 1) the banding patterns were assumed to be the expression of independent loci; 2) the variation was assumed to be genetic in origin; 3) the variation was described in terms of alleles, which were assumed to be codominant, with products only differing in electrophoretic mobility (see section 4.1.1). The putative alleles were designated alphabetically, with *A* representing the allele migrating furthest in each case.

4.2.4 The enzymes

The enzymes examined in this study were chosen on the basis of 1) the cost and complexity of the stain, 2) previous experience in our laboratory, and 3) previous electrophoretic studies of spiders in which some or all of these enzymes were successfully utilised (e.g. Elliott *et al.*, 1982; Lubin and Crozier, 1985; Smith, 1986; Terranova and Roach, 1987a, 1987b; Porter and Jakob, 1990; Smith and Engel, 1994; Ramirez and Beckwitt, 1995; Ramirez and Fandino, 1996; Gillespie and Oxford, 1998).

Fourteen enzyme systems were initially examined for scorability and variability in 18 individuals of *T. saeva* and 18 *T. gigantea* from zones 1 and 6 respectively. Five enzymes were eliminated at this stage on the grounds of poor scorability or repeatability: alcohol dehydrogenase (E.C. 1.1.1.1); glucose-6-phosphate dehydrogenase (E.C. 1.1.1.49); malate dehydrogenase (E.C. 1.1.1.37); mannose-6-phosphate isomerase (E.C. 5.3.1.8); and peptidase-B

(E.C. 3.4.11/13). However, it seems likely that further development of electrophoresis and/or staining conditions might render them usable systems as all appeared polymorphic. A further five enzymes and seven loci were screened for an additional 12 *T. saeva* (six from zone 1, six from zone 3; giving a total of 30), and an additional 12 *T. gigantea* (six from zone 4, six from zone 6; giving a total of 30). These enzymes were as follows: glycerol-3-phosphate dehydrogenase (E.C. 1.1.1.8); hexokinase (E.C. 2.7.1.1); glutamic-oxaloacetic transaminase (E.C. 2.6.1.1); isocitrate dehydrogenase (E.C. 1.1.1.42) (two loci: *Idh-1*, *Idh-2*); and lactate dehydrogenase (E.C. 1.1.1.27) (two loci: *Ldh-1*, *Ldh-2*). However, each of these systems proved to be identically monomorphic in both species. The remaining four polymorphic enzymes (four loci) were examined in 119 *T. saeva* (30 from zone 1, 30 from zone 2, and 59 from zone 3), and 120 *T. gigantea* (30 from zone 6, 30 from zone 5, and 60 from zone 4). These four systems were also examined in 80 specimens from the York area (20 *T. saeva*, 12 *T. saeva?*, 20 *T. saeva/gigantea?*, 8 *T. gigantea?* and 20 *T. gigantea*). In addition four *T. atrica* from Nancy, France were examined. The four polymorphic systems were as follows: aconitase (E.C. 4.2.1.3); amylase (E.C. 3.2.1.1); glucose-6-phosphate isomerase (E.C. 5.3.1.9); and phosphoglucomutase (E.C. 5.4.2.2). These yielded the following loci: *Aco-1*; *Amy-1*; *Gpi-1*; and *Pgm-1*.

4.2.5 Analyses

As emphasized above, sampling was not executed in such a way as to allow analyses of population structure. Therefore estimates of Wright's (1951) F_{ST} and $N_e m$ (gene flow) have not been made, although patterns of gene flow and genetic exchange are illustrated by other means. This study provides a **taxonomic** baseline from which more detailed population-based studies can be developed in the future. The analyses of the data are divided into three sections:

1) The allele frequencies and their patterns of geographical distribution in southern England and among morphotypes in the York area are examined.

2) The samples are analysed in terms of genotypes following a taxonomic approach originally devised for use in species diagnosis. Ayala (1983) quotes Mayr (1969) that "a taxonomic character is 'any attribute of a member of a taxon by which it differs or *may* differ from a member of a different taxon'" (my emphasis). A diagnostic character (state) is one that uniquely specifies a given taxon but, as Ayala (1983) goes on to emphasize, taxonomic characters are not just attributes of individuals but of populations and that it is differences between populations, i.e. frequency distributions, that are taxonomically significant. Enzyme loci are often polymorphic and should therefore provide good taxonomic characters, however it is frequently observed that closely related species have some alleles in common at every locus studied. This unsurprising observation initially led to the rejection of allozymes as taxonomic characters due to an emphasis on allelic frequencies (Ayala, 1983). It is the frequency of genotypes, and not of alleles, that is taxonomically significant (Ayala, 1983). This is particularly important as although one may be interested in population differences it is ultimately individuals that are to be classified.

Ayala (1983) presents a method for calculating the probability of a correct diagnosis (identification) of an individual when using allozyme data as a taxonomic character. The technique is developed here to give a tangible measure of the genotypic similarity between the different sample groups. Measures of genotypic overlap and the probability of a correct diagnosis are calculated for all possible pairwise combinations of sample groups. The procedure is as follows.

The expected genotypic frequencies at each locus in each sample are calculated from the allele frequencies using the Hardy-Weinberg polynomial expansion. So, immediately the assumption has been made that the sample represents an outcrossing *population* that is in Hardy-Weinberg equilibrium and clearly this is

not true for the data presented here. This is unlikely to be a major source of error in this context and is considered better than using the observed genotypic frequencies for the following reason. Probable error in estimation decreases as sample size increases. The number of alleles at a polymorphic locus is necessarily smaller than the number of possible genotypes and so 'sample sizes' are greater (sample sizes are also larger because there are two alleles for each genotype). Therefore the observed allele frequencies will always estimate the population allele frequencies with a smaller probable error than the observed genotypic frequencies estimate the actual genotypic frequencies (Ayala, 1983). (In fact very little difference is found between the present results if parallel calculations are made using both observed and expected genotypic frequencies (*Pers. Obs.*)).

To illustrate the method of calculation consider the grouped samples from sample zones 1 and 2, and from zones 5 and 6 (see section 4.3.1); the expected genotypic frequencies at the *Gpi-1* locus are:

	AA	***AB***	***BB***
1 & 2	0.9834	0.0165	6.944×10^{-5}
5 & 6	0.4444	0.4444	0.1111

The overlap of the genotypic distributions for this locus is simply the sum of the smallest frequencies for each genotype: $0.4444 + 0.0165 + 6.944 \times 10^{-5} = 0.4610$. The genotypic overlap is then calculated for each of the other loci and the total genotypic overlap (S_g) is the product of these values for each locus, which for 1&2/5&6 = 0.0480 or 4.8 %. S_g can be thought of as an index with range zero to one (genotypic identity or 100 % overlap).

The theoretical criterion for assigning an individual to one or other of the groups is simply to assign the individual to the group in which its genotype is found with the highest frequency. The assumption is made that both groups are equally likely (or common) in the sample, this gives for *Gpi-1* in the above

example a probability of a wrong diagnosis (P_W) of 0.4610/2 = 0.2305, which is clearly quite high. If this value is calculated for all loci and the product taken, then for 1&2/5&6, P_W = 0.0030. The probability of a correct diagnosis P_C using all four loci is therefore 0.9970 or 99.7 %, which is extremely high.

3). The third part of the analysis employed estimates of genetic distance to generate 'population' trees in order to describe further the relationships between samples. The genetic distances were estimated in two ways: 1) using conventional techniques based on allele frequencies, and 2) by converting pairwise estimates of genotypic overlap into a novel measure of genetic distance. The approaches to generating the distance matrices and trees are introduced as appropriate in section 4.3.3.

4.3 Results

4.3.1 Allele frequencies, and distributions.

The four variable systems exhibited the following characteristics: *Aco-1* exhibited three putative alleles of which *C* was the most frequent. Allele *B* was only found in the *T. atrica* samples. *Amy-1* exhibited three alleles with allele *B* being more frequent and allele *C* relatively rare. *Gpi-1* exhibited only two alleles of which allele *A* was the most frequent. *Pgm-1* exhibited four alleles of which allele *B* was the most frequent and allele *D* was relatively rare. Allele *C* was only found in one individual of *T. gigantea* from Gosport, Hampshire, as a homozygote. This could represent an imported specimen from Europe where allele *C* may occur or an unusual sub-population. A proportion of individuals showed weak or no activity for *Aco-1* and despite four or five re-runs (with increased substrate, increased numbers of sample applications, or change of application position) failed to produce scorable bands. The specimens concerned mostly originated from zones 2 and 5 and had been stored at -80°C longer than many of the other specimens which were collected during later surveys. It seems possible that *Aco-1* is more sensitive to storage and treatment conditions than the other enzymes (which showed no reduction in activity). The existence of null alleles (non-functional enzymes or non-transcribable genes) is not likely to offer an explanation because null allele homozygotes for enzymes of intermediary metabolism are usually lethal (Richardson *et al.*, 1986).

Allele frequencies were calculated from the observed genotypes by direct count. These are illustrated as pie-charts, above the maps showing the sample distributions and zones, for southern England in Figure 4.1 and the sample distributions for the York area in Figure 4.3. Figures 4.2 and 4.4 illustrate, in more detail, the sample distributions in the York City area and in the contact zone in southern England. The allele frequency pie-charts for southern England directly correspond to the geographically distributed sample zones and those for the York area correspond to the morphological categories. The actual allele frequencies are given in Tables 4.1 and 4.2. The appearance of the allele electromorphs for these

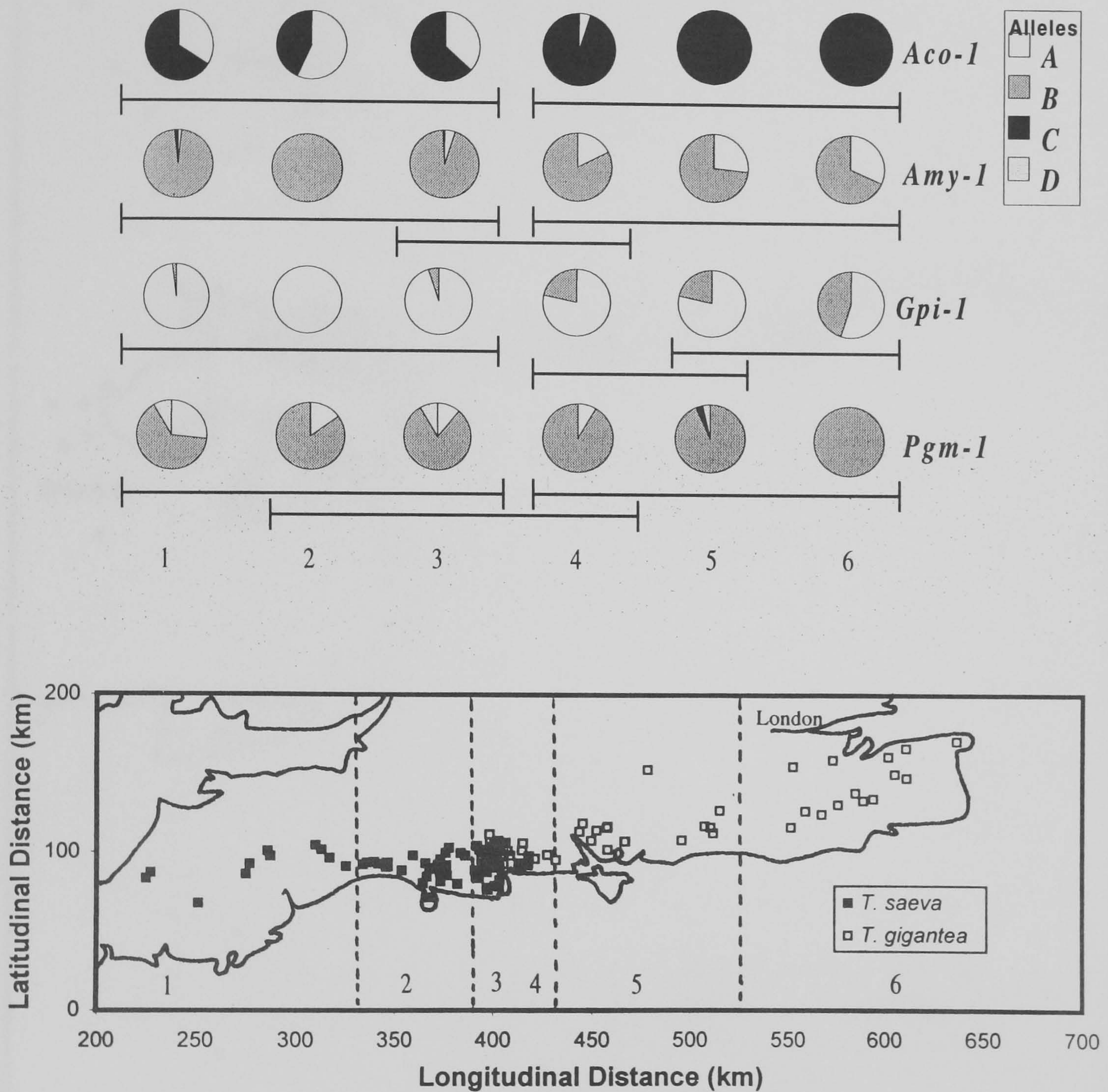


Figure 4.1. Southern England sample sites, sample zones, and allele frequencies. The sample sites for individuals used in the electrophoresis, along with the sample zones (see Chapter 2) are shown on the map (lower half); the corresponding allele frequencies of the four polymorphic loci in each sample zone are illustrated above the map. The bars below the pie charts show homogeneous groupings from the G_H test (see text). Sample zones 1, 2 and 3 consist of *T. saeva* and sample zones 4, 5 and 6 consist of *T. gigantea*. Sample zones 3 and 4 represent the contact area (shown in more detail in Figure 4.2).

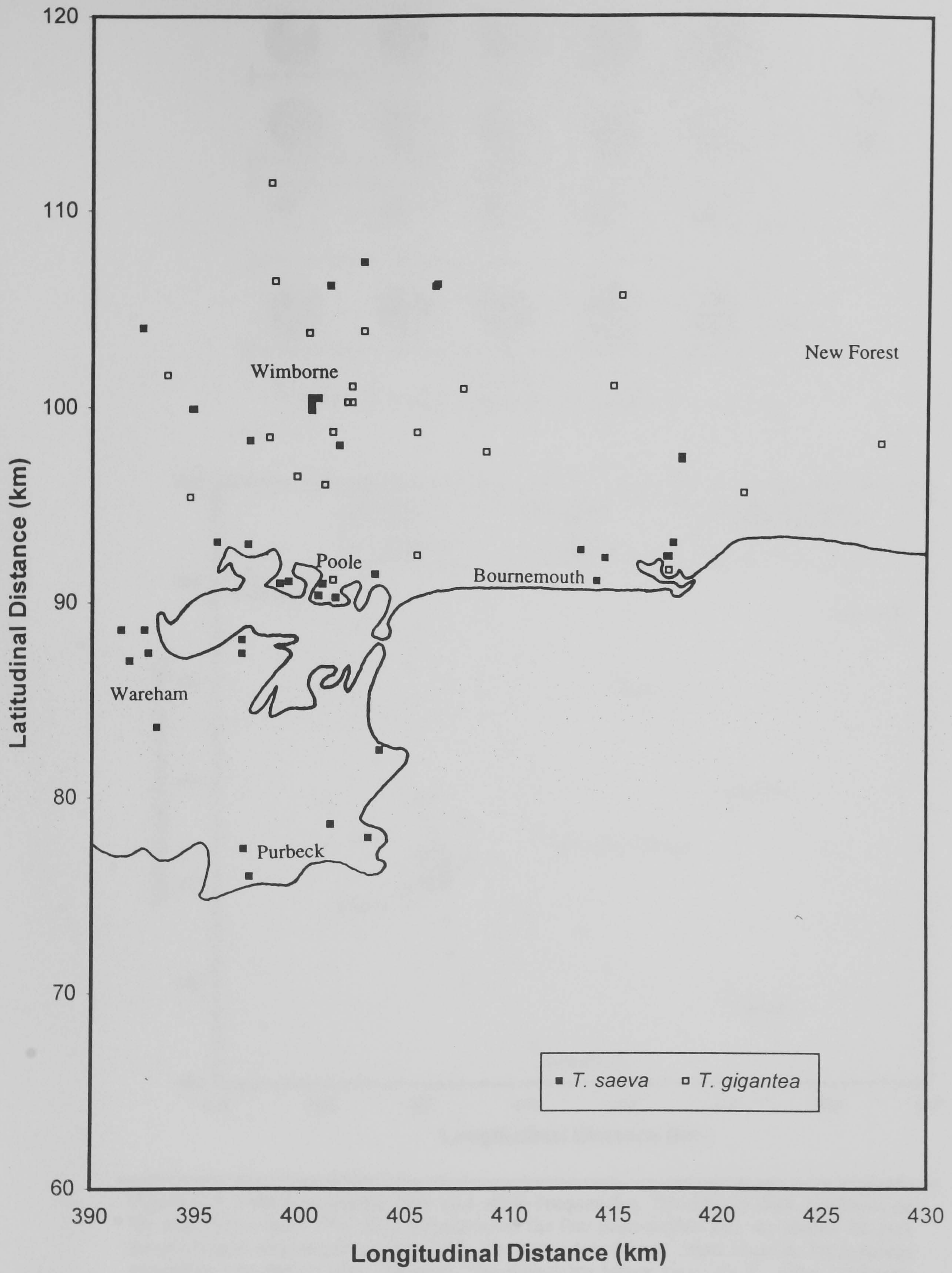
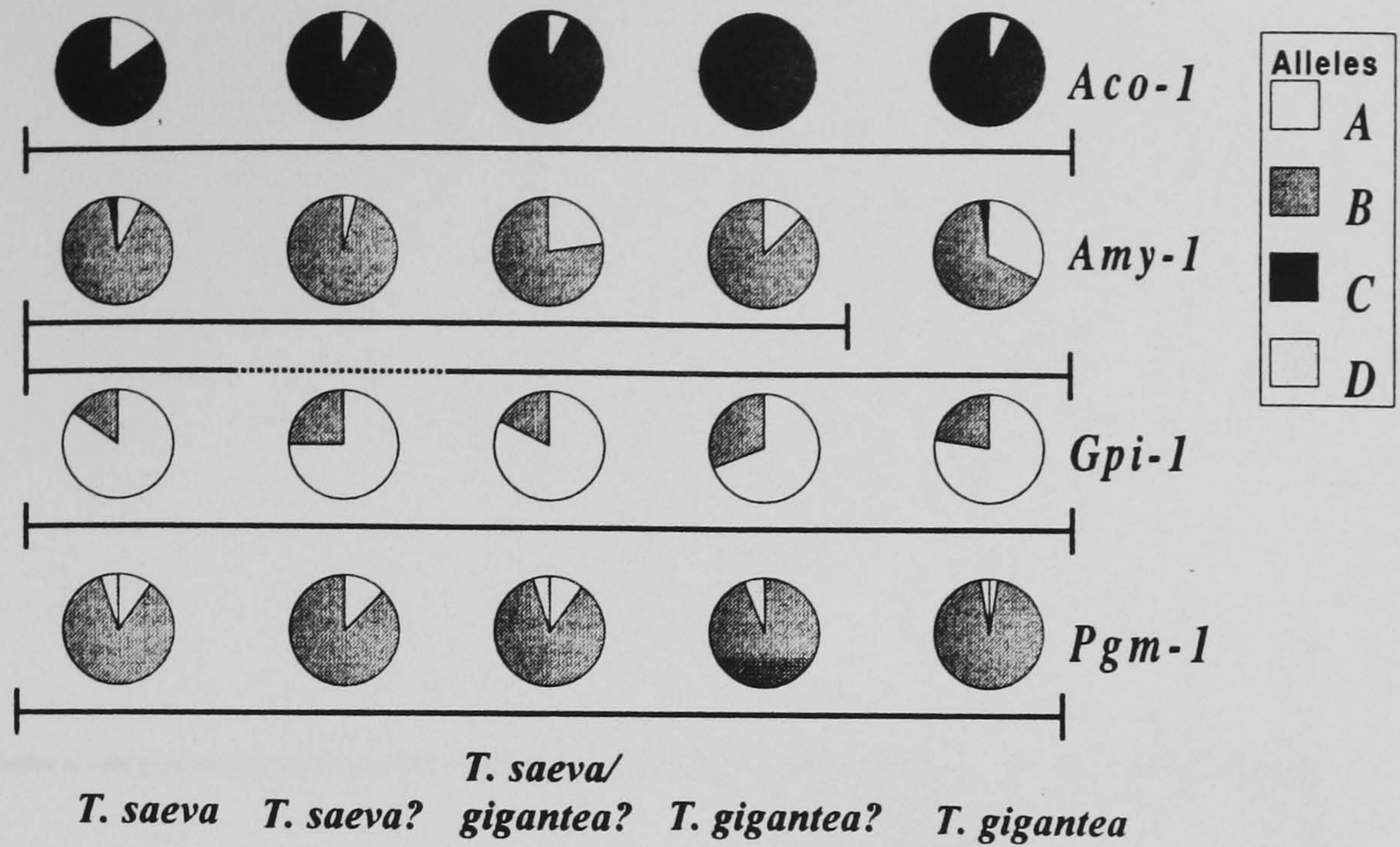


Figure 4.2. Southern England sample sites in the contact area. The sample sites for zones 3 and 4 are shown in greater detail.



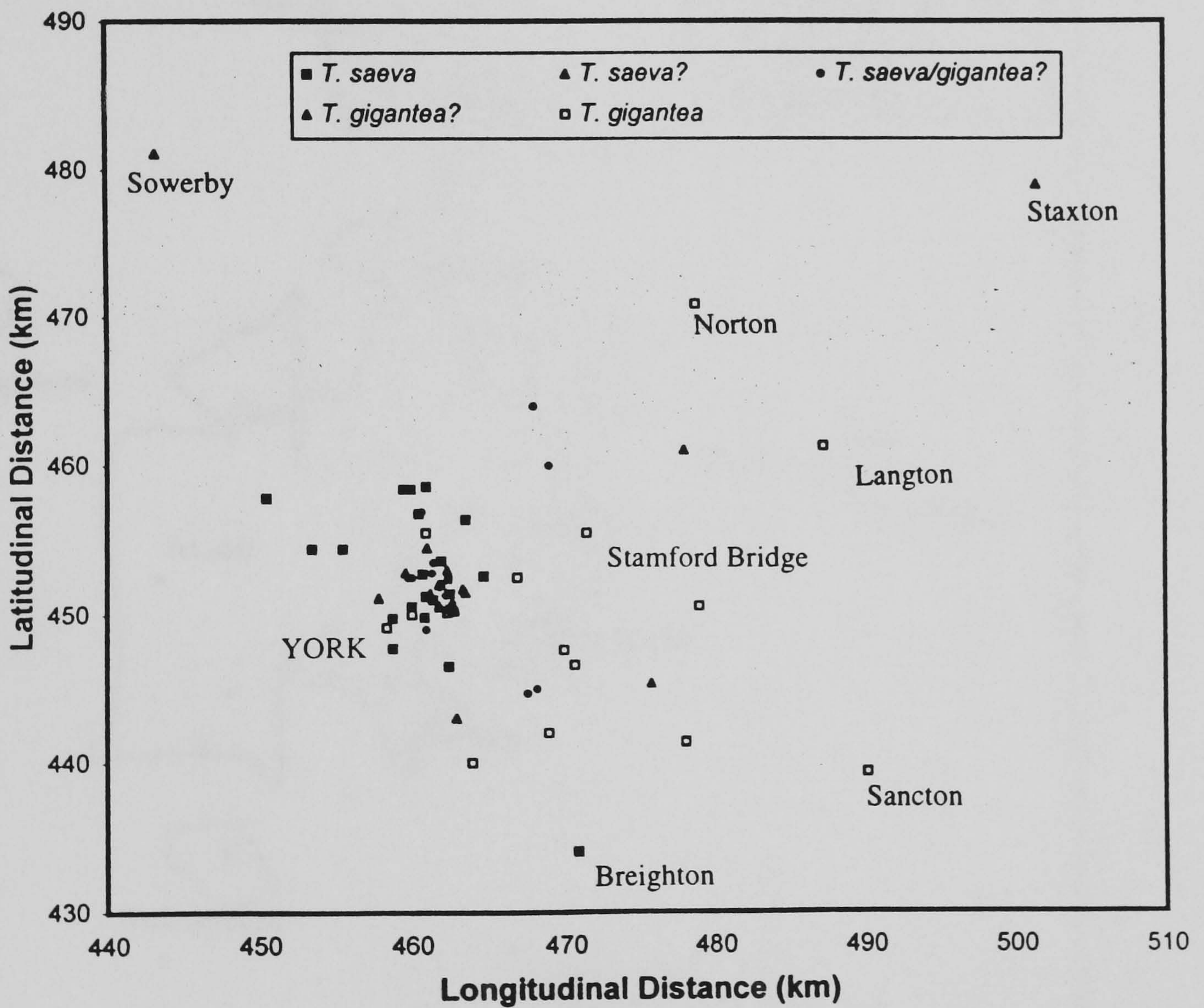


Figure 4.3. York area sample sites and allele frequencies. The sample sites are shown on the map (lower half). The allele frequencies of the four polymorphic loci are shown for each morphological sample group (upper half). The bars below the pie charts illustrate homgeneous groupings from the G_H tests (see text). The broken bar below *Amy-1* for *T. saeva?* indicates that all other sample groups were homogeneous for this locus except *T. saeva?*.The area around York is shown in more detail in Figure 4.4.

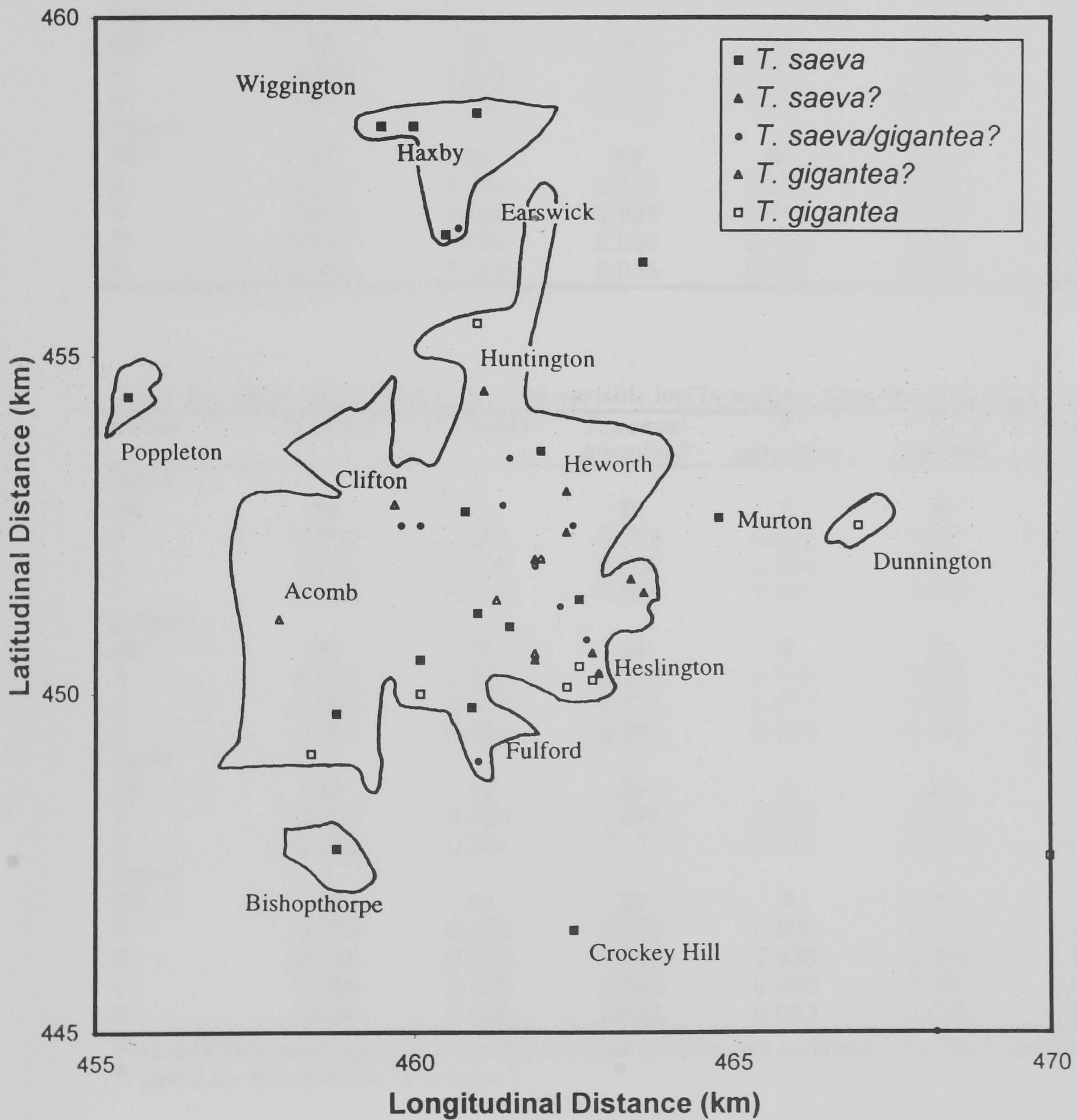


Figure 4.4. York area sample sites (close-up on York). The sample sites for the area in and around York City are shown in more detail.

Table 4.1. Allele frequencies at the four variable loci in the sample zones from Southern England.

	T. saeva			*T. gigantea*		
Locus	**Zone 1**	**Zone 2**	**Zone 3**	**Zone 4**	**Zone 5**	**Zone 6**
Aco-1						
(N)	**28**	**15**	**57**	**58**	**16**	**30**
A	0.339	0.567	0.368	0.052	0.000	0.000
B	0.000	0.000	0.000	0.000	0.000	0.000
C	0.661	0.433	0.632	0.948	1.000	1.000
Amy-1						
(N)	**30**	**30**	**59**	**60**	**30**	**30**
A	0.017	0.000	0.051	0.175	0.267	0.317
B	0.967	1.000	0.941	0.825	0.733	0.683
C	0.017	0.000	0.008	0.000	0.000	0.000
Gpi-1						
(N)	**30**	**30**	**59**	**60**	**30**	**30**
A	0.983	1.000	0.949	0.783	0.783	0.550
B	0.017	0.000	0.051	0.217	0.217	0.450
Pgm-1						
(N)	**30**	**30**	**59**	**60**	**30**	**30**
A	0.267	0.150	0.110	0.083	0.000	0.000
B	0.650	0.850	0.805	0.917	0.933	1.000
C	0.000	0.000	0.000	0.000	0.033	0.000
D	0.083	0.000	0.085	0.000	0.033	0.000

Table 4.2. Allele frequencies at the four variable loci in samples from the York area.

Locus	*T. saeva*	*T. saeva?*	*T. saeva/ gigantea?*	*T. gigantea?*	*T. gigantea*	*T. atrica*
Aco-1						
(N)	**20**	**12**	**20**	**8**	**20**	**4**
A	0.150	0.083	0.075	0.000	0.075	0.500
B	0.000	0.000	0.000	0.000	0.000	0.500
C	0.850	0.917	0.925	1.000	0.925	0.000
Amy-1						
(N)	**20**	**12**	**20**	**8**	**20**	**4**
A	0.075	0.042	0.225	0.125	0.325	0.750
B	0.900	0.958	0.775	0.875	0.650	0.250
C	0.025	0.000	0.000	0.000	0.025	0.000
Gpi-1						
(N)	**20**	**12**	**20**	**8**	**20**	**4**
A	0.850	0.750	0.825	0.688	0.775	1.000
B	0.150	0.250	0.175	0.313	0.225	0.000
Pgm-1						
(N)	**20**	**12**	**20**	**8**	**20**	**4**
A	0.100	0.125	0.100	0.000	0.025	0.000
B	0.850	0.875	0.850	0.938	0.950	1.000
C	0.000	0.000	0.000	0.000	0.000	0.000
D	0.050	0.000	0.050	0.063	0.025	0.000

York area individuals grouped according to visual identification (columns 2-6). The 4 specimens of *T. atrica* are also included in column 7.

four systems is illustrated stylistically in Figure 4.5. (The full list of specimens, including their locations of origin and genotypes, has been retained by the author and by G. S. Oxford, University of York).

The bars below the pie-charts in Figure 4.1 and Figure 4.3 represent homogeneous groupings derived from log-likelihood ratio (G_H) tests of heterogeneity between sample zones/morphological categories; samples contributing to a significant heterogeneity being identified with a simultaneous test procedure, again using G (Oxford, 1991; Sokal and Rohlf, 1995). Critical values of G_H were read from a χ^2 table at n-1 degrees of freedom (where n is the number of sample groups). This test requires two series of paired alleles, ordered according to the most common allele; therefore the most common allele was taken as one series and all other alleles were pooled to form the other series. Because this test requires the calculation of *ln*, zeros (which occur when only one allele is present in a sample) are not tolerated; some sample sets therefore had to be pooled. The acceptability of this was tested using a Fisher's Exact Test (2-tailed) or χ^2 (if pooling more than two samples). The simultaneous test procedure is better than calculating numerous pairwise χ^2 values which could lead to an excess of type I errors. See Table 4.3, below.

Examination of Figure 4.1 immediately reveals that there is a geographic pattern to the data from southern England. Firstly it is very clear that *T. saeva* and *T. gigantea* separate in terms of allele frequencies and this is supported by the G_H statistics. Allopatric populations (zones 1 and 2, and 5 and 6) are always significantly different from one another at all four loci. *Aco-1 A* and *Pgm-1 A* are largely restricted to *T. saeva* (zones 1, 2, and 3) and *Amy-1 A* and *Gpi-1 B* are largely restricted to *T. gigantea* (zones 4, 5, and 6). Secondly it is clear that there is a small amount of 'overspill' by these alleles into the other species at the contact zone. For example *Aco-1 A* is frequent in *T. saeva* in zones 1, 2, and 3 and absent in *T. gigantea* from zones 5 and 6, but present at low frequency in *T. gigantea* from zone 4 (the contact zone). A similar pattern can be seen for the other three loci and is highly suggestive of some hybridization and subsequent introgression in

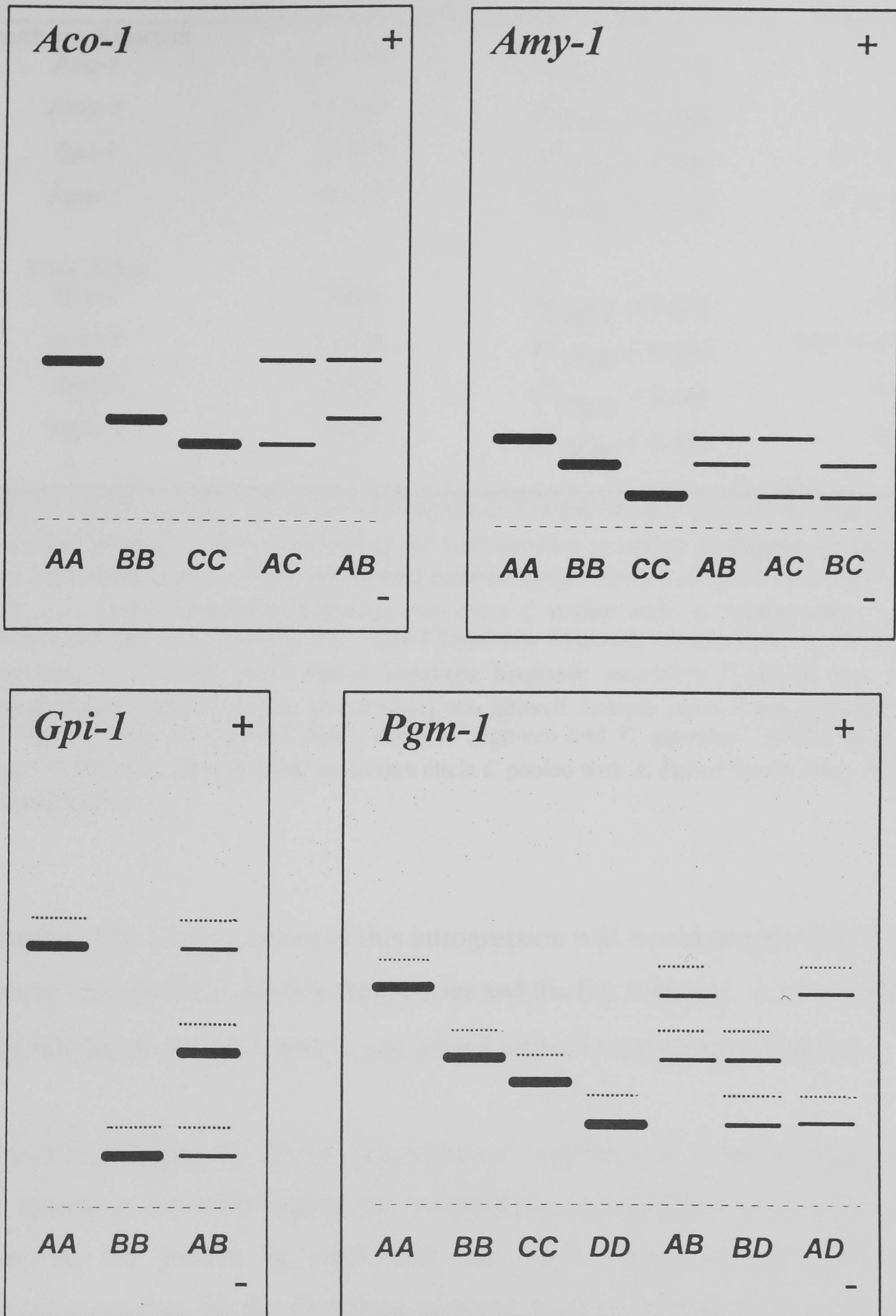


Figure 4.5. Electromorph patterns for the four polymorphic loci.
A diagrammatic representation of the electromorphs. The letters correspond to the genotypes and the broken line above the letters is the origin of application. The broken lines behind the main bands in *Gpi-1* and *Pgm-1* represent the positions of sub-bands which are occasionally observed.

Table 4.3. Log-likelihood ratio tests of heterogeneity for the four variable loci.

Locus	G_H	$G_{critical}$ P=0.05	Overall Significance
Southern England			
Aco-1	96.776	$\chi^2_{.05(3)} = 7.815$	$P << 0.001$
Amy-1	51.180	$\chi^2_{.05(4)} = 9.488$	$P << 0.001$
Gpi-1	79.301	$\chi^2_{.05(4)} = 9.488$	$P << 0.001$
Pgm-1	46.158	$\chi^2_{.05(4)} = 9.488$	$P << 0.001$
York Area			
Aco-1	2.656	$\chi^2_{.05(3)} = 7.815$	n.s.
Amy-1	13.599	$\chi^2_{.05(4)} = 9.488$	$0.01 > P > 0.001$
Gpi-1	2.428	$\chi^2_{.05(4)} = 9.488$	n.s.
Pgm-1	3.466	$\chi^2_{.05(4)} = 9.488$	n.s.

G_H is the overall G_H value for the data; the significance of this value is given in the final column. The critical value of G used in delimiting the homogeneous groupings in Figures 4.1 and 4.3 is given in the third column. ***Aco-1*** **Southern England:** sample zones 4, 5, and 6 pooled ($\chi^2_{.05(2)}$ = 4.900, n.s.). ***Amy-1*** **Southern England:** rare allele *C* pooled with *A*. Sample zones 1 and 2 pooled (Exact test: $P_{1/2}$ = 0.496, n.s.). ***Gpi-1*** **Southern England:** sample zones 1 and 2 pooled (Exact test: $P_{1/2}$ = 1.000, n.s.). ***Pgm-1*** **Southern England:** rare allele *D* pooled with *A*. The unusual allele *C* (present in one homozygote) was ignored. Sample zones 5 and 6 pooled (Exact test: $P_{5/6}$ = 0.239, n.s.). ***Aco-1*** **York area:** *T. gigantea* and *T. gigantea?* pooled (Exact test: $P_{G/G?}$ = 0.550, n.s.). ***Amy-1*** **York area:** rare allele *C* pooled with *A*. ***Pgm-1*** **York area:** rare allele *D* pooled with *A*.

parapatry. The relative extent of this introgression will be examined below. On the basis of these patterns in allele frequencies and the G_H statistics, it was considered legitimate to pool zone 1 with 2, and zone 5 with 6 in subsequent analyses.

Inspection of Figure 4.3 shows suggestions of a gradient in allele frequencies from one species to the other (across the morphological categories from the York area). However, the pattern is weak and the allele frequencies are much more homogeneous than for the data from southern England. Indeed the G_H tests failed to find any heterogeneity with the exception of *Amy-1* which gave two groupings - one excluding *T. gigantea* and one excluding *T. saeva?*. The significance of this result is low when compared to the values obtained with the data from southern England and may in part be due to the relatively small sample sizes of *T. saeva?* and *T. gigantea?* available for electrophoresis. However when all four loci are

considered together the data are clearly quite homogeneous. These results suggest that there has been much hybridization and introgression in the York area when compared to Southern England. In comparing Figure 4.1 and Figure 4.3 it is interesting to note that the York *T. gigantea* appears to be almost identical to the allopatric *T. gigantea* from zones 5 and 6, but the York *T. saeva* is more like *T. gigantea* at all loci than the allopatric *T. saeva* from zones 1 and 2.

4.3.2 Genotypic Overlap

The observed and expected genotypic frequencies for Southern England are shown in Table 4.4, and for the York area in Table 4.5. The expected genotype frequencies were used in the calculation of genotypic overlap (S_g) as described previously. The pairwise values of S_g and P_c ('probability of a correct diagnosis') are shown in Table 4.6.

For a particular locus to be diagnostic it has been suggested that a minimum criterion should be that it assigns individuals to the correct group with a 99% probability (Ayala, 1983). Clearly none of the four variable loci alone in this study fits this criterion, but when taken together they provide a quite powerful diagnostic tool. Examination of Table 4.6 reveals those pairwise values of P_c which fit this criterion. The G_H statistics of section 4.3.1 are clearly supported with no significant values separating the York area samples from each other and with the sample zones from Southern England splitting into two groups (zones 1&2 and 3 (*T. saeva*) versus zones 4 and 5&6 (*T. gigantea*)). The P_c value of zone 3 versus zone 4 just fails to meet the 99% criterion. Also, the P_c values of the pairwise comparisons of zones 1&2 and 3 (southern *T. saeva*) versus York *T. gigantea* and York *T. gigantea?* are highly significant, but interestingly the opposite comparison (zones 5&6 and 4 (southern *T. gigantea*) versus York *T. saeva* and *T. saeva?*) markedly fail to meet the criterion. In other words, as suggested from the examination of the allele frequencies described above, *T. saeva* in the York area resembles *T. gigantea* from both York and southern England in terms of genotype,

Table 4.4. Observed and expected genotype frequencies in the sample zones from Southern England.

	T. saeva		*T. gigantea*	
Locus	Zone 1&2	Zone 3	Zone 4	Zone 5&6
Aco-1				
(N)	**43**	**57**	**58**	**46**
AA	0.2093 (0.1752)	0.2281 (0.1357)	0.0000 (0.0027)	0.0000 (0.0000)
AC	0.4186 (0.4867)	0.2807 (0.4654)	0.1034 (0.0981)	0.0000 (0.0000)
CC	0.3721 (0.3380)	0.4912 (0.3989)	0.8966 (0.8992)	1.0000 (1.0000)
Amy-1				
(N)	**60**	**59**	**60**	**60**
AA	0.0000 (0.0001)	0.0169 (0.0026)	0.0333 (0.0306)	0.1000 (0.0851)
AB	0.0167 (0.0164)	0.0678 (0.0957)	0.2833 (0.2888)	0.3833 (0.4132)
AC	0.0000 (0.0001)	0.0000 (0.0009)	0.0000 (0.0000)	0.0000 (0.0000)
BB	0.9667 (0.9669)	0.8983 (0.8849)	0.6833 (0.6806)	0.5167 (0.5017)
BC	0.0167 (0.0164)	0.0169 (0.0159)	0.0000 (0.0000)	0.0000 (0.0000)
CC	0.0000 (0.0001)	0.0000 (0.0001)	0.0000 (0.0000)	0.0000 (0.0000)
Gpi-1				
(N)	**60**	**59**	**60**	**60**
AA	0.9833 (0.9834)	0.9153 (0.9009)	0.6333 (0.6136)	0.5000 (0.4444)
AB	0.0167 (0.0165)	0.0678 (0.0965)	0.3000 (0.3394)	0.3333 (0.4444)
BB	0.0000 (0.0001)	0.0169 (0.0026)	0.0667 (0.0469)	0.1667 (0.1111)
Pgm-1				
(N)	**60**	**59**	**60**	**60**
AA	0.0667 (0.0434)	0.0000 (0.0121)	0.0333 (0.0069)	0.0000 (0.0000)
AB	0.2667 (0.3125)	0.2034 (0.1774)	0.1000 (0.1528)	0.0000 (0.0000)
AD	0.0167 (0.0174)	0.0169 (0.0187)	0.0000 (0.0000)	0.0000 (0.0000)
BB	0.5833 (0.5625)	0.6610 (0.6482)	0.8667 (0.8403)	0.9500 (0.9344)
BC	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0322)
BD	0.0667 (0.0625)	0.0847 (0.1365)	0.0000 (0.0000)	0.0333 (0.0322)
CC	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0167 (0.0003)
CD	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0006)
DD	0.0000 (0.0017)	0.0339 (0.0072)	0.0000 (0.0000)	0.0000 (0.0003)

Expected frequencies given in parentheses. Values are given to 4 decimal places.

Table 4.5. Observed and expected genotype frequencies for samples from the York area.

Locus	*T. saeva*	*T. saeva?*	*T. saeva/ gigantea?*	*T. gigantea?*	*T. gigantea*
Aco-1					
(N)	**20**	**12**	**20**	**8**	**20**
AA	0.0000 (0.0225)	0.0000 (0.0069)	0.0500 (0.0056)	0.0000 (0.0000)	0.0500 (0.0056)
AC	0.3000 (0.2550)	0.1667 (0.1528)	0.0500 (0.1388)	0.0000 (0.0000)	0.0500 (0.1388)
CC	0.7000 (0.7225)	0.8333 (0.8403)	0.9000 (0.8556)	1.0000 (1.0000)	0.9000 (0.8556)
Amy-1					
(N)	**20**	**12**	**20**	**8**	**20**
AA	0.0000 (0.0056)	0.0000 (0.0017)	0.0500 (0.0506)	0.000 (0.0156)	0.1500 (0.1056)
AB	0.1500 (0.1350)	0.0833 (0.0799)	0.3500 (0.3488)	0.250 (0.2188)	0.3500 (0.4225)
AC	0.0000 (0.0038)	0.0000 (0.0000)	0.0000 (0.0000)	0.000 (0.0000)	0.0000 (0.0163)
BB	0.8000 (0.8100)	0.9167 (0.9184)	0.6000 (0.6006)	0.750 (0.7656)	0.4500 (0.4225)
BC	0.0500 (0.0450)	0.0000 (0.0000)	0.0000 (0.0000)	0.000 (0.0000)	0.0500 (0.0325)
CC	0.0000 (0.0006)	0.0000 (0.0000)	0.0000 (0.0000)	0.000 (0.0000)	0.0000 (0.0006)
Gpi-1					
(N)	**20**	**12**	**20**	**8**	**20**
AA	0.7000 (0.7225)	0.5833 (0.5625)	0.7000 (0.6806)	0.5000 (0.4727)	0.5500 (0.6006)
AB	0.3000 (0.2550)	0.3333 (0.3750)	0.2500 (0.2888)	0.3750 (0.4297)	0.4500 (0.3488)
BB	0.0000 (0.0225)	0.0833 (0.0625)	0.0500 (0.0306)	0.1250 (0.0977)	0.0000 (0.0506)
Pgm-1					
(N)	**20**	**12**	**20**	**8**	**20**
AA	0.0000 (0.0100)	0.0000 (0.0156)	0.0000 (0.0100)	0.0000 (0.0000)	0.0000 (0.0006)
AB	0.2000 (0.1700)	0.2500 (0.2188)	0.1500 (0.1700)	0.0000 (0.0000)	0.0500 (0.0475)
AD	0.0000 (0.0100)	0.0000 (0.0000)	0.0500 (0.0100)	0.0000 (0.0000)	0.0000 (0.0013)
BB	0.7000 (0.7225)	0.7500 (0.7656)	0.7500 (0.7225)	0.8750 (0.8789)	0.9000 (0.9025)
BC	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
BD	0.1000 (0.0850)	0.0000 (0.0000)	0.0500 (0.0850)	0.1250 (0.1172)	0.0500 (0.0475)
CC	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
DD	0.0000 (0.0025)	0.0000 (0.0000)	0.0000 (0.0025)	0.0000 (0.0006)	0.0000 (0.0006)

Expected frequencies given in parentheses. Values are given to 4 decimal places.

Table 4.6 Pairwise measures of genotypic overlap (S_g) and probability of a correct diagnosis (P_c) between any two samples.

	T.saeva		*T. gigantea*						
	Zone 1&2	***Zone 3***	***Zone 4***	***Zone 5&6***	***T.saeva***	***T. saeva?***	***T.saeva/ gigantea ?***	***T. gigantea ?***	***T. gigantea***
Zone 1&2	-	0.9588	0.9913*	0.9970*	0.9814	0.9866	0.9894	0.9936*	0.9944*
Zone 3	0.6590	-	0.9860	0.9945*	0.9703	0.9813	0.9829	0.9905*	0.9905*
Zone 4	0.1392	0.2241	-	0.9678	0.9668	0.9606	0.9547	0.9629	0.9611
Zone 5&6	0.0480	0.0885	0.5155	-	0.9842	0.9793	0.9722	0.9593	0.9612
T.saeva	0.2980	0.4759	0.5315	0.2530	-	0.9627	0.9615	0.9751	0.9766
T. saeva?	0.2147	0.2992	0.6298	0.3308	0.5963	-	0.9534	0.9690	0.9757
T.saeva/ gigantea?	0.1695	0.2740	0.7242	0.4446	0.6157	0.7454	-	0.9714	0.9612
T. gigantea ?	0.1023	0.1521	0.5939	0.6520	0.3979	0.4961	0.4583	-	0.9716
T. gigantea	0.0895	0.1524	0.6222	0.6216	0.3745	0.3885	0.6200	0.4543	-

S_g lower diagonal. P_c upper diagonal. The values of P_c which are boxed and marked * represent pairwise comparisons for which the four variable loci are diagnostic at the 99% confidence level.

whereas *T. saeva* from southern England is genotypically distinct. This suggests that *T. saeva* in the York area has experienced substantial introgression of *T. gigantea* nuclear genes.

What patterns do the values of S_g reveal? The data from Southern England are based on geographical sample zones whereas the data from the York area are based on morphological groupings so a direct comparison is not possible. Estimation of genotypic overlap was originally intended for use in calculating the probability of a wrong diagnosis when using genotype to assign individual specimens to species (or populations) (Ayala, 1983). The use of S_g to describe the relationships between 'populations' (the sample zones from southern England or the morphological groupings from the York area), or gene flow, is a novel approach and demands caution in interpretation. The behaviour of S_g as a statistical function is undefined and is likely to be very complex. Figure 4.6 illustrates the complexity of S_g under the simplest possible scenario: two populations, 1 and 2 with one locus possessing two alleles, *A* and *B*. The figure shows the value of S_g at all possible frequencies of allele *A* in each population. The behaviour of S_g when there is more than one locus and/or multiple alleles requires computer modelling. It is not possible to test observed values of S_g against expectation because no model or expectation exists. S_g can only be used as an indicator of the relative similarity between samples in terms of genotype. Further, it should be stressed that the genetic background of the individuals in each of the samples analysed here is unknown, for example although the York *T. saeva/gigantea?* samples may have appeared intermediate in terms of morphology it is not reasonable to assume that they represent first generation hybrids.

Despite the cautionary notes, above, the relative values of S_g are still informative. Examining the values for southern England (Table 4.6) a pattern that might have been expected from the pie-charts in Figure 4.3 is seen. Zones 1&2 and zones 5&6 have the lowest genotypic overlap (0.0480). Zone 3 and zone 4 also have a low S_g (0.2241) but this is 4.67 times greater than that of zone 1&2 and zone 5&6 suggesting some gene flow between these parapatric zones. The values for zone

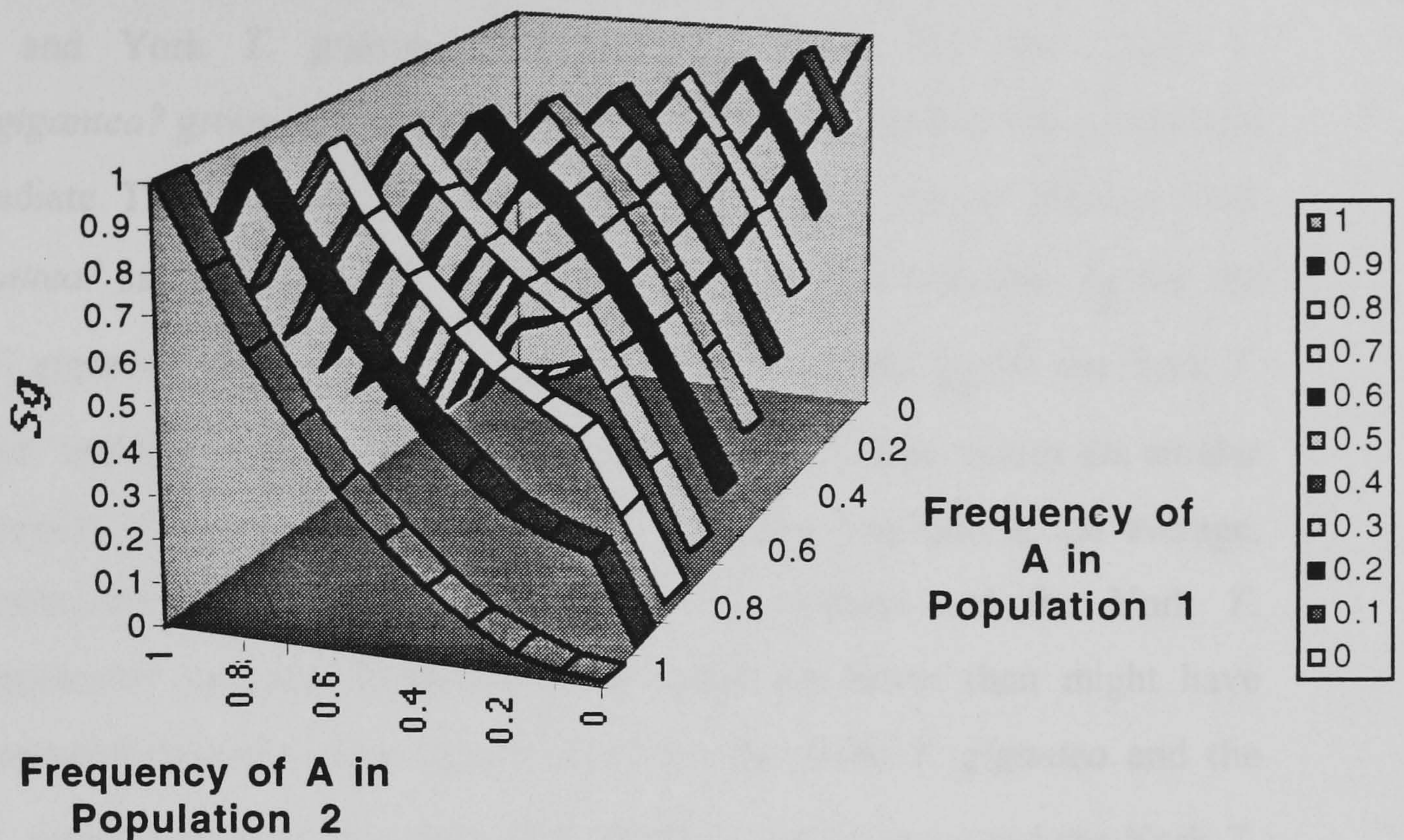


Figure 4.6. The behaviour of the genotypic overlap, S_g, between two theoretical populations, 1 and 2, for a single locus with two alleles, *A* and *B*. Values of S_g are plotted for all possible frequencies of *A* in each population (in 0.1 increments from 0 to 1).

1&2 and zone 3 (0.6590) and for zone 4 and zone 5&6 (0.5155) are, as might be expected, high. These values differ by a factor of 1.28 suggesting that there may be some asymmetry in hybridization or gene flow which makes zone 4 (parapatric *T. gigantea*) appear more distinct from zone 5&6 (allopatric *T. gigantea*) than does zone 3 (parapatric *T. saeva*) from zone 1&2 (allopatric *T. saeva*). A better idea of this possible asymmetry can be achieved by comparing the S_g values for zone 4 and zone 1&2 (0.1392) and zone 3 and zone 5&6 (0.0885) from which it is clear that parapatric *T. gigantea* resemble allopatric *T. saeva* 1.57 times more than parapatric *T. saeva* resemble allopatric *T. gigantea*.

Examining the York values of S_g it can be seen that York *T. saeva* and the York *T. gigantea* samples have the lowest value (0.3745) as expected. S_g for the York *T. gigantea* and the York *T. saeva/T. gigantea?* = 0.6200, and S_g for the

York *T. saeva* and the York *T. saeva/gigantea?* = 0.6157. These values are very similar and larger in magnitude than the values of S_g between the York *T. saeva* and York *T. gigantea* samples, suggesting that the York *T. saeva/gigantea?* group is not only morphologically intermediate but genetically intermediate. The values of S_g associated with the York *T. saeva?* and the York *T. gigantea?* samples are a little harder to interpret. For example, S_g for the York *T. gigantea* and the York *T. gigantea?* = 0.4543, and S_g for the York *T. gigantea?* and the York *T. saeva/gigantea?* = 0.4583. These values are similar and therefore may suggest that the York *T. gigantea?* sample is, on average, genetically intermediate between the York *T. gigantea* and the York *T. saeva/gigantea?* samples. However, these values are lower than might have been expected from the value S_g of 0.6200 for the York *T. gigantea* and the York *T. saeva/gigantea?* The value of S_g for the York *T. saeva* and the York *T. saeva?* = 0.5963 and S_g for the York *T. saeva?* and the York *T. saeva/gigantea?* = 0.7454, suggesting the York *T. saeva?* is genetically more similar to the York *T. saeva/gigantea?* sample than it is to the York *T. saeva*. Once more it is important to consider the lack of knowledge about the genealogy of the individuals in the sample groups, the uncertainty about the behaviour of S_g and the small sample sizes for the York *T. saeva?* and York *T. gigantea?*.

The S_g value for York *T. gigantea* and zone 5&6 (allopatric southern *T. gigantea*) is large as expected (0.6216) but S_g for York *T. saeva* and zone 1&2 (allopatric southern *T. saeva*) is much smaller than expected (0.2147). In fact the York *T. saeva* has a larger value of S_g with zone 5&6 (allopatric southern *T. gigantea*) (0.2530) than it does with zone 1&2. This value is 2.83 times as large as the value for York *T. gigantea* and zone 1&2 (allopatric southern *T. saeva*) (0.0895). These again suggest a genetic shift in the York *T. saeva* towards *T. gigantea*, but with no such shift in *T. gigantea*.

Figure 4.7 has independent two curves. The lower curve illustrates the relationships between the S_g values for the sample zones from southern England with zone 1&2 and zone 5&6. It can be seen that the points fall onto a smooth hyperbola and that zone 3 is more similar to zone 1&2 (S_g = 0.6590) than zone 4 is to zone 5&6 (S_g = 0.5155) (suggesting that the parapatric *T. gigantea* of zone 4 contain more *T. saeva* genes than the parapatric *T. saeva* of zone 3 contain *T. gigantea* genes). The curve for the York area samples (S_g values with 'good' *T. saeva* and 'good' *T. gigantea*) is shifted to the right reflecting the greater amount of overall overlap in this data set (for example, for the York *T. gigantea* and the York *T. saeva* S_g = 0.3745, whereas for zone 1&2/zone 5&6 S_g = 0.0480). The discontinuity between the morphological groupings from the York area compared to the continuous geographic groupings from the south is clearly indicated by the inflexion in the York curve. The *T. saeva/gigantea?* point lies halfway between the *T. saeva* and *T. gigantea* points indicating the intermediacy of this group. The *T. saeva?* point is closer to the *T. saeva* point than the *T. gigantea?* point is to the *T. gigantea* point.

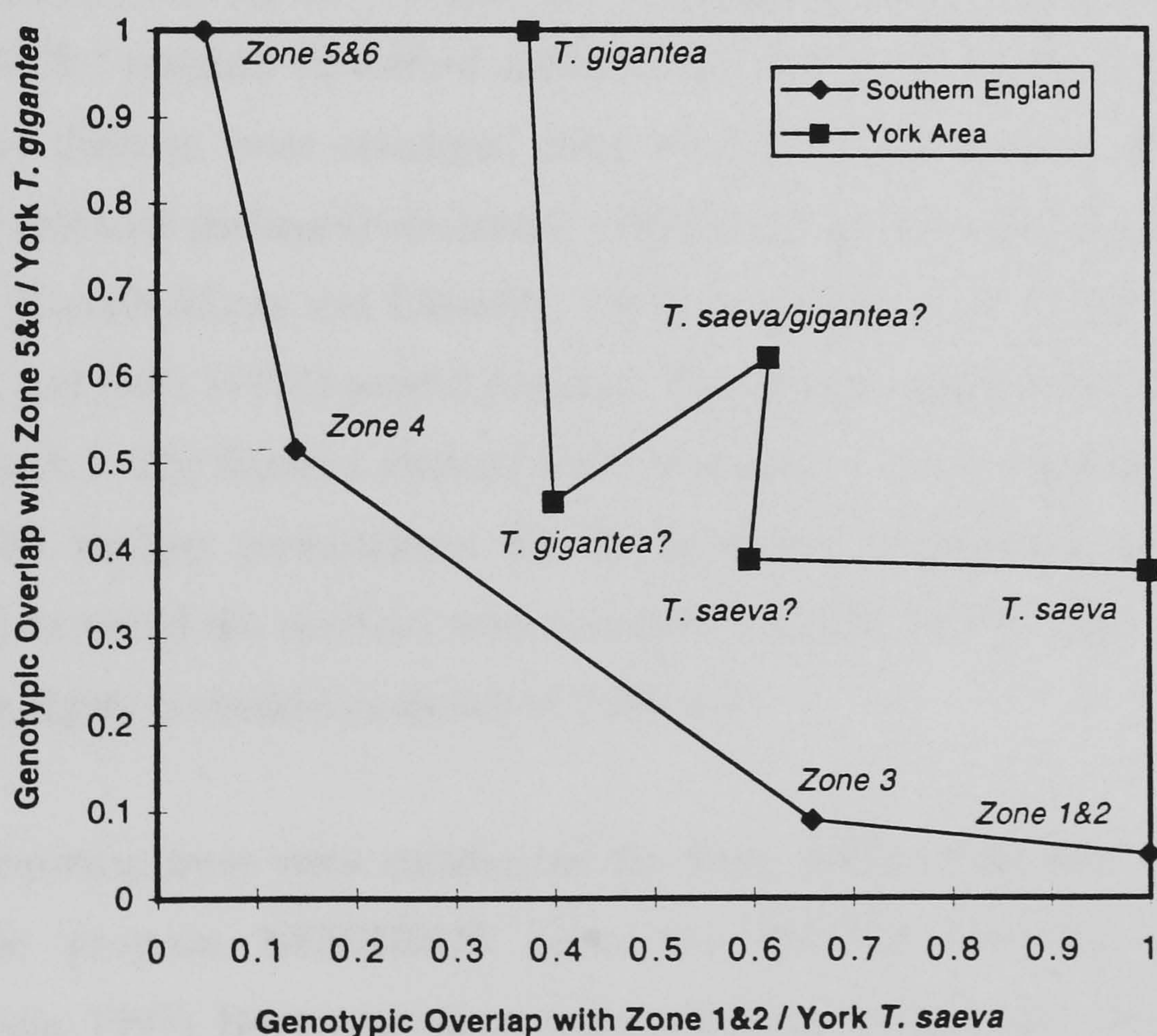


Figure 4.7. Genotypic overlaps of samples with zone 1&2 and zone 5&6 (southern England samples), and with the 'good' *T. saeva* and the 'good' *T. gigantea* (York area samples).

4.3.3 Genetic distances and 'population' trees

'Population' trees were generated in order to simplify interpretation of the relationships between the different sample groups. These are not meant to imply any phylogenetic relationship but merely illustrate the genetic relationships between the different groups. A variety of algorithms was used to calculate pairwise distance matrices among the five morphological groups from the York area and for zone 1&2, zone 3, zone 4, and zone 5&6 from southern England. The measure of genotypic overlap (S_g) was converted into a novel distance measure which takes account of genotype (D_g), using the simple relationship:

$$D_g = 1 - S_g$$

The other distance measures used are based upon allele frequencies. Rogers' (1972) genetic distance, with Wright's (1978) modification, was calculated using the BIOSYS-1 program (Swofford and Selander, 1981). Three other measures of genetic distance were calculated using the GENDIST program from the PHYLIP software package (Felsenstein, 1995); these are Cavalli-Sforza's chord distance (Cavalli-Sforza and Edwards, 1967), Reynolds *et al.* (1983) genetic distance, and Nei's (1972) genetic distance. The distance matrices are given in Appendix A.1. The distance matrices were subjected to a one-way Mantel test with 1000 random permutations (G. K. Roderick, unpublished computer program) to test if the matrices were correlated (Manley, 1994). All were, as expected, highly correlated as shown in Table 4.9

Neighbor-joining trees were constructed for these different distance matrices using the program NEIGHBOR from the PHYLIP software package (Felsenstein, 1995). Bootstrap values were calculated for the trees constructed using Nei's (1972) distance, the Cavalli-Sforza (1967) distance, and Reynolds *et al.*'s (1967) distance by subjecting the allele frequencies to 100 bootstrap replicates using the programs SEQBOOT, GENDIST, NEIGHBOR and

Table 4.7. Mantel tests on different measures of genetic distance.

	D_g	*Rogers*	*Cavalli-Sforza*	*Nei*
D_g	-			
Rogers	0.949	-		
Cavalli-Sforza	0.851	0.915	-	
Nei	0.904	0.969	0.945	-
Reynolds	0.930	0.981	0.944	0.991

The values in the table are the Mantel correlation statistic. $P < 0.01$ for all results under a one-way test. A one-way test was appropriate as distance measures are unlikely to be negatively correlated.

CONSENSE from the PHYLIP software package (Felsenstein, 1995). Bootstrap values for D_g and Roger's (1972)distance could not be calculated with the software available. The trees for D_g and Roger's (1972) distance are shown in Figure 4.8. The trees for the other three distance measures are shown in Figure 4.9. Different tree building algorithms (UPGMA, Fitch-Margoliash - from the PHYLIP package) yielded similar topologies.

The motive for employing a variety of distance estimates and tree-building algorithms was not related to the various genetic models and assumptions underlying each of these algorithms. It has been demonstrated (Nei *et al.*, 1983) that it is improbable that any combination of distance methods and tree-building algorithms will give the 'correct' topology when there are few genetic systems (less than the number of taxa - samples or 'populations' in the present case) and only small levels of genetic divergence. Therefore, following the approach of Goldberg and Ruvolo (1997), a variety of methods were employed for comparative purposes. Despite this, all the trees exhibited similar topologies. In all the trees zone 1&2 and zone 3 are separated from the rest of the tree by long branch lengths which are well supported by bootstrap values for the three trees for which these could be calculated. The next node in all the trees is for *T. saeva* from the York area, another well supported node. The York *T. saeva?* is the next node in all trees, though not so well supported. In all trees except the Cavalli-Sforza and Edwards (1967) tree the putative hybrid

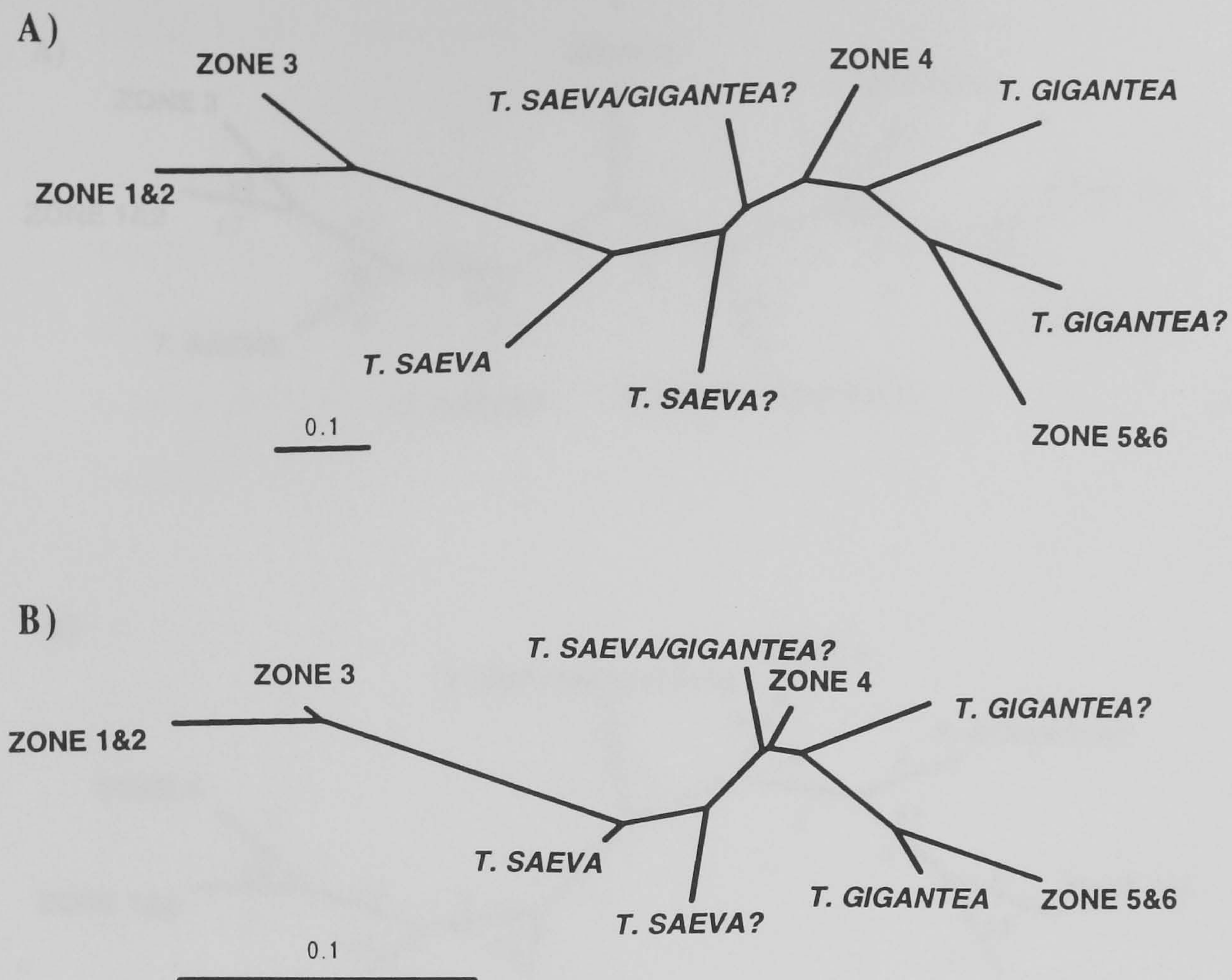


Figure 4.8. Neighbor-joining trees for all sample groups generated from A) D_g and B) Rogers' (1972) genetic distance with Wright's (1978) modification. Scale bar represents branch length. Samples from southern England: zone 1&2 = allopatric *T. saeva*; zone 3 = parapatric *T. saeva*; zone 4 = parapatric *T. gigantea*; zone 5&6 = allopatric *T. gigantea*. All other samples are from York (according to the morphological groupings - see section 4.2.1).

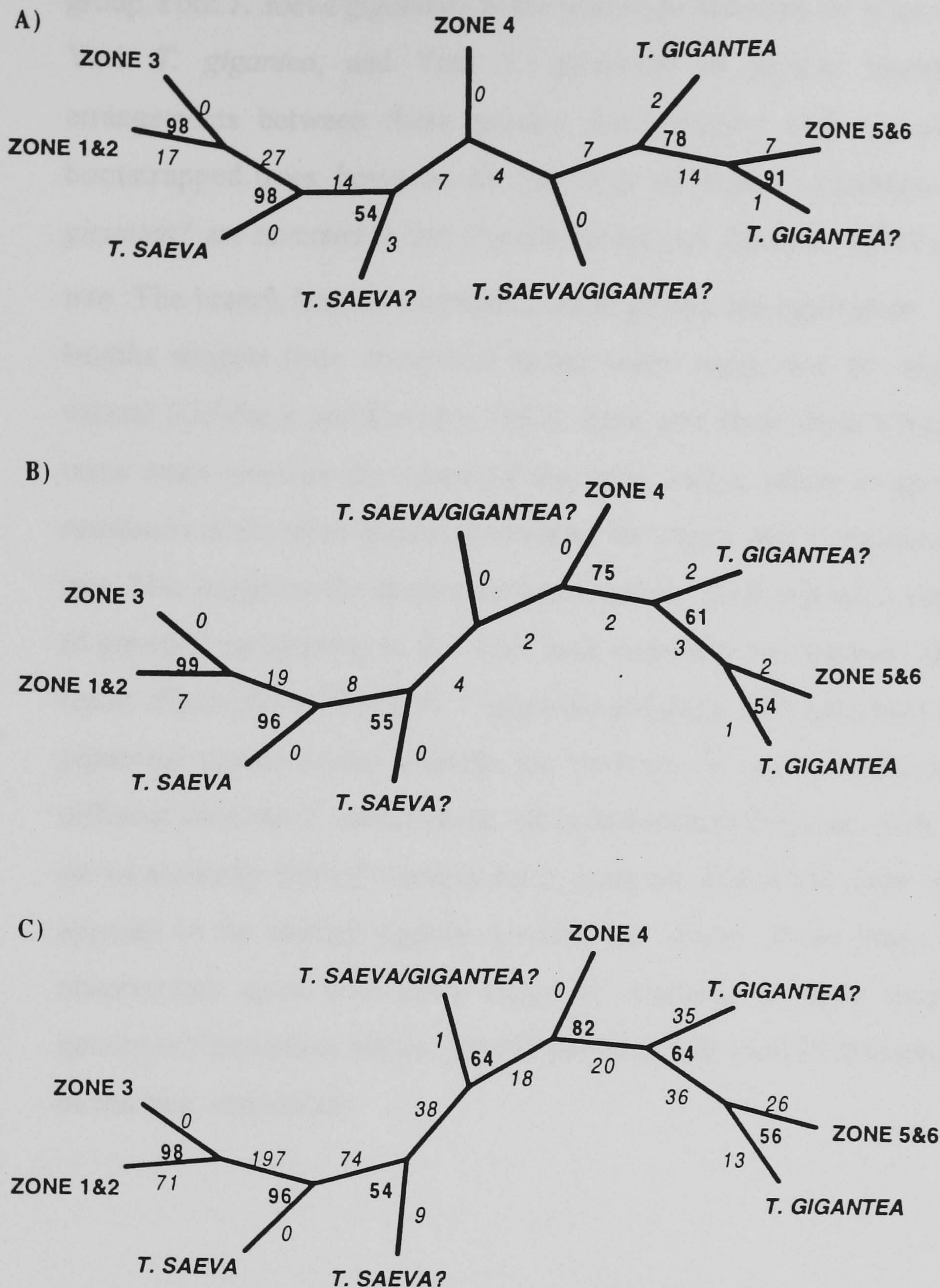


Figure 4.9. Neighbor-joining trees for all sample groups generated using A) Cavalli-Sforza and Edwards chord distance (1967), B) Nei's (1972) genetic distance and C) Reynolds *et al.* (1983) genetic distance. The numbers at the nodes, in bold, are bootstrap values (only values above 50% are shown). The numbers by the branches, in italics, are relative branch lengths (zero (or negative) branches score zero, the shortest branch scores 1 and all other branches are relative to this (after Goldberg and Ruvolo, 1997; and T. Goldberg, *Pers. Comm.*)). Samples from southern England: zone 1&2 = allopatric *T. saeva*; zone 3 = parapatric *T. saeva*; zone 4 = parapatric *T. gigantea*; zone 5&6 = allopatric *T. gigantea*. All other samples are from York (according to the morphological groupings - see section 4.2.1).

group York *T. saeva/gigantea?* is the next node followed by zone 4. Zone 5&6, York *T. gigantea*, and York *T. gigantea?* all cluster together and the arrangements between these groups are generally well supported in the bootstrapped trees; however, the positions of York *T. gigantea* and York *T. gigantea?* are reversed in the Cavalli-Sforza and Edwards (1967) and the D_g tree. The branch lengths connecting these groups are quite short. Zero branch lengths suggest little resolution (even when supported by high bootstrap values) (Goldberg and Ruvolo, 1997). Zero and short branch lengths tend to occur more towards the centre of the trees, and it seems in general that the resolution of the trees is poorer towards the centre and *T. gigantea* end of the tree. This supports the results discussed earlier which indicated a greater degree of genetic homogeneity in the York area, compared to southern England, as a result of gene flow. The York *T. gigantea* and zone 5&6 (southern allopatric *T. gigantea*) appear similar whereas the York area *T. saeva* appears to be very different from the *T. saeva* (zone 1&2) in Southern England, with the York *T. saeva* markedly shifted towards the *T. gigantea* end of the trees. Also, zone 4 appears to be shifted slightly towards the centre of the trees. Both these observations agree with those from the analyses of allele frequencies and genotype frequencies, above. Overall the Reynolds *et al.* (1983) tree appears to be the best supported.

4.4 Discussion

The limited screening of enzyme systems failed to find any fixed allelic differences in allozymes between the allopatric populations of *T. saeva* and *T. gigantea* from southern England (zone 1&2 and zone 5&6). Perhaps this was not surprising given the evidence for hybridization and gene flow between these species, although a more extensive allozyme survey could reveal some fixed differences. The identification of fixed allelic differences between the species would be particularly valuable because it could facilitate the identification of F_1 and backcross individuals with a high degree of certainty (Avise, 1994). This can provide a powerful diagnostic tool in conjunction with mtDNA analyses which would allow assignment of the female (and hence male) parent for each allozymically characterized specimen (Avise, 1994). If fixed allozyme differences remain elusive a more sensitive approach such as the development if microsatellite markers for these species could prove valuable. There were however extensive differences between the allopatric populations of the two species in terms of allele and genotype frequencies.

The variation between the sample groups was analysed by a number of approaches: empirical descriptions of allele frequencies and genotype (genotypic overlap) frequencies, and tree-based descriptions of the genetic relationships between the samples. All of these treatments of the data revealed similar patterns and therefore provide some confidence in interpretation.

Tegenaria saeva and *T. gigantea* are clearly differentiated in southern England, with the allopatric samples showing marked differences in allele and genotype frequencies. There was evidence of an overspill of alleles between the parapatric populations which is highly suggestive of hybridization and gene flow. There was little evidence of introgression of nuclear markers over a long distance, however it should be emphasized that the species were not sampled at a sufficiently detailed level to examine fine patterns in gene flow and introgression across the hybrid zone. There is a need for more detailed transects across the hybrid zone at the population level, employing more enzyme systems, and the calculation of F_{ST}

statistics and N_em estimates. One intriguing observation was that the parapatric *T. gigantea* sample appeared to contain a higher proportion of *T. saeva* genes than the parapatric *T. saeva* sample contained *T. gigantea* genes. This could suggest some asymmetry in hybridization/introgression at the species boundary, for instance more backcrosses may occur (or more progeny survive) between *T. gigantea* and F_1 hybrids than between *T. saeva* and F_1 hybrids. However, the results for the York area suggest that the asymmetry in introgression is in quite the opposite direction (see below). Another possible explanation could be that the parapatric *T. gigantea* sample contains a significant proportion of F_1 hybrids; the morphometrical analyses in Chapter 3 suggested that F_1 hybrid females may be frequently overlooked on examination of morphology. This possibility is discussed further in Chapter 7.

Compared to southern England, the material from the York area showed a much greater degree of genetic homogeneity which suggests a much greater degree of gene flow and hybridization in the this area. The pattern of gene flow was asymmetric. The York *T. gigantea* were genetically similar to the allopatric *T. gigantea* from southern England in all analyses. However, the York *T. saeva* did not resemble the allopatric *T. saeva* but were more similar to *T. gigantea*. This indicates a marked asymmetry in introgression, with the York *T. saeva* experiencing an influx of *T. gigantea* genes and a subsequent shift in the genetic profile in this species towards *T. gigantea*. Given the congruency between the patterns observed in the allozyme data and those revealed from morphometry, it would seem unlikely that the apparent shift of *T. saeva* towards *T. gigantea* in the York area could be the chance result of geographical variation in *T. saeva*. Rather, it seems likely that the asymmetry in introgression must result from genuine asymmetries in behavioural or mechanical isolation (see Chapter 6), or postzygotic isolation. Asymmetrical introgression of nuclear markers has been reported for other taxa. For instance, at the contact zone between *Mus musculus* and *M. domesticus* in Denmark Hunt and Selander (1973) observed free interbreeding within the hybrid zone (as indicated by agreement between allozyme genotype frequencies and random-mating expectations), yet adjacent to

the hybrid zone they found extensive introgression of some *M. domesticus* alleles into *M. musculus* but little gene movement in the opposite direction. This asymmetry was proposed to result from selection against introgression of the genes (or chromosomal segments that they mark) through backcross generations experiencing (asymmetrically) reduced fitness from the disruption of coadapted parental gene complexes (Avise, 1994; Hunt and Selander, 1973).

The patterns of hybridization and introgression revealed by the nuclear markers are discussed further, in light of the morphometrical, mtDNA and behavioural data, in Chapter 7.

Although Nei's (1972) genetic distance (Nei's *D*) was calculated, estimates of divergence times were not calculated for Zone 1&2 and Zone 5&6 (which are taken to represent relatively pure samples of *T. saeva* and *T. gigantea*). Any estimate of Nei's *D* from electrophoretic data will be prone to huge errors, which will only be reduced by considering a *large* number of random loci (Hillis and Moritz, 1990). Very few loci were analysed in the current study. Further, no electrophoretic molecular clock exists for spiders. In fact Hillis and Moritz (1990) failed 'to locate any applicable data' for any species in attempting to provide confidence limits for an allozyme clock. Allozyme clocks suffer from a lack of reliable independent information about times of speciation and because of this estimates of time are probably no better than arbitrary guesses (Hillis and Moritz, 1990).

A number of areas require further work. The hybrid zone in southern England should be examined in more detail (see above). It would also be valuable to compare the results presented here with data from transects across Spain or western France (where all three species occur). The four samples of *T. atrica* appeared to resemble *T. saeva* more than *T. gigantea* in terms of alleles, with the exception that only *T. atrica* possessed *Aco-1* allele *B*. The allozyme variation in *T. atrica* warrants further examination. It would also be interesting to explore the mathematical behaviour of genotypic overlap (S_g) by computer modelling.

5 Mitochondrial DNA Species Markers: Introgression and Phylogeny

5.1 Introduction

5.1.1 Aspects of Mitochondrial Genetics and Function

The metazoan mitochondrial genome is a circular molecule typically containing 13 protein-encoding genes, two ribosomal RNA genes, 22 tRNA genes and a control region containing an origin of replication (Avise *et al.* 1987; Moritz *et al.*, 1987; Harrison, 1989; Crozier, 1993). The minimum amount of DNA required for all these functional regions is about 15 kb, with animal mtDNAs typically being about 16.5 kb (Crozier, 1993). Vertebrate control regions contain a displacement loop referred to as the D-loop. In invertebrates there is no such loop and the control region is referred to as the A + T-rich region because of its base composition. Heteroplasmy, the possession of two different mtDNA genomes in the same individual (generally observed as intra-individual variation in mtDNA length and resulting from duplications or extensions in the A+T region), is rarely reported but has been described in several taxa including the *Pissodes* beetles (Boyce *et al.*, 1989), *Gryllus* crickets (Rand and Harrison, 1986) and some *Drosophila* (Hale and Singh, 1987). The organization of the *Drosophila yakuba* mtDNA molecule is illustrated in Figure 5.1 (taken from Simon *et al*, 1994 - originally redrawn from Clary and Wolstenholme, 1985).

The origin of mitochondria (and chloroplasts) in eukaryotic cells is generally accepted to be endosymbiotic (Margulis, 1981), whereby engulfed prokaryotes were retained by the eukaryotic ancestors because of mutually beneficial relationships (Gray, 1989a; 1989b; Li and Graur, 1991). Two important

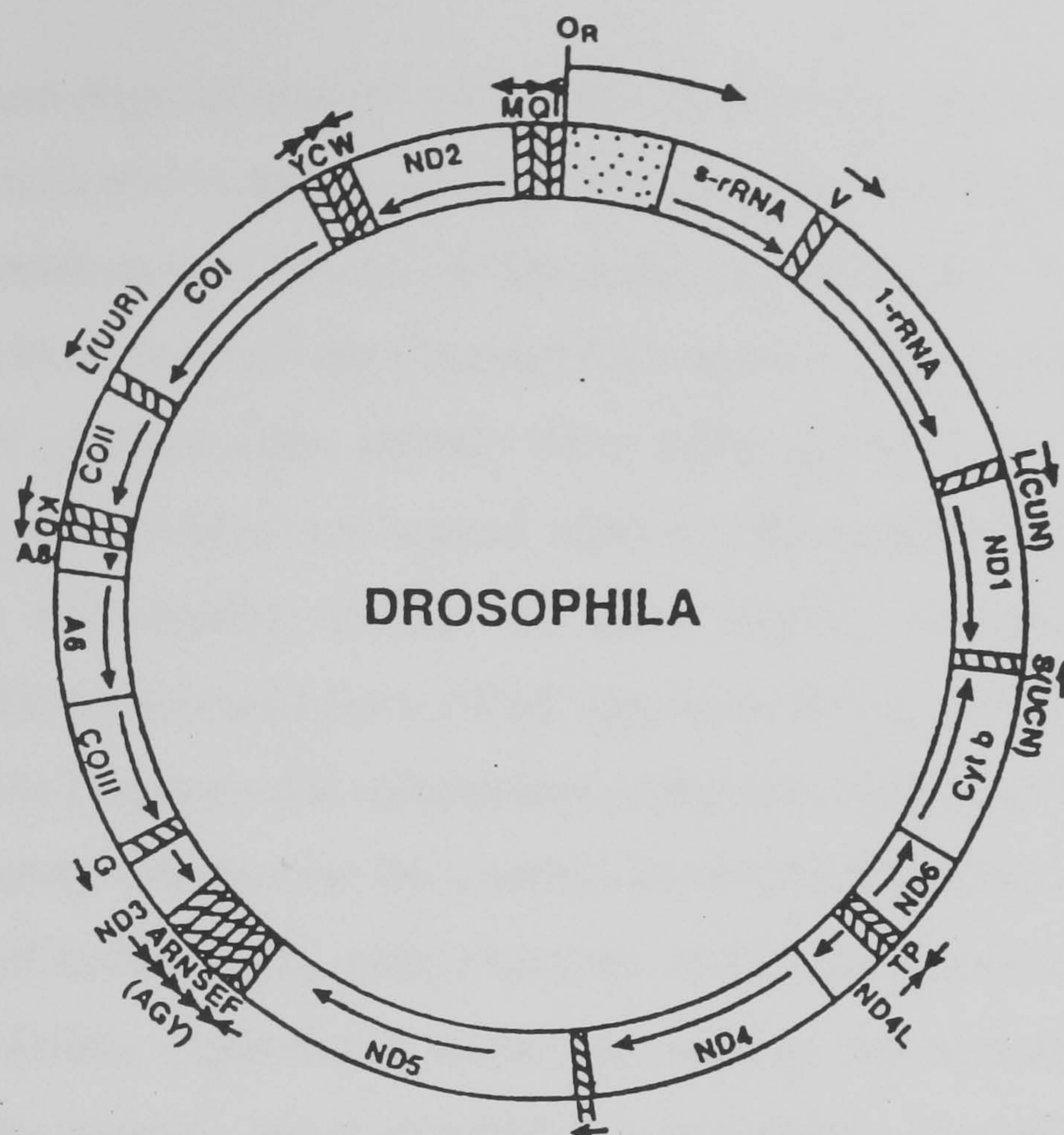


Figure 5.1. The organization of the *Drosophila yakuba* mitochondrial DNA molecule. O_R = origin of replication. Stippled box = A + T-rich region. Shaded regions = tRNA genes with single-letter amino acid code. l-rRNA = large ribosomal subunit RNA (16S). s-rRNA = small ribosomal subunit RNA (12S). Cyt b = cytochrome b gene. A1 to A8 = ATPase subunit genes. COI to COIII = cytochrome oxidase subunit genes. ND1 to ND5 = nicotinamide adenine dinucleotide dehydrogenase subunit genes. Arrows indicate direction of transcription. (From Simon *et al.* 1994 - originally redrawn from Clary and Wolstenholme, 1985).

assumptions in this hypothesis are that nuclear and mitochondrial (and chloroplast) genomes derive from demonstrably distinct lineages with a long period of independence before endosymbiosis, and that there has been a massive and consistent transfer of genetic control from the organelle to the nucleus with time. Thirteen of the mitochondrial protein-encoding genes are subunits of enzymatic complexes required in electron transport or ATP synthesis (Harrison, 1989). Hence, although the mitochondrion has its own

genome and origin of replication, it lacks the necessary enzymes for replication and synthesis and is therefore dependent upon the nucleus. The existence of nuclear pseudogenes - non-functional copies of mitochondrial DNA sequences in the nucleus - supports the hypothesis that there is a mechanism of transferral of control to the nucleus (Mirol, 1996) (although of course the transfer of control must have been very ancient relative to the transfer of pseudogenes - all metazoan mitochondria contain the same number of genes (see above)). Overwhelming evidence from rRNA sequences (Gray, 1989a; 1989b; Li and Graur, 1991) supports the independent prokaryotic origins of these organelles as does recent evidence for the common evolutionary history of the heat-shock proteins of mitochondria, purple bacteria, and hydrogenosomes (an enigmatic ATP producing organelle in eukaryotes lacking mitochondria) (Bui *et al.*, 1996). The genetic content of mtDNA is remarkably constant across taxa but the positional arrangement has differed in every phylum so far examined (Moritz *et al.*, 1987). The full mtDNA of an arachnid has yet to be characterized and so the positional arrangements of the genes are uncertain. Although such rearrangements are potentially useful, the great utility of mtDNA for phylogenetic studies stems from the conservation of homologous genes, lacking introns (Moritz *et al*, 1987), across taxa.

In addition to the existence of simply arranged homologous genes, other characteristics of mtDNA have proved very attractive to evolutionary biologists: 1) mitochondria exist in high copy number (up to 50 % of the DNA in eggs) which facilitates isolation; 2) mitochondria exhibit a maternal inheritance that is clonal and lacks recombination (Moritz *et al.*, 1987; Harrison, 1989; Crozier, 1993). Maternal inheritance may partly stem from a dilution effect: eggs have many more mitochondria than sperm. However there is now much evidence to suggest that there is specific exclusion of paternal mtDNA (see Crozier, 1993). Paternal leakage has been demonstrated in *Mus*, *Drosophila* and *Mytilus* but generally occurs at very low levels. One intriguing

observation is that all of these studies have involved interspecific crosses, suggesting that genetically distinct mtDNA from foreign lineages may occasionally be able to escape the exclusion mechanisms normally operating on closely related paternal molecules (Avise, 1991; Gyllensten *et al.*, 1991). Recombination of mtDNA molecules is likely to be very rare, and would seldom result in new genotypes if it did occur (Brown, 1985). The maternal transmission of mtDNA implies that only a fraction of the population passes on its mtDNA to the offspring, making the effective population size for mitochondrial genes smaller than that for nuclear genes. Smaller effective population size implies that mtDNA alleles will become fixed in the population more rapidly than their nuclear counterparts. Fixation occurs by an inevitable process of 'stochastic lineage sorting' following from the simple fact that at each generation not all mothers will leave offspring: the mtDNA gene-tree is 'self-pruning'. Avise *et al.* (1984) showed that, in populations founded by n females, it would be highly probable that within $4n$ generations all descendants would trace their ancestry to a single female founder. The length of time this will take depends upon the initial size and growth-rate of the female population and the variance in offspring number.

Another attractive feature of mtDNA is that different regions of the molecule evolve at greatly differing rates. This means that different regions may be used to explore phylogenetic relationships at various depths, from populations through to orders and even phlya. The control region shows the greatest substitutional variation and the two ribosomal RNA genes the least, with the protein-encoding genes showing their own characteristic rates. This variation between different regions of the molecule limits the usefulness of determining a general rate of evolution, or molecular clock, for the entire molecule. Even within individual genes different regions evolve at different rates dependent upon the functional constraints of the amino acid composition in their products. Individual genes also appear to evolve at differing rates in different taxa

(Moritz *et al.*, 1987; Harrison, 1989; Avise, 1994). Uncertainties in rate constancy between species and problems in calibration - generally from an imperfect fossil record or other molecular data (Avise, 1994) - suggest caution in applying molecular clocks. However for certain genes, and among certain groups of taxa, such molecular clocks can prove useful, at least for recent to moderate divergence times (see below). A rate of evolution for the Cytochrome oxidase I gene of invertebrates has been devised by Brower (1994), and appears to be of broad applicability across invertebrate taxa. This rate, of 2.3 % divergence per million years between sequences, has been employed in this study.

Rates of mammalian mtDNA evolution have been suggested to be on average 5 - 10 times as fast as rates of nuclear DNA evolution (Brown *et al.*, 1979), but in invertebrates the average rates of evolution appear to be similar for both genomes (Powell *et al.*, 1986; Harrison, 1989; Crozier, 1993).

The mtDNA molecule in animals has a low G + C content, perhaps reflecting a bias towards more efficient use of these bases by polymerases (Mortiz *et al.*, 1987). Mitochondrial genetic codes (invertebrates, vertebrates and plants have subtly different codes) are more redundant than the "universal" genetic code - fewer tRNAs are employed in the translation of protein encoding genes. Most substitutions in the third codon position are silent (invoking no amino acid changes), and most first position changes produce changes to similar amino acids. There is also a bias towards codons ending in A and T which means that most of the A + T bias in the mtDNA molecule is absorbed by the third position. Silent substitutions occur far more frequently than replacement substitutions, as a result of the constraint that replacement substitutions will generally be deleterious. As a direct result of this the substitution frequency in each codon position follows the order: third > first > second. Transitions (A $\leftrightarrow$

G, C $\leftrightarrow$ T) are far more frequent than transversions (purine $\leftrightarrow$ pyrimidine) for two reasons: 1) transition mispairs do not require a conformational change in the double helix; 2) many transversions in the third position cause an amino acid replacement whereas no transition does (Wakeley, 1996). These properties cause predictable patterns in sequence evolution with increasing divergence. Typically, for a protein encoding gene, the frequency of transitions is far greater than the frequency of transversion at low levels of divergence. As divergences increase this bias gradually disappears. Also, sequence divergence tends to be linear over short divergence times, but as divergences increase above about 20% the rate of divergence decreases. This may correspond to most sites with weak functional constraints having been altered - only the less easily altered sites are left and reversal substitions may start to accrue in those that are already altered (Moritz *et al.*, 1987).

5.1.2 Mitochondrial DNA as a Phylogenetic Marker

For the reasons outlined above mtDNA has become the tool of choice for the reconstruction of phylogenies from molecular data. By judicious selection of the portion of the molecule to be analysed, different levels of phylogenetic inference can be examined, bounded at recent levels by polymorphism and stochastic lineage sorting and at deep levels by sequence saturation (Moritz, 1987). Maternal inheritance and lack of recombination make mtDNA particularly suited to tracing recent history, including colonization histories, introductions and bottlenecks (Harrison, 1989). Approximately constant divergence rates between similar taxa, at least for low to moderate levels of divergence, can be used to date the point at which two individuals or taxa last shared a common ancestor (although great caution should be used when applying molecular clocks; see below and section 5.4.2). Phylogenies of very recently diverged taxa may occasionally be hindered by ancestral

polymorphisms (Avise *et al.*, 1987): when insufficient time has passed for the processes of stochastic lineage sorting (described above in section 5.1.1) and mutation to lead to complete monophyly. Stochastic lineage sorting can also result in the perceived time of divergence from molecular data preceding the actual time of population differentiation (especially for recent splits) as mtDNA lineages that were divergent before the original population split may be the ones to survive. A similar problem can beset more ancient speciation events that occurred close together in time: gene-trees may show incongruencies with species trees (Avise, 1994).

Intraspecific studies, especially in species with limited dispersal abilities often reveal strong geographic structuring. Studies may reveal structure in both spatial and temporal dimensions, ranging from local subdivisions to ancient processes associated with for example Pliocene or Pleistocene events (Avise, 1994). Deeper phylogenetic relationships eventually become unresolvable as a result of the saturation of sequences with substitutions and multiple substitutions resulting in homoplasy (see section 5.2.7). This problem may be alleviated by focusing on the less variable regions, for example the rRNA and tRNA genes, and weighting against transitions or third position changes. However, studies of very deep phylogenetic events will eventually run into a wall when structural rearrangements and very slowly evolving nuclear sequences (for example histones) may be better markers (Avise, 1994).

There have been few published studies on arachnids using mtDNA sequence data. Notable cases include Hedin's (1997) population/species level studies on *Nesticus* cave spiders, the demonstration of polyphyly within the spider family Ctenidae (Huber *et al.*, 1993), the revelation that the radiation of the spider genus *Tetragnatha* in Hawaii has resulted from multiple colonizations by different species (Gillespie *et al.*, 1994), and the identification of extreme

genetic differentiation and subdivision in a cryptic species complex of pseudoscorpions (Wilcox *et al.*, 1997).

5.1.3 Mitochondrial DNA as a Measure of Gene Flow and a Marker of Introgression

The variability of mtDNA both within and between populations makes it an ideal tool for exploring the genetic structure of populations. The evolutionary forces that generate population differentiation include gene flow, random genetic drift, various modes of natural selection, mutational divergence, and differences in mating system, life-history and ecology (Avise, 1994). The amount of gene flow will clearly affect the rate and ability of populations to diverge from each other. Under equilibrium conditions the level of divergence among populations is a function of the numbers of migrants exchanged and one of the most common approaches to estimate this is to calculate estimates of Wright's (1951) N_em. This may be calculated from hierarchical F-statistics (Wright, 1951), private alleles (Slatkin, 1985) or allelic phylogenies (gene trees) (Slatkin and Maddison, 1989). As in Chapter 4 (allozymes), the current study was not structured in such a way as to allow the estimation of N_em across the hybrid zone in southern England; this is fodder for a future project. However, the concept of gene flow is introduced because it is so fundamentally entwined with introgressive hybridization and speciation.

Mitochondrial DNA is an ideal marker for examining introgression - the backcross-mediated movement of genes between species or well-marked, genetically differentiated populations (Avise, 1994). Other molecular markers, such as allozymes, are also important tools in the study of introgressive hybridization and have advantages such that one may be able to discriminate between parental types and different classes of hybrid (F_1, backcrosses etc.) on

the basis of genotype (hybrid classes may not always be reliably identified from morphology (Avise, 1994); see Chapters 3 and 4). The combination of such nuclear markers and mtDNA provides a very powerful way to investigate hybrid zones. The maternal inheritance of mtDNA and lack of recombination allows one to follow the maternal line. Hence one can infer which species was the mother in an F_1 hybrid (identified using other characters) and which sort of F_1 hybrid was the parent in a backcross. This is of great value as hybridization and introgression are frequently asymmetrical. Differences in behaviour between species or populations and between sexes, subtleties in morphology (especially of the genitalia), and the differential compatibilities of various alleles and disruption of coadapted parental gene complexes, when placed in a mixed genetic background, may all generate asymmetries.

Perhaps the classic example of asymmetrical introgression of mtDNA haplotypes is the well studied hybrid zone between the western race of the house mouse (*Mus musculus domesticus*) and the eastern house mouse (*M. m. musculus*). There are on-going arguments as to whether these two forms constitute species or sub-species, nonetheless, they meet and hybridize along a narrow north-south zone throughout central Europe. The extensively studied hybrid zone in East Holstein, northern Germany, revealed apparently extensive introgression of *M. m. domesticus* mtDNA into the *M. m. musculus* populations north of the hybrid zone, but with very little gene movement in the opposite direction (Prager *et al.*, 1993). Detailed surveys suggest that southern Sweden was colonized from East Holstein, about 4000 years ago, by a few recombinant *M. m. musculus* individuals after a series of island hopping and founder events. The Swedish *M. m. musculus* subsequently colonized northern Denmark, with the result that the *M. m. musculus* populations in southern Sweden, and north of the hybrid zone with *M. m. domesticus* in northern Jutland, all possess an *M. m. domesticus* mtDNA, but one that differs from that currently found elsewhere in Europe. (Avise, 1994; Gyllensten and Wilson, 1987; Prager *et al.*,

1993; Vanlerberghe, 1988) Other possible examples of asymmetric introgression include the hybrid zone between the field crickets *Gryllus pennsylvanicus* and *G. firmus* in which crosses of *G. firmus* males and *G. pennsylvanicus* females produce viable and fertile hybrids but the reciprocal cross does not (Harrison *et al.*, 1987), and the transfer of mtDNA from the northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*) (Tegelström, 1987).

In hybrid zones where Haldane's rule is operating (the absence, scarceness, inviability, or sterility of the heterogametic sex), and males are the heterogametic sex, one may see a high rate of mtDNA exchange relative to nuclear introgression as females leave open an avenue for cytoplasmic exchange, whereas males do not (Avise, 1994). In cases where females are the heterogametic sex, for example the flycatcher species *Ficedula albicollis* and *F. hypoleuca* (Tegelström and Gelter, 1990), one may see a high level of nuclear introgression mediated by fertile male hybrids but little cytoplasmic introgression (Avise, 1994).

Introgression can clearly confound phylogenies generated from mtDNA. Instances of ancient reticulate evolution may produce erroneous species phylogenies, and it may be difficult to untangle ancestral polyphyly and paraphyly from introgression in very recently diverged species.

5.1.4 Aims

The aims of the work presented in this chapter were: 1) to examine whether there is any phylogenetic support for the existence of a discrete *T. atrica* group sister-species clade in relation to other *Tegenaria* species; 2) to explore mtDNA variation between *T. saeva* and *T. gigantea*, and specifically to examine if there is any asymmetry in the flow (if any) of mtDNA haplotypes

between these two species. This constitutes the first study of hybridization and introgression in an arachnid using mtDNA markers.

5.2 Materials and Methods[1]

5.2.1 Sample Collection and Storage

The majority of the samples utilised for the DNA studies came from the public mediated surveys and field work in southern England and the York area (as described in Chapter 2). The specimen of *T. agrestis* was collected by G. S. Oxford from Liverpool. In addition there were several specimens collected abroad: *T. parietina* from northern Italy (coll. P. J. P. Croucher); *T. gigantea* from the Pacific Northwest of the USA (coll. G. Binford, University of Arizona); *T. atrica* from southern France (coll. M. Emerit, Montpellier), *T. atrica* and *T. saeva* from northern France (coll. R. Leborgne (University of Nancy), and *T. atrica* from Eire (coll. J. O'Connor, National Museum of Ireland). The specimens sequenced and their origins are detailed in Table 5.1 (page 176).

A major obstacle to molecular analyses is the nuclease degradation and shearing of the DNA in stored specimens. The best preservation conditions for insects have been shown to be freezing and storage in liquid nitrogen or freezer storage at -80°C. However, storage in ethanol (preferably at 4°C) denatures nucleases and provides a suitable alternative with minimal shearing at high ethanol concentrations (Post *et al.*, 1993). Arachnological collections are generally stored in 70% ethanol to facilitate examination, therefore all specimens were stored in 95% ethanol (to minimise shearing), with the exception of those collected by other workers which were stored in 70% ethanol. DNA was extracted from a few individuals freshly killed by freezing at -80°C.

---

[1]For details of solutions and reagents refer to Appendix A.2.

5.2.2 DNA Extraction

DNA extraction and purification was performed using the salt-extraction technique of Medrano *et al.*, (1990), as used in spiders by Gillespie *et al.* (1994), and modified as necessary for our laboratory equipment. This technique was originally developed for extracting DNA from nucleated vertebrate erythrocytes and is especially suited to small volumes of tissue.

DNA was extracted from single spider legs (generally leg IV). The leg was removed from the spider with a pair of sterile (ethanol flamed) forceps and placed into a clean, plastic weighing boat. The tissue was rinsed once with 1 ml 70% ethanol, which was applied and removed using a pipette, and then with 1 ml of distilled water. A further 1 ml of water was applied and the leg cut into small sections using a sterile razor blade. It is essential that all ethanol is removed from the tissue, consequently the minced tissue was allowed to rest in this water for 1 minute, after which the water was carefully pipetted away and a further 1 ml distilled water added for another minute. This too was discarded. The tissue was transferred to a 1.5 ml microcentrifuge tube to which 550 μl lysis buffer and 11 μl proteinase K solution (approximately 2 units) were added. The mixture was ground briefly with a P1000 pipette tip, previously sealed over a Bunsen, and then placed at 55°C in a water bath for a few hours or overnight until all the muscle had disintegrated. One hour before the end of the incubation time 5.5 μl RNAse A solution (approximately 5 units) was added.

Following digestion, 350 μl 5 M NaCl were added (giving a final concentration of 2 M NaCl). The mixture was vortexed gently for 15 seconds and then centrifuged at 10,200*g* for 30 minutes at room temperature. This precipitates most of the proteinaceous matter. 450 μl of the supernatant was transferred to a

new labelled tube and 900 μl of ice-cold (-20°C) absolute ethanol added. DNA was generally visible as a cloud at this stage. The tube was inverted gently several times to mix and placed at -20°C for two hours or at 4°C overnight for complete precipitation. The use of monovalent Na^+ cations (from the NaCl) reduces to coprecipitation of deoxynucleotides (Mirol, 1996). The tube was then centrifuged at 10,200*g* for 30 minutes at room temperature, the supernatant poured off and the pellet washed twice with 70% ethanol (with a spin of 5 minutes at 2550 *g* between washes). The pellet was air-dried until all the ethanol had evaporated. The nucleic acid was resuspended in 100 μl TE buffer and stored at -20°C. The quality and quantity of DNA was checked by running 5 μl on an 1.5% agarose gel (see section 5.2.5).

5.2.3 Amplification using the Polymerase Chain Reaction

The polymerase chain reaction (PCR) utilises thermal cycling to amplify a fragment of DNA between two regions of known sequence. Oligonucleotides of between about 15 to 30 nucleotides in length are designed to be complementary to these flanking regions. The oligonucleotides act as primers for amplification of the intervening fragment in the presence of a DNA polymerase which adds bases in a 5'→3' direction. *Conserved* primers are typically used; these are primers whose recognition sequence has been shown to be more or less constant across a wide range of species - thus facilitating PCR of the same gene fragment in previously unsampled species. A typical PCR cycle consists of three steps. In a thermal cycler, the reaction mixture is first heated to around 93 to 97°C for a few seconds to a few minutes. This denatures the double-stranded DNA to produce single-stranded DNA which can act as a template. The temperature and duration of this denaturing step is a compromise between sufficiently high temperature and duration to fully

denature the template and the damaging effect that prolonged or too high temperatures have on the polymerase. The thermal cycler then rapidly ramps the temperature down allowing the primers to anneal specifically to the template. This temperature should, ideally, be the highest possible temperature at which the primers anneal to the template: lower temperatures can lead to mis-priming and multiple (erroneous) PCR products. The optimum temperature will depend upon the degree of identity between the primers and the target sequences and also upon factors such as the relative proportions of purines and pyramidines in these sequences (primers rich in C and G residues have higher annealing and denaturing temperatures). Annealing temperatures as high as 65°C may be used when there is complete primer-target identity (Innis *et al.*, 1990). The third phase is the extension step during which the polymerase adds nucleotides to the 3'-ends of the primers. *Taq* DNA Polymerase (Promega) was used. This enzyme is derived from the thermophilic bacterium *Thermophilus aquaticus* and has a number of advantages over other available DNA polymerases. Although it carries a 5'→3' polymerization-dependent exonuclease activity, it lacks a 3'→5' exonuclease activity, has a high accuracy, a wide temperature tolerance, and an optimal activity at about 74°C (Sambrook *et al.*, 1989; Promega Protocols and Applications Guide, 1996). Consequently the extension step usually takes place at around this temperature, the duration depending upon the length and concentration of the template.

The thermal cycle is repeated between 25 and 45 times resulting in an exponential increase in the rate of DNA production (Mullis *et al.*, 1986; Mullis and Faloona, 1987; Saiki *et al.*, 1988). In addition to the primers, template DNA, and *Taq* DNA Polymerase, the PCR reaction mixture also contains: buffer (KCl facilitating primer annealing; $MgCl_2$ as a cofactor to the polymerase; Tris HCl pH 8.3); and deoxynucleotides in excess (dATP, dCTP, dGTP, dTTP).

Three different fragments of the mtDNA were amplified. Initially only a fragment of the cytochrome oxidase I gene (CO1) was to be analysed, however, initial sequencing experiments suggested that this fragment was not sufficiently variable to act as a marker for *T. gigantea* and *T. saeva* (ironically, this apparent lack of variability turned out to be because of introgression). Therefore overlapping fragments of the large ribosomal subunit (16S) and Nicotinamide adenine dinucleotide dehydrogenase subunit I (ND1) genes were also amplified. The following conserved primers were employed:

C1-J-1718 (Simon *et al.*, 1994)
5'- GGA GGA TTT GGA AAT TGA TTA GTT CC -3' (26mer)

C1-N-2191 (NANCY) (Simon *et al.*, 1994)
5'- CCC GGT AAA ATT AAA ATA TAA ACT TC -3' (26mer)

N1-J-12261 (SPIDND1) (Hedin, 1997)
5'- TCR TAA GAA ATT ATT TGA GC -3' (20mer)

LR-N-12945 (N116S) (Simon *et al.*, 1994; M. Hedin, pers. comm.)
5'- CGA CCT CGA TGT TGA ATT AA -3' (20mer)

LR-J-12887 (16SBR) (Simon *et al.*, 1994)
5'- CCG GTC TGA ACT CAG ATC ACG T -3' (22mer)

LR-N-13398 (16SAR) (Simon *et al.*, 1994)
5'- CGC CTG TTT AAC AAA AAC AT -3' (20mer)

The letters J and N refer to the maJority strand (that coding for most genes) and the miNority strand of the invertebrate mtDNA, with the numbers referring to the position of the 3' base in the complete mtDNA sequence of *Drosophila*

yakuba (Clary and Wolstenholme, 1985) following the nomenclature of Simon *et al.* (1994). The values in parentheses refer to the length of the oligonucleotide polymers. LR-N-12945 is identical to the primer given in Simon *et al.* (1994), LR-N-12945 (alias Faw16S2), except that an initial G at the 5' end has been deleted (M. Hedin, pers. comm.).

These primer pairs amplified a CO1 fragment of approximately 450 base pairs (bp), an ND1 fragment (including the tRNA *leu*^{CUN} and part of the 16S gene) of approximately 600 bp, and a 16S fragment (overlapping with the ND1 fragment) of approximately 500 bp.

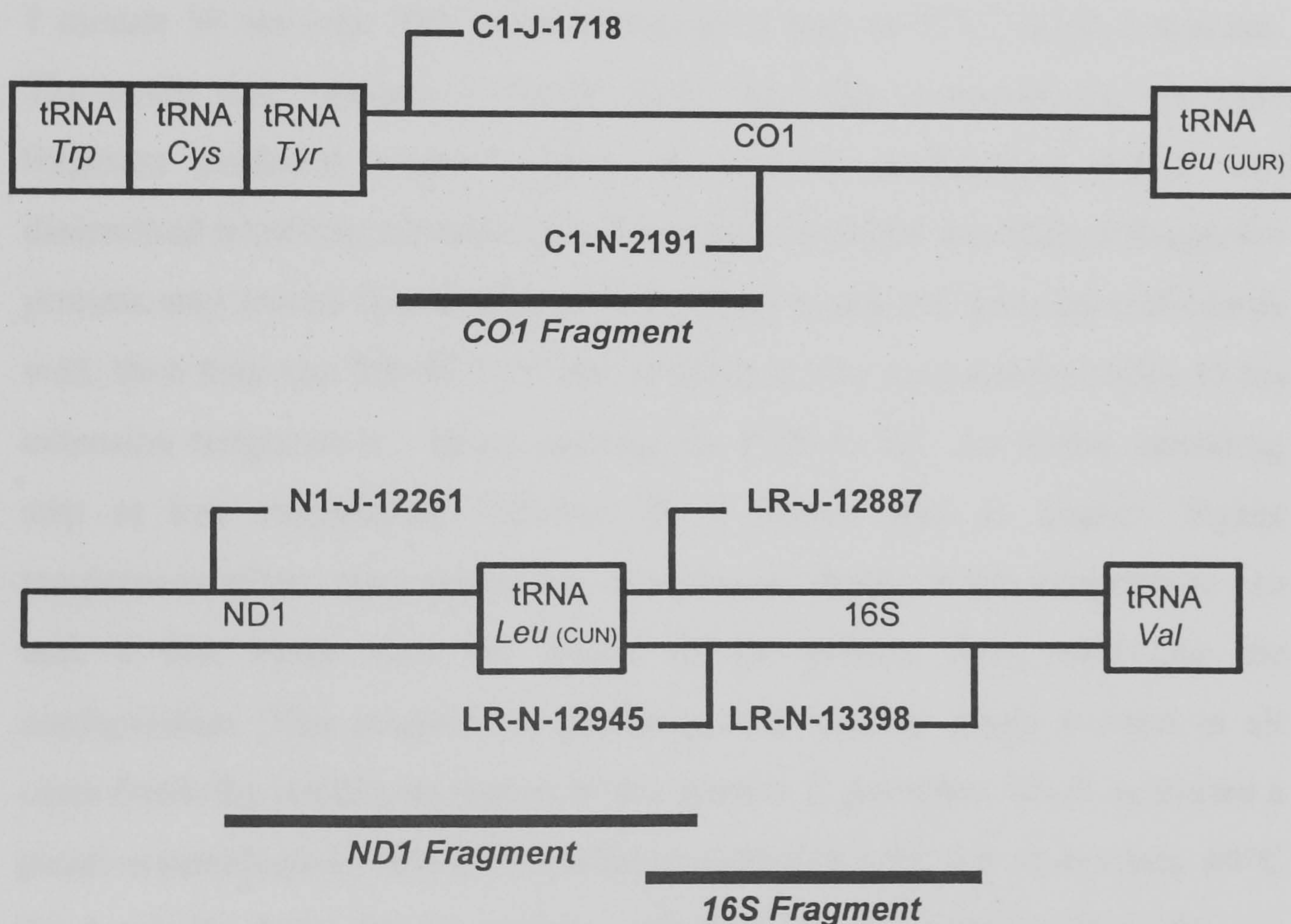


Figure 5.2. Strategy used for the amplification of the mtDNA gene fragments. The orientation of the primers is shown and the numbers correspond to the *Drosophila yakuba* mitochondrial DNA sequence (Clary and Wolstenholme, 1985; Simon *et al.*, 1994).

$MgCl_2$, primers and deoxynucleotides were titrated until optimal concentrations were obtained. A standard reaction mix of 50 µl or 25 µl, at final concentration, consisted of 1X *Taq* buffer, 2.5 mM $MgCl_2$, 200 µM of each of the four deoxynucleotides, 0.2 µM of each primer[2], 1 unit *Taq* DNA Polymerase (Promega), 20-30 ng template DNA, and filtered sterile distilled water to final volume. The thermal cycler possessed a heated lid and therefore the addition of mineral oil to prevent refluxing was unnecessary.

The PCR reaction was performed in an M.J. Research, Inc. PTC-100 programmable thermal cycler. The temperature profile for the ND1 and 16S fragments consisted of 95°C for 30 seconds, 47°C for 30 seconds, and 72°C for 1 minute 30 seconds. The initial denaturation step at 95°C lasted 2 minutes. The cycle was repeated 35 times. Early attempts at amplifying the CO1 fragment produced erratic results. An unusual temperature profile was determined involving a double annealing step. The logic was that, although the primers may anneal specifically, if they do not match the template sufficiently well, then they can lift-off from the template as the temperature ramps to the extension temperature - hence causing the PCR to fail. An initial annealing step at low temperature followed by a further step at slightly higher temperature allows the polymerase - functioning slowly at this temperature - to tack a few bases onto the 3'-end of the primer, thus stabilizing the configuration. This temperature profile gave a reliable, single product in all cases (with the notable exception of the species *T. parietina* which exhibited a putative pseudogene). The CO1 profile consisted of 95°C for 30 seconds, 40°C for 1 minute, 50°C for 30 seconds, and 72°C for 2 minutes. The cycle was repeated 35 times. The results were checked by running 5 µl of each PCR product, along with a control (no template DNA), and a DNA marker (ΦX174 Hae III digest), in a 1.5% agarose gel. PCR products were stored at 4°C until purification.

[2]Primer N1-J-12261 possesses a degenerate site and was therefore used at 0.4 µM for PCR.

The specimens of *T. parietina*, when amplified for the CO1 fragment, consistently exhibited a double PCR band (the faster - lower molecular weight - band apparently representing a pseudogene). These products were separated after gel electrophoresis by excising the two bands from a 3% agarose gel which had been run as long as possible to maximise the separation of the two products. The bands were cut out of the gel with a sterile razor blade and the slither of gel placed in an Microcentrifuge tube with 100 µl of filtered sterile distilled water and kept at 4°C overnight. The following day, 2 µl of the water, containing the leached product, was used as a template for a further amplification resulting in single bands.

5.2.4 Purification of PCR Products

PCR products were purified with the Wizard PCR Preps DNA Purification System (Promega), according to the manufacturer's recommended protocols. The procedure for using these 'minicolumns' can be summarized as follows:

The total volume of the PCR product was mixed with 100 µl of Direct Purification Buffer, briefly mixed, and 1 ml of Purification Resin added. The mixture was vortexed briefly three times over a one minute period and then added to a syringe barrel which had previously been inserted into a minicolumn. The mixture was then gently pushed through the column using the syringe plunger. The barrel was then removed from the column and the plunger removed before re-attaching the barrel to the column. 2 ml of 80% isopropanol were then similarly pushed through the assembly to wash the column. The minicolumn was placed onto a 1.5 ml microcentrifuge tube and centrifuged at 10,200*g* for two minutes to remove any isopropanol. After transferral to a new microcentrifuge tube, 50 µl of filtered sterile distilled water was added to the

column and left for 1 minute. The bound DNA fragment was eluted by centrifugation at 10,200*g* for 20 seconds.

After purification, 5 µl of the product was run in a 1.5% agarose gel to verify the presence and purity (e.g. the absence of primers) of the DNA. In initial PCRs the concentration and purity of the DNA was measured by spectrophotometry as the ratio of absorbance at wavelengths of 260 nm and 280 nm (Sambrook *et al.*, 1989).

5.2.5 Gel Electrophoresis of DNA[3]

DNA, from extractions, PCR products, and PCR product purifications, was visualized after agarose gel electrophoresis. Agarose was dissolved (by heating in a microwave oven) in 1X TAE buffer to give a 1.5% agarose gel (3% for separation of the *T. parietina* CO1 double-product). Electrophoresis was performed at 100 V in 1X TAE buffer on samples mixed with an equal volume of orange G loading dye. When the dye had migrated approximately three quarters of the way up the gel, the gel was transferred to a 1 µg/ml solution of ethidium bromide and the DNA stained for 20 minutes. The DNA was visualized under UV light and recorded using an UVP Image-Store 7500 PC and CCD camera. In order to estimate the size and concentration of PCR products, lanes containing the DNA marker ΦX174 Hae III Digest (Promega) were also run in these gels. The marker was diluted such that loading 5 µl would result in the intensity of the band corresponding to 603 bp (close to the PCR product size) being equal to approximately 10 ng of DNA (which is a quantity that is easily visualized). This was calculated according to the following formula:

---

[3] For details of solutions and reagents refer to Appendix A.2

$$x = 10 \times \frac{\sum MW}{MW}$$

where x is the amount of marker in ng to be added from the stock solution and MW is the molecular weight, in bp, of the fragments in the marker, i.e. the total length of all fragments (the uncut ΦX174 plasmid) is divided by the length of the desired fragment.

5.2.6 Sequencing

Direct (without cloning) chain-termination DNA sequencing (Sanger *et al.*, 1977) was employed to determine the nucleotide sequence of the purified PCR products. Chain-termination sequencing functions by *in vitro* polymerase mediated synthesis of a DNA strand, from a single-stranded template, proceeding from the site to which a specific oligonucleotide primer anneals. Strand synthesis is terminated when a nucleotide analogue (a 2',3'-dideoxynucleoside 5'-triphosphate (ddNTP)) is incorporated. The lack of the 3'-OH group, necessary for the formation of the phosphodiester bond, prevents concatenation. When the four dNTPs and the appropriate ddNTP are mixed in the correct proportions the reaction will terminate in a fraction of all the sites where the appropriate ddNTP can be incorporated. A labelled nucleotide is incorporated into the synthesis and the different length fragments separated by high-resolution electrophoresis and the results visualised. Two approaches to chain-termination were employed: manual (most of the CO1 sequences) and automated. In manual sequencing four reactions are performed - each one containing one of the four ddNTPs plus the four dNTPs - and the results are visualised by autoradiography. In automated sequencing one reaction is performed - all four ddNTPs plus dNTPs - and the results read from the base of the electrophoresis gel by a laser and analysed by a computer.

All PCR products were sequenced in both directions (providing verification of the sequence) using the same primers as employed in the PCR.

Manual sequencing was performed using the Sequenase Version 2.0 DNA Sequencing Kit (United States Biochemical/Amersham International plc.). The first step was to denature the double-stranded template and anneal the primers. In a typical annealing mixture, 1 µl (2 pmol) of the primer, up to 7 µl (approximately 50 ng) of PCR template, and sufficient sterile filtered distilled water to make 10 µl final volume, were placed into a 0.5 ml Microcentrifuge tube and boiled (100°C) for three minutes in a thermal cycler. The mixture was then plunged directly into ice for five minutes, briefly centrifuged and returned to ice. The termination reactions were performed in the wells of microtitre plates. For each reaction 2.5 µl of C, T, A or G termination mixtures (the appropriate ddNTP plus the four dNTPs) was pipetted into four separate wells - one for A, one for C and so on. The termination reaction involves two steps. Firstly the fragments must be labelled. Each labelling reaction consisted of 2 µl Reaction Buffer, 1 µl 0.1M dithiothreitol solution (DTT), 2 µl labelling mix (previously diluted approximately 5 times in filtered sterile distilled water), 0.5µl [α-^{35}S]dATP (at 10 µCi/µl and 10 µM (1000 Ci/mmol)) (Du Pont NEN), and 2 µl pre-diluted polymerase (Sequenase diluted 1:8 and inorganic pyrophosphatase diluted 1:16 with Enzyme Dilution Buffer). This constitutes enough 'master mix' for one reaction - dilution calculations were greatly simplified if four reactions were performed at once. 7.5 µl of the labelling reaction master mix was added to the 10 µl of annealing mixture and the tube given a pulse of centrifugation to mix and bring the contents to the base of the tube. The labelling reaction was left to incubate for 5 minutes at room temperature - during this period the primers are extended under limiting concentrations of dNTPs and incorporate the labelled nucleotide to generate a range of labelled fragments of randomly distributed length. The second step of

the termination reaction extends the fragments further and terminates them. The termination mixtures were heated to 37°C for 1 minute on a hot-plate and 3.5 µl of the labelling reaction added to each of the four wells (A, C, G, T). The reaction was incubated for 5-10 minutes at 37°C. The reaction was then stopped by the addition of 4 µl of a stop solution containing EDTA, formamide and two electrophoresis dyes: bromophenol blue and xylene cyanol FF. The sequencing reactions were stored at -20°C until electrophoresis, immediately prior to which they were denatured for 2-3 minutes at >80°C on a hot-plate and placed on ice. The four termination mixtures, 2.5-3.0 µl of each, was electrophoresed at 1500-3000 V (approximately 55°C plate temperature) in a 5% polyacrylamide gel using a Biorad Sequi-gen GT Sequencing Cell. The gel was cast and set for two hours and pre-run for 15-30 minutes prior to loading. The electrophoresis was stopped after the bromophenol blue had run off and as the xylene cyanol was about 4 cm from the base of the gel (xylene cyanol corresponds to fragments of about 130 nucleotides in a 5% gel (Sambrook *et al.*, 1989). The gel was allowed to adhere to a sheet of Whatman chromatography paper, dried, and exposed to an X-ray film for 2-3 days. Sequences were read from the developed autoradiograph.

Automated sequencing was performed using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq DNA Polymerase, FS (Perkin Elmer). Cycle sequencing uses a thermal cycler to produce a single stranded PCR with chain-termination. The use of thermal cycling greatly reduces the labour of sequencing as all reactions occur in a single tube and require very little template. Fluorescent dye terminators are employed removing the need to handle radioactive sources. For each reaction 4 µl of Terminator Ready Reaction Mix was pipetted into a 0.2 ml PCR tube on ice. To this was added 1-2 µl (5-15 ng) PCR product, 1.6 pmol primer, and filtered sterile distilled water to a final volume of 10 µl. The reaction was placed in an

M.J. Research, Inc. PTC-100 programmable thermal cycler and subjected to 25 cycles of 96°C for 30 seconds, 50°C for 30 seconds, and 60°C for 4 minutes. After the last cycle the machine was held at 4°C, until the samples were removed to a 4°C refrigerator. Excess dye terminators were removed by ethanol precipitation. Two µl 3M sodium acetate and 50 µl 95% ethanol were placed in a 1.5 ml microcentrifuge tube and the 10 µl of sequence reaction transferred to this mixture. The tubes were vortexed and placed on ice for 10 minutes, and then centrifuged at 10,200*g* for 15-30 minutes. The ethanol solution was carefully pipetted away and the pellet rinsed once with 70% ethanol. The pellet was then air-dried. Three µl of loading buffer was added to each sample before electrophoresis on an ABI PRISM 377 DNA Sequencer (Perkin Elmer) connected to a 7200/90 Power Macintosh. The electropherogram sequence output was edited and verified using the FACTURA and AUTOASSEMBLER programs (Perkin Elmer).

5.2.7 Phylogenetic Analysis

Sequences were edited as necessary using the SEQED program (ABI) and transferred to the GCG suite of programs running on the Daresbury Seqnet computer system (Program Manual for the Wisconsin Package, 1994). The coding regions of the sequences were translated with the TRANSLATE and MAP programs (Program Manual for the Wisconsin Package, 1994) using an invertebrate mitochondrial translation table. Sequences were then compared to other sequences deposited in the EMBL database using the FASTA routine (Program Manual for the Wisconsin Package, 1994). Sequences were aligned using PILEUP (Program Manual for the Wisconsin Package, 1994) and CLUSTALW version 1.7 (Thompson *et al.*, 1994).

There are numerous methods of phylogenetic reconstruction for molecular data and detailed reviews of these have been published (see for example Felsenstein, 1988; Hillis *et al.*, 1993; Swofford and Olsen, 1990). They can be divided into four basic categories: distance, character-state (parsimony), invariants, and maximum likelihood methods. The method of invariants (Lake, 1987; Swofford and Olsen, 1990) was a method developed to deal with the problem of homoplasies (see MAXIMUM PARSIMONY) of independent evolutionary origin bringing divergent sequences together. A problem that might ideally, but not always in practice, be solved by the addition of intermediate taxa. The method is complex and is not easily applied to more than four taxa and will be discussed no further. Only those methods relevant to the present work will be described.

1) Distance methods:

These methods construct phylogenetic trees using a matrix of evolutionary distances between all the pairs of taxa (the number of nucleotide or amino acid substitutions separating two taxonomic units - generally given as the average number of substitutions per nucleotide or amino acid). However the calculation of evolutionary distance is not straightforward and depends upon which of the numerous models of sequence evolution is employed. There are three commonly used models. The Jukes-Cantor model (Jukes and Cantor, 1969) is perhaps the simplest and includes among its assumptions that all four nucleotides occur equally frequently and that transitions and transversions are equally likely - this is rarely realistic. The Kimura two-parameter model (Kimura, 1980) still assumes that all four nucleotides occur equally frequently but allows for independent rates of transition and transversion substitutions. It can be shown that as divergence increases the Kimura model tends to be more accurate than the Jukes-Cantor model (Li and Graur, 1991). The third model is the generalization of the two-parameter model to allow the four nucleotides to occur at different frequencies (Swofford and Olsen, 1990).

Distance matrices were estimated using the Kimura two-parameter model, partly for the reasons outlined above, and partly as this model is most commonly used and allows a direct comparison with published divergence times/rates of evolution. These models of sequence evolution are of course vital to all methods of phylogeny reconstruction, for instance the maximum likelihood program DNAML (Felsenstein, 1995) uses the generalized two-parameter model.

CLUSTER ANALYSIS (UPGMA): The unweighted pair group method with arithmetic averages is the simplest method for tree reconstruction (Li and Graur, 1991). The tree is reconstructed from the distance matrix by stepwise addition of the least distant pair of taxa (or sequences). The data are assumed to be both additive and ultrametric such that the distance between two taxa is equal to the sum of the branches joining them, and that all taxa are equidistant from the root when the tree is rooted. This implies that there should be a completely uniform rate of evolution in the different lineages and minimal systematic errors in the data. This is rarely the case (Swofford and Olsen, 1990).

NEIGHBOR-JOINING (Saitou and Nei, 1987): Perhaps the most widely used method. The data are assumed to be additive such that the distance between two taxa is equal to the sum of the branches joining them, but need not be ultrametric so that all taxa are not necessarily equidistant from the root when the tree is rooted. The raw distance matrix is modified to generate a new distance matrix in which the separation between each pair of nodes is adjusted according to their average divergence from all other nodes (this effectively normalizes the divergence for each taxon for its average molecular clock rate). The least distant (nearest neighbor) taxa are linked, their ancestral node replaces these two terminal nodes, and the process is repeated. This sequential

pooling of nearest neighbors to form what will become the internal nodes of the tree minimizes the total length of the tree (Swofford and Olsen, 1990; Li and Graur, 1991).

FITCH-MARGOLIASH (Fitch and Margoliash, 1967): This method uses distance estimates to generate a least squares estimate of the fit of the data to an additive tree - the aim is to select the tree that best fits (minimizes) this estimate.

If the relative rates of evolution in the different lineages are the same, and the data are free of systematic errors, all the distance methods should give equivalent results.

2) Character-state methods:

MAXIMUM PARSIMONY: This method searches for trees of lengths that minimize the amount of evolutionary change required to explain the data. In other words, the shortest, or most parsimonious, tree is that one over which most characters undergo the least number of steps. In the context of molecular data the characters are the nucleotide (or amino acid) positions and the character states are A, C, G and T. Character states are *unordered* or free to transform into any other state - sometimes known as Fitch Parsimony (Fitch, 1971). Polymorphic sites are only considered to be potentially informative if at least two taxa contain the character states - they are shared and therefore favour one grouping over another. Such characters are termed synapomorphies. Character states observed in only one taxon are uninformative and defined as autapomorphic. The maximally parsimonious (MP) tree may contain characters that individually do not support the tree, for instance there may be a change A→G at nucleotide position 291 which occurs in two separate lineages on the MP tree, and therefore represents two evolutionary steps. Such characters represent homoplasy (convergence, parallelism, or reversal) (Swofford and Olsen, 1990).

The MP tree (or trees) may be searched for by exact or heuristic methods. Exact methods include exhaustive searches of all possible trees and branch and bound searches which set an initial upper bound (length) and then search exhaustively until a shorter tree is found which is used to set a new shorter upper bound. Both the exhaustive and the branch and bound approaches guarantee to find the shortest tree but as the number of possible trees increases exponentially with the number of taxa these approaches may be prohibitively time-consuming. Heuristic searches take an initial random tree and rearrange it to minimise the length, either by stepwise addition of taxa to an initial random tree with three taxa, or by branch-swapping - trying all possible rearrangements of the initial tree. There are many ways to maximise the efficiency of heuristic searches but they cannot guarantee to find the shortest tree.

3) Maximum likelihood methods (Felsenstein, 1981b):

These methods evaluate the net likelihood that a specified model of sequence evolution (Jukes-Cantor, Kimura two-parameter, generalized two-parameter model) would yield the observed sequences given the branching order and branch lengths of a certain tree. The inferred phylogenies are those with the highest likelihoods (Swofford and Olsen, 1990).

Character-state methods are often regarded as being more powerful than distance-based methods as they use the raw data, which is in the form of a string of character states, and concentrate on shared character states - thus most likely to reflect accurately evolutionary history. However parsimony methods only use a limited fraction of the raw data. If the divergence between sequences is very low then parsimony may not find enough informative sites to be effective (Saitou and Nei, 1986; Li and Graur, 1991). It has also been shown that if the divergence or rate of change between two lineages is very high, or unequal, parsimony may make inconsistent estimates of the evolutionary tree

(Felsenstein, 1981b; 1983; 1988). In general maximum likelihood, using all the data, may be the most efficient method (Nei, 1996), however it is computationally intensive and few large trees have been published (Swofford and Olsen, 1990). In general, if the data are robust and contain a strong phylogenetic signal, most methods will produce comparable results.

The confidence of the results can be tested in many ways. The most common method is to use the bootstrap (Efron, 1979). Successive random samples of the original data-matrix are constructed by random removal of data columns (with replacement), and the construction of a tree for each new data-matrix. This is repeated many times (typically 100-1000 times). An estimate of the statistical reliability of the internal nodes in the phylogenetic tree from the original data-matrix is the frequency of occurrence of these nodes in the set of trees derived from the resampled data-sets.

Maximum Parsimony analyses on the aligned sequences were run on a Motorola StarMax 3000/200 Power PC using PAUP, Version 3.1.1 (Swofford, 1993). For each data set the frequency distribution of the lengths of 10,000 unrooted random trees was examined. The shape of this distribution provides a good indicator of structure in the data; random sequence variation would produce a normal distribution whereas data containing a strong phylogenetic signal produces a left-skewed distribution (Hillis, 1991; Hillis and Huelsenbech, 1992). This indicates non-random correlation among the characters and few trivial solutions close to the optimum one. Hillis and Huelsenbech (1992) suggest a statistic (g_1) to describe the distribution, and provide tables of probability values:

$$g_1 = \frac{\sum_{i=1}^{n}(T_i - \overline{T})^3}{ns^3}$$

where n is the total number of trees, T_i is the individual tree length, and s is the standard deviation of tree lengths. For a left-skewed distribution $g_1 < 0$. Exhaustive searches were performed to find the MP trees and bootstrap values were estimated by generating 1000 replicates using a branch and bound algorithm. The number of unambiguous changes, and the frequency of these changes in each substitution class were estimated from the MP trees using MACCLADE (Maddison and Maddison, 1992).

Distance matrices for the aligned sequences were generated using the DNADIST program in the PHYLIP package (Felsenstein, 1995) running under the Daresbury Laboratory Seqnet Service (Program Manual for the Wisconsin Package, 1994). Neighbor-joining trees were generated with the PHYLIP program NEIGHBOR. One thousand bootstrap replicates were generated using the PHYLIP program SEQBOOT, and a consensus neighbor-joining tree drawn using the PHYLIP programs DNADIST, NEIGHBOR and CONSENSE to give bootstrap values. Neighbor-joining trees and 1000 bootstrap replicates were also generated using CLUSTALW version 1.7 (Thompson *et al.*, 1994). UPGMA trees were generated using NEIGHBOR, Fitch-Margoliash trees using FITCH, and maximum likelihood trees using DNAML (all PHYLIP programs). Trees were viewed using TREEVIEW (Page, 1996).

MP and neighbor-joining trees are presented in the results. All the above tree building algorithms produce unrooted trees (there is no *a priori* inference of an ancestral character state). Trees were rooted for orientation and display purposes to an outgroup taxon. The outgroup taxon should neither be so distant as to make it difficult to gain reliable estimates of the distance between the outgroup and the other taxa, nor so close that one cannot be certain that it *is* an outgroup - the outgroup must have diverged from the other taxa before they diverged from each other. For the CO1 sequences *Tetrix denticulata*, representing a sister genus, was employed. For the other analyses (i.e. for

specific consideration of the relationships and interaction of *T. saeva* and *T. gigantea* using the ND1 and the combined CO1 and ND1 data) *T. atrica* was used as the outgroup as it was the most divergent taxon and had consistently branched before the *T. saeva* and *T. gigantea* haplotypes in the CO1 analyses. The results section will make this clearer.

5.2.8 Relative Rate of CO1 Evolution

The binomial relative rate test of Mindell and Honeycutt (1990) was performed to assess if there were significant differences in the rate of CO1 evolution between taxa, and between taxa and a putative pseudogene. The test proceeds by taking two taxa and a taxon that is an outgroup to these; as the ingroup taxa are more related to each other than to the outgroup they are of equal age relative to the outgroup. If the two ingroup taxa have experienced equal evolutionary rates then 50 % of the substitutions relative to the outgroup should be found in each ingroup taxon. The binomial test was calculated according to the following formula:

$$P = [\ \sum_{r=0}^{m} m!/(m-r)!r!\ ]0.5^{m}$$

where r is the number of unique differences in one ingroup species relative to the outgroup (whichever has fewer differences), and m is the combined number of unique differences for each ingroup species. The probability was doubled for a two-tailed test (DeWalt *et al.*, 1993).

5.3 Results A: Mitochondrial DNA Phylogeny

(Major analyses of sequences reasonably presumed to be mitochondrial are described here. Data on a putative pseudogene are given in section 5.4 Results B).

5.3.1 DNA Extraction, PCR, and Sequencing.

Extraction of spider DNA, from individual legs of ethanol-stored specimens, generally worked well. It was clear from DNA extracted from frozen material that ethanol-stored specimens suffered some DNA degradation but this was insufficient, in general, to prevent its use as a PCR template. In all cases a single PCR product was consistently obtained, with the exception of CO1 from *T. parietina* which yielded two products, the smaller of which apparently representing part of a pseudogene. The following sets of figures (Figures 5.3 to 5.6) show agarose gels of DNA extractions and PCR products, a section of a manual sequencing gel autoradiograph, and electropherogram output from an automated sequencing gel.

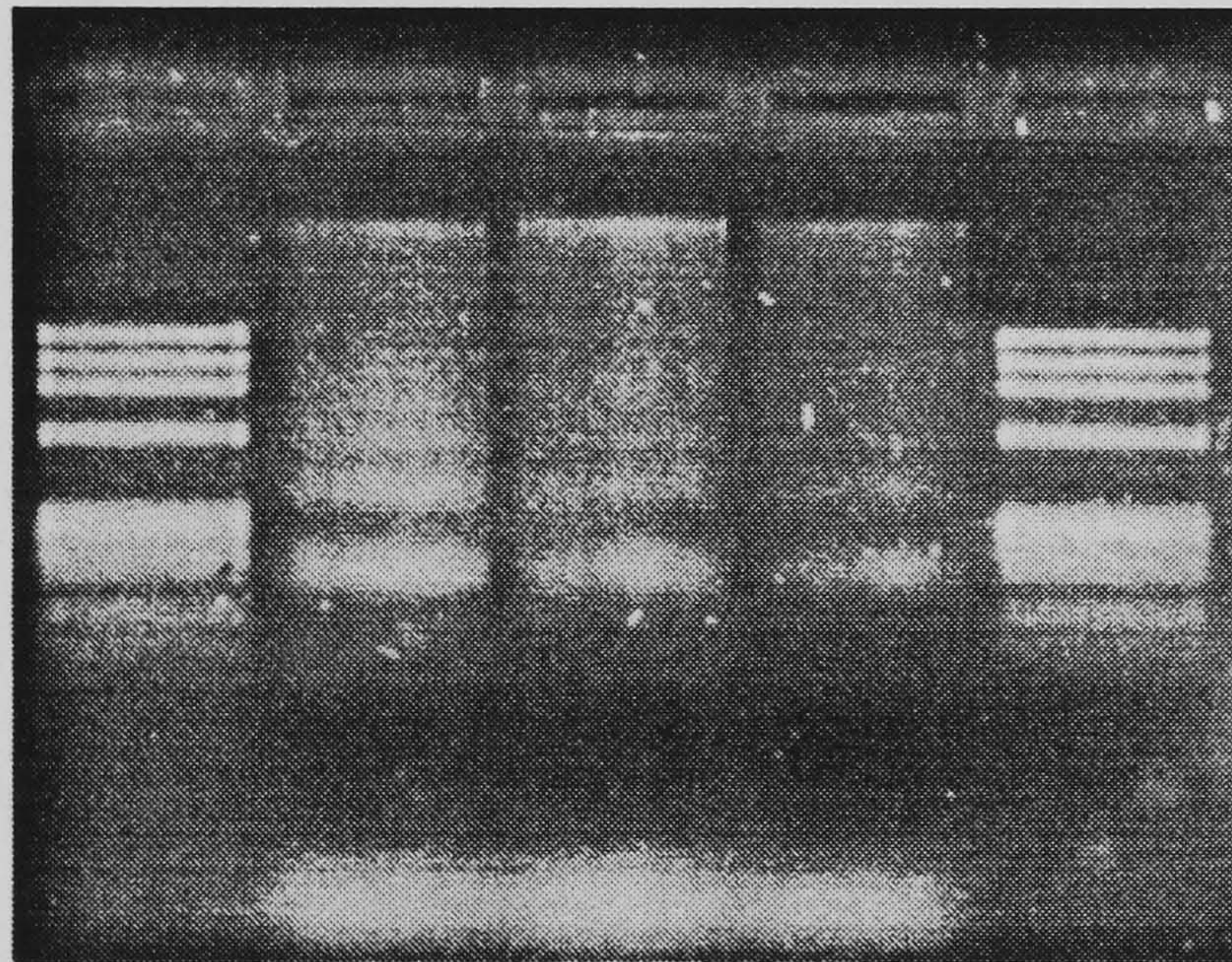


Figure 5.3. Agarose gel electrophoresis of DNA extractions. Three DNA extractions are shown between two lanes containing the size marker ΦX174 Hae III digest (Promega). Marker fragment sizes are given in Figure 5.4.

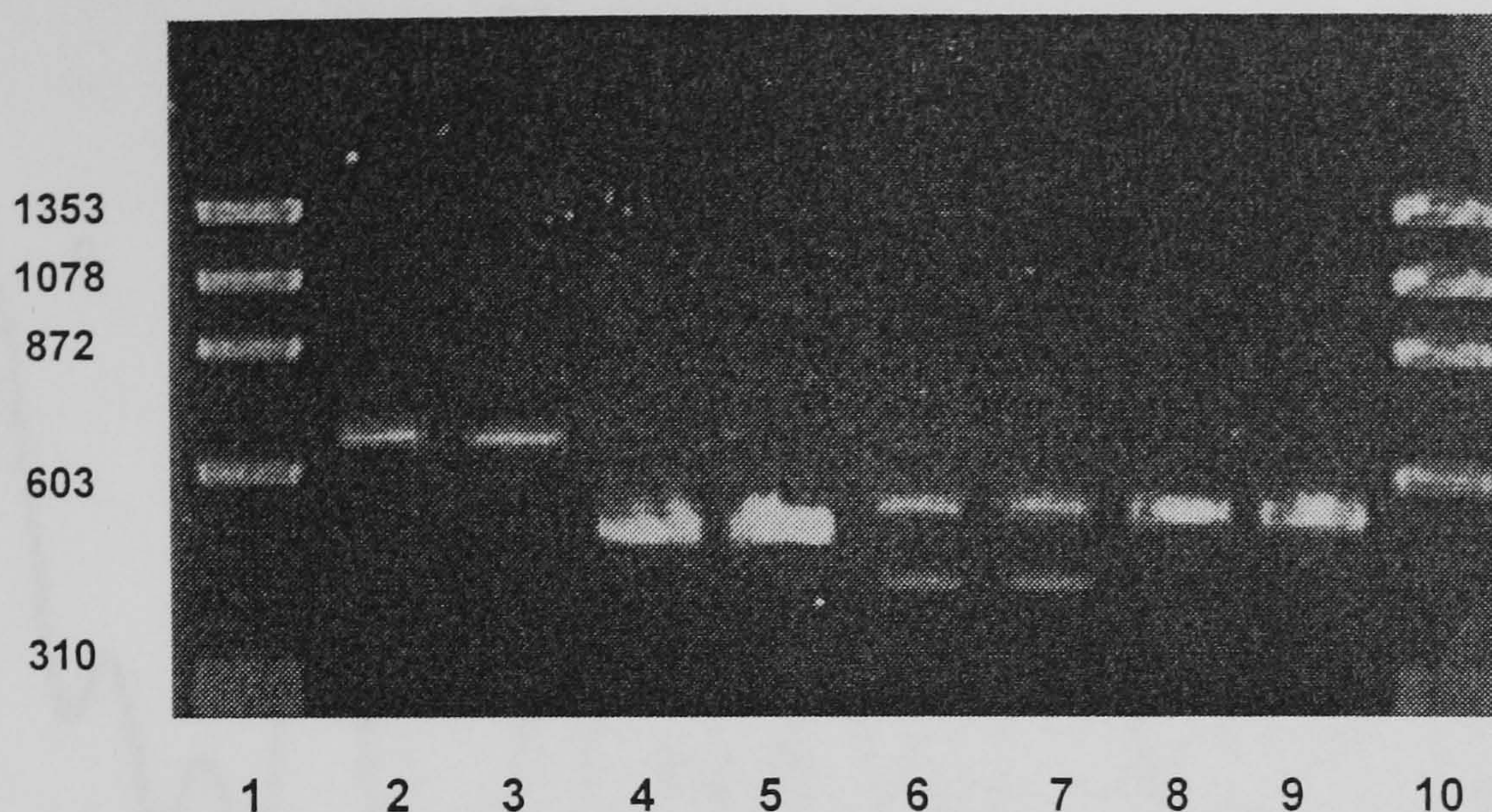


Figure 5.4. Agarose gel electrophoresis of PCR products (after purification). Lanes 1 and 10 contain the size marker ΦX174 Hae III digest (Promega). Relevant marker band sizes (bp) are given alongside the picture. Lanes 2 and 3 show the amplification of the ND1 fragment using primers N1-J-12261 and LR-N-12945 to yield a fragment of a little over 650 bp. Lanes 4 and 5 show the amplification of the 16S fragment using primers LR-J-12887 and LR-N-13398 to yield a fragment of approximately 450 bp. Lanes 6 and 7 show the amplification of the CO1 fragment using primers C1-J-1718 and C1-N-2191 for *T. parietina*. A fast (lower molecular weight) band is just visible at approximately 350 bp, in addition to the usual 470 bp fragment. Lanes 8 and 9 show the typical CO1 fragment for *T. gigantea*.

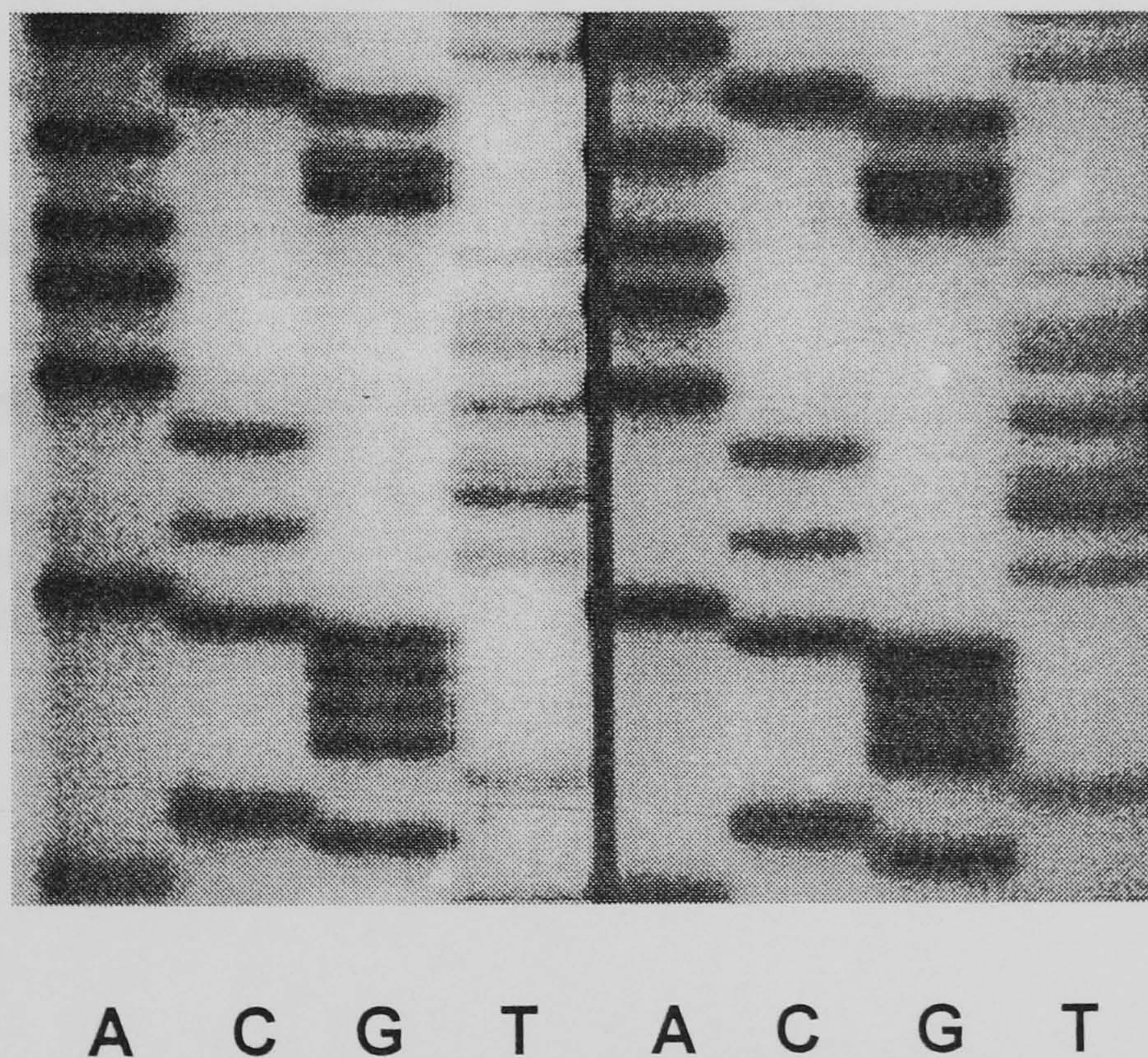


Figure 5.5. A section from a typical manual sequencing autoradiograph. Part of two CO1 sequences from *T. gigantea* generated using primer C1-J-1718. Both the sequences are identical for the region shown. There are four lanes corresponding to the four dinucleotides in the order: A, C, G, T. The sequence read from the bottom-left is: AGCTGGGGCA TCTTCTATTA TAGGAGCTA.

[Original in colour]

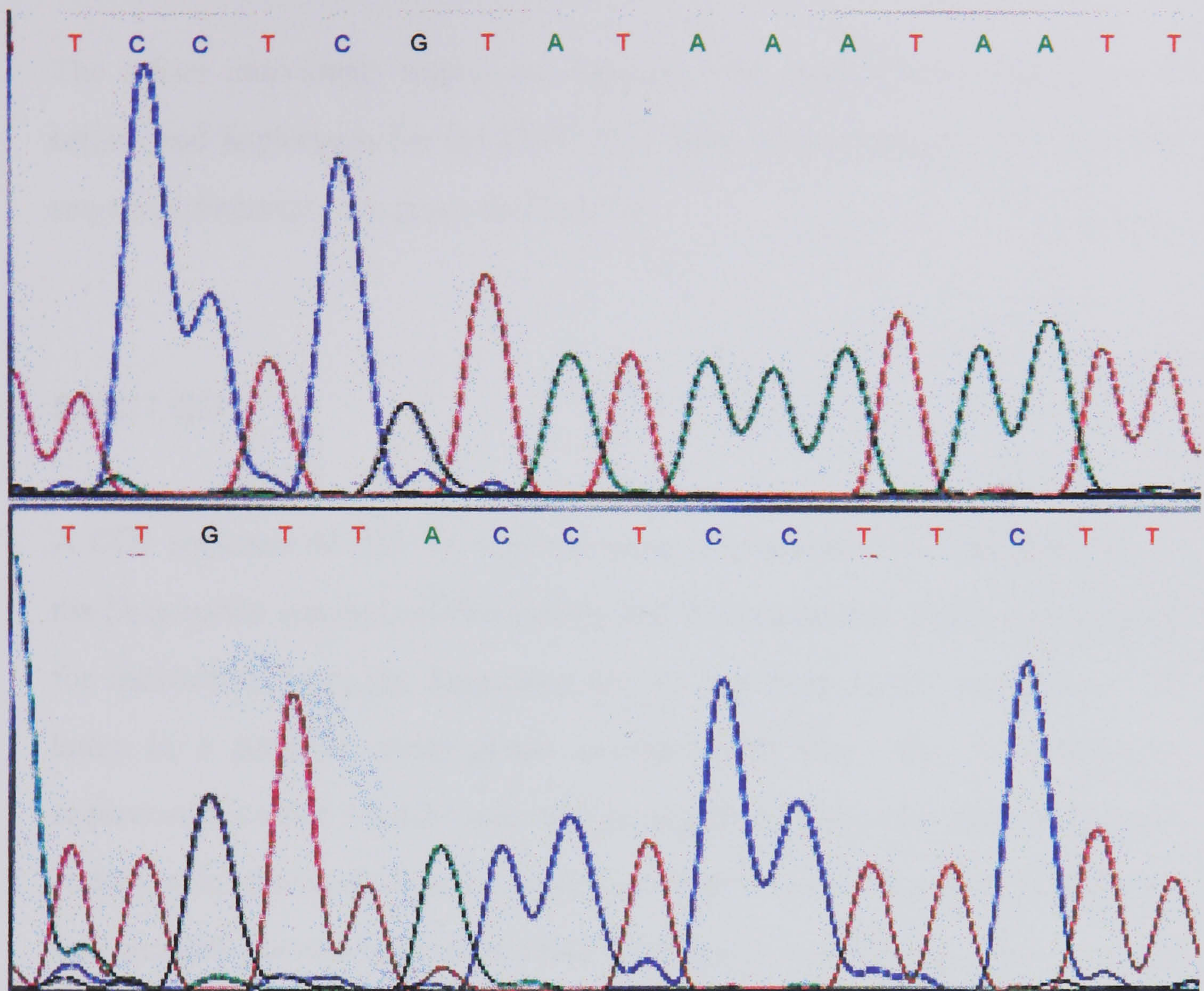


Figure 5.6. A section from a typical automated sequencer electropherogram. Part of the electropherogram output from the program FACTURA (Perkin Elmer) for a CO1 sequence from *T. gigantea* generated using primer C1-J-1718. The peaks of the trace represent peaks in flourescense for each of the four dye-terminator ddNTPs. The corresponding sequence is given above the trace.

5.3.2 Sequence Characteristics

The spider individuals sequenced, together with their species identification, origin, and haplotypes for the CO1, 16S, ND1, and combined CO1 and ND1 sequence fragments are given in Table 5.1.

5.3.2.1 CO1

A CO1 sequence of 325 bp, corresponding to positions 1777 through 2101 of the *Drosophila yakuba* mtDNA (Clary and Wolstenholme, 1985) was analysed for individuals from six *Tegenaria* species and from *Tetrix denticulata*, the latter in a putative sister genus of the Agelenidae. The 31 individuals sequenced revealed 9 haplotypes with no length variation. No stop or missense codons were detected when the sequences were translated using an invertebrate mitochondrial codon translation table. Alignment with the *D. yakuba* mtDNA sequence (Clary and Wolstenholme, 1985) and comparison with arthropod CO1 sequences deposited in EMBL suggested the sequences were from functional CO1 genes. The alignment for the CO1 sequences is given in Table 5.2. The *D. yakuba* sequence has been included in this alignment for comparison.

For the five haplotypes corresponding to the *T. atrica* group, of the 325 bp analysed, 300 (92.3%) were invariant, 19 (5.9%) were autapomorphic, and 6 (1.9%) were potentially phylogenetically informative. The numbers of substitutions at each codon position for the *T. atrica* group are given in Figure 5.7. All of the potentially phylogenetically informative sites occurred in the third position.

Table 5.1. Individuals sequenced, identification, location of origin, and haplotypes for the CO1, 16S, ND1, and combined CO1 and ND1 fragments.

		Map Reference		Haplotypes			
Species	Origin	O.S. G.R.	Coordinates	CO1	16S	ND1	Combined CO1 and ND1
Tetrix denticulata	Horton Heath, Dorset, UK	4065 1063	50°52'N 1°58'W	*T. denticulata*	-	-	-
T. agrestis	Liverpool, UK	3350 3913	53°25'N 2°55'W	*T. agrestis*	-	-	-
T. domestica	York, UK	459- 453-	53°58'N 1°05'W	*T. domestica*	-	-	-
T. domestica	Honiton, Devon, UK	3160 1000	50°48'N 3°13'W	*T. domestica*	-	-	-
T. domestica	Axmouth, Devon, UK	3256 0909	50°43'N 3°00'W	*T. domestica*	-	-	-
T. parietina	Oare, Kent, UK	5000 1630	31°18'N 00°35'W	*T. parietina*	-	-	-
T. parietina	Pavia, Italy	-	45°12'N 9°09'E	*T. parietina*	-	-	-
T. parietina	Pavia, Italy	-	45°12'N 9°09'E	*T. parietina*	-	-	-
T. atrica	S. France	-	Ca. 43°N 4°E	*T. atrica*	-	*T. atrica*	*T. atrica*
T. atrica	York, UK	4560 4533	53°58'N 1°05'E	*T. atrica*	-	-	-
T. atrica	Castleknock, Co. Dublin, Eire	-	Ca. 53°20'N 6°00'W	*T. atrica*	*T. atrica*	*T. atrica*	*T. atrica*
T. saeva	Nancy, N. France	-	48°42'N 6°12'E	*T. saeva*	*T. saeva*	-	-
T. saeva	Axmouth, Devon, UK	2865 1009	50°43'N 3°00'W	*T. saeva*	-	-	-
T. saeva	Weymouth, Dorset, UK	3659 0791	50°36'N 2°28'W	*T. saeva*	-	-	-
T. saeva	Shobrooke, Devon, UK	2865 1009	50°47'N 3°39'W	*T. saeva*	-	*T. saeva 1*	*T. saeva*
T. saeva	Nancy, N. France	-	48°42'N 6°12'E	-	-	*T. saeva 1*	-
T. saeva	Newton St. Cyres, Devon, UK	2878 0975	50°47'N 3°39'W	-	-	*T. saeva 1*	-

Continued →

Table 5.1 Continued.

		Map Reference		Haplotypes			
Species	Origin	Grid Ref.	Coordinates	CO1	16S	ND1	Combined CO1 and ND1
T. gigantea	Southwater, W. Sussex, UK	5153 1268	51°04'N 0°21'W	*T. gigantea + T. saeva 1*	-	-	-
T. saeva	Cheselbourne, Dorset, UK	3764 0996	50°44'N 2°20'W	*T. gigantea + T. saeva 1*	-	-	-
T. gigantea/saeva	Wheldrake, York, UK	4676 4447	53°58'N 1°05'W	*T. gigantea + T. saeva 1*	-	-	-
T. saeva	Tadcaster, York, UK	459- 453-	53°53'N 1°16'W	*T. gigantea + T. saeva 1*	-	-	-
T. saeva	Acomb, York, UK	459- 453-	53°58'N 1°05'W	*T. gigantea + T. saeva 1*	-	-	-
T. saeva	Church Green, Devon, UK	3172 0964	50°42'N 3°08'W	*T. gigantea + T. saeva 1*	-	-	-
T. gigantea	Southwater, W. Sussex, UK	5153 1268	50°58'N 0°24'W	*T. gigantea + T. saeva 1*	-	-	-
T. gigantea	Swanmore, Hants, UK	4583 1163	50°53'N 1°12'W	*T. gigantea + T. saeva 1*	-	-	-
T. gigantea	Horton Heath, Dorset, UK	4065 1063	50°48'N 1°55'W	*T. gigantea + T. saeva 1*	-	-	-
T. gigantea	Whitstable, Kent, UK	6109 1662	51°20'N 1°56'E	*T. gigantea + T. saeva 1*	-	*T. gigantea 2*	*T. gigantea + T. saeva 3*
T. gigantea	Woodchurch, Kent, UK	5943 1340	51°02'N 0°42'E	*T. gigantea + T. saeva 1*	*T. gigantea + T. saeva*	*T. gigantea + T. saeva*	*T. gigantea + T. saeva 4*
T. saeva	Church Green, Devon, UK	3172 0964	50°42'N 3°08'W	-	-	*T. gigantea + T. saeva*	-
T. gigantea	Dunton Green, Kent, UK	5530 1550	51°14'N 0°07'E	*T. gigantea + T. saeva 2*	-	-	-
T. saeva	Trewen, Cornwall, UK	2250 0831	50°35'N 4°25'W	*T. gigantea + T. saeva 2*	*T. gigantea + T. saeva*	*T. saeva 2*	*T. gigantea + T. saeva 1*
T. gigantea	Biddenden, Kent, UK	5850 1378	51°04'N 0°35'E	-	-	*T. gigantea 1*	-
T. gigantea	Godmersham, Kent, UK	6053 1500	51°14'N 0°48'E	*T. gigantea + T. saeva 2*	-	*T. gigantea 1*	*T. gigantea + T. saeva 2*
T. gigantea	Snohomish Co., WA., USA	-	47°55'N 122°05'W	*T. gigantea + T. saeva 2*	-	*T. gigantea 1*	*T. gigantea + T. saeva 2*
T. gigantea	Broadstone, Dorset, UK	3998 0965	50°42'N 2°01'W	*T. gigantea*	-	-	-

The species identification is given in the first column. The second column shows the location of capture. Approximate map references indicated by dashes in the grid reference. O.S. G.R. = British national Ordnance Survey 8-figure grid reference (easting first). Coordinates are latitude and longitude. The haplotypes for each gene fragment sequenced refer to the species in which the haplotype was found, for example: *T. gigantea + T. saeva 2* for CO1 was the second haplotype of CO1 found to occur in both specimens of *T. gigantea* and *T. saeva*. Haplotypes marked '-' were not sequenced. Specimens listed in order of haplotype.

Table 5.2. Alignment for the 325 bp of CO1 in the 5'→3' direction for the J strand.

	1777			1780										1790							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
D. yakuba	T	T	A	C	T	A	C	C	T	C	C	T	G	C	T	C	T	T	T	C	T
													A	*l*	*a*				*S*	*e*	*r*
T. atrica	T	T	G	T	T	A	C	C	T	C	C	T	T	C	T	C	T	A	A	T	T
	L	*e*	*u*	*L*	*e*	*u*	*P*	*r*	*o*	*P*	*r*	*o*	*S*	*e*	*r*	*L*	*e*	*u*	*I*	*l*	*e*
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
T. gigantea	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
T. agrestis	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	T	.	G	T	.	.
																			P	*h*	*e*
T. domestica	.	.	.	.	.	G	.	.	.	.	.	.	.	.	G	T	.	.	T	.	.
																			P	*h*	*e*
T. parietina	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	T	.	G	T	.	.
																			P	*h*	*e*
Tetrix denticulata	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	.
																			P	*h*	*e*

			1800										1810								
	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
D. yakuba	T	T	A	T	T	A	T	T	A	G	T	A	A	G	A	A	G	A	A	T	A
							L	*e*	*u*	*V*	*a*	*l*									
T. atrica	T	T	A	T	T	A	T	T	T	A	T	T	T	C	T	T	C	T	A	T	A
	L	*e*	*u*	*L*	*e*	*u*	*P*	*h*	*e*	*I*	*l*	*e*	*S*	*e*	*r*	*S*	*e*	*r*	*M*	*e*	*t*
T. saeva	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 1	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 2	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. agrestis	A	.	G	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	M	*e*	*t*																		
T. domestica	A	.	G	C	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
	M	*e*	*t*																		
T. parietina	A	.	G	C	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	G
	M	*e*	*t*																		
Tetrix denticulata	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	M	*e*	*t*																		

Continued →

Table 5.2. Continued

		1820										1830									
	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463
D. yakuba	G	T	T	G	A	A	A	A	C	G	G	A	G	C	T	G	G	T	A	C	A
							A	*s*	*n*				*A*	*l*	*a*				*T*	*h*	*r*
T. atrica	G	T	T	G	A	A	A	T	G	G	G	T	G	T	T	G	G	A	G	C	A
	V	*a*	*l*	*G*	*l*	*u*	*M*	*e*	*t*	*G*	*l*	*y*	*V*	*a*	*l*	*G*	*l*	*y*	*A*	*l*	*a*
T. saeva	A	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G
	I	*l*	*e*																		
T. gigantea	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 1	.	.	.	.	.	.	G	.	A	.	.	.	.	.	.	.	.	.	.	.	.
							V	*a*	*l*												
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
T. agrestis	.	.	.	.	.	G	.	.	A	.	.	.	.	.	.	.	.	G	.	.	T
T. domestica	.	.	G	.	.	T	.	.	A	.	.	G	.	.	.	.	.	T	.	.	T
				A	*s*	*p*															
T. parietina	.	.	G	.	.	T	.	.	A	.	.	A	.	.	.	.	.	.	.	.	G
				A	*s*	*p*															
Tetrix denticulata	.	.	A	.	.	T	.	.	A	.	.	A	.	.	G	.	.	.	.	.	.
				A	*s*	*p*															

	1840										1850										1860
	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684
D. yakuba	G	G	T	T	G	A	A	C	T	G	T	T	T	A	C	C	C	T	C	C	T
										V	*a*	*l*									
T. atrica	G	G	A	T	G	A	A	C	T	A	T	T	T	A	T	C	C	T	C	C	T
	G	*l*	*y*	*T*	*r*	*p*	*T*	*h*	*r*	*I*	*l*	*e*	*T*	*y*	*r*	*P*	*r*	*o*	*P*	*r*	*o*
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
T. gigantea	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
T. gigantea + T. saeva 1	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
T. gigantea + T. saeva 2	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
T. agrestis	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. domestica	.	.	T	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
T. parietina	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Tetrix denticulata	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Continued →

Table 5.2. Continued

										1870										1880	
	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105
D. yakuba	T	T	A	T	C	T	T	C	A	G	G	T	A	T	C	G	C	T	C	A	T
				S	*e*	*r*				*G*	*l*	*y*				*A*	*l*	*a*			
T. atrica	T	T	G	G	C	T	T	C	T	T	C	T	A	T	T	G	G	T	C	A	T
	L	*e*	*u*	*A*	*l*	*a*	*S*	*e*	*r*	*S*	*e*	*r*	*I*	*l*	*e*	*G*	*l*	*y*	*H*	*i*	*s*
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. agrestis	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. domestica	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.
													M	*e*	*t*						
T. parietina	.	.	.	.	.	G	.	.	.	.	.	.	C	.	.	.	.	G	.	.	C
													L	*e*	*u*						
Tetrix denticulata	.	.	A	.	.	A	.	.	.	.	.	.	C	.	.	.	.	.	.	.	C
													L	*e*	*u*						

										1890										1900		
	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126
D. yakuba	G	G	T	G	G	A	G	C	T	T	C	T	G	T	A	G	A	T	T	T	A
	G	*l*	*y*				*A*	*l*	*a*										*L*	*e*	*u*
T. atrica	T	T	T	G	G	T	A	A	A	T	C	T	G	T	T	G	A	T	T	T	T
	P	*h*	*e*	*G*	*l*	*y*	*L*	*y*	*s*	*S*	*e*	*r*	*V*	*a*	*l*	*A*	*s*	*p*	*P*	*h*	*e*
T. saeva	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.
							S	*e*	*r*												
T. gigantea	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.
							S	*e*	*r*												
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.
							S	*e*	*r*												
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.
							S	*e*	*r*												
T. agrestis	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.
							S	*e*	*r*												
T. domestica	A	.	A	.	.	G	.	G	.	G	.	.	A	.	A	.	.	.	.	.	.
	M	*e*	*t*				*S*	*e*	*r*	*A*	*l*	*a*	*M*	*e*	*t*						
T. parietina	A	.	A	.	.	A	.	G	T	.	.	.	A	.	G	.	.	.	.	.	.
	M	*e*	*t*				*S*	*e*	*r*				*M*	*e*	*t*						
Tetrix denticulata	A	.	A	.	.	A	.	G	.	.	.	.	A	.	G	.	.	.	.	.	.
	M	*e*	*t*				*S*	*e*	*r*				*M*	*e*	*t*						

Continued →

Table 5.2. Continued

								1910										1920			
	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147
D. yakuba	**G**	**C**	**T**	**A**	**T**	**T**	**T**	**T**	**T**	**T**	**C**	**T**	**C**	**T**	**T**	**C**	**A**	**T**	**T**	**T**	**A**
T. atrica	**G**	**C**	**T**	**A**	**T**	**T**	**T**	**T**	**T**	**T**	**C**	**T**	**T**	**T**	**G**	**C**	**A**	**T**	**T**	**T**	**G**
	A	*l*	*a*	*I*	*l*	*e*	*P*	*h*	*e*	*S*	*e*	*r*	*L*	*e*	*u*	*H*	*i*	*s*	*L*	*e*	*u*
T. saeva	.	**T**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
	V	*a*	*l*																		
T. gigantea	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	**A**
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	**A**
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	**A**
T. agrestis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. domestica	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
T. parietina	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	**A**
Tetrix denticulata	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	**A**

							1930										1940				
	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168
D. yakuba	**G**	**C**	**T**	**G**	**G**	**A**	**A**	**T**	**T**	**T**	**C**	**T**	**T**	**C**	**A**	**A**	**T**	**T**	**T**	**T**	**A**
							I	*l*	*e*										*L*	*e*	*u*
T. atrica	**G**	**C**	**T**	**G**	**G**	**A**	**G**	**C**	**A**	**T**	**C**	**T**	**T**	**C**	**T**	**A**	**T**	**T**	**A**	**T**	**G**
	A	*l*	*a*	*G*	*l*	*y*	*A*	*l*	*a*	*S*	*e*	*r*	*S*	*e*	*r*	*I*	*l*	*e*	*M*	*e*	*t*
T. saeva	.	.	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
T. gigantea	.	.	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
T. gigantea + T. saeva 1	.	.	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
T. gigantea + T. saeva 2	.	.	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
T. agrestis	.	.	.	.	.	**G**	.	.	**T**	.	.	.	.	.	.	.	.	.	.	.	**A**
T. domestica	.	.	**A**	.	.	**G**	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.
T. parietina	.	.	.	.	.	**G**	.	.	**T**	.	.	.	.	.	**A**	.	.	.	.	.	**A**
Tetrix denticulata	.	.	.	.	.	**G**	.	.	**T**	.	.	.	.	.	**A**	.	.	.	.	.	**A**

Continued →

Table 5.2. Continued

	1669	1670	1671	1672	1673	1674 (1950)	1675	1676	1677	1678	1679	1680	1681	1682	1683	1684 (1960)	1685	1686	1687	1688	1689
D. yakuba	G	G	A	G	C	T	G	T	A	A	A	T	T	T	T	A	T	T	A	C	G
							V	a	l										T	h	r
T. atrica	G	G	A	G	C	T	A	T	T	A	A	T	T	T	T	A	T	T	T	C	T
	G	l	y	A	l	a	I	l	e	A	s	n	P	h	e	I	l	e	S	e	r
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 2	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. agrestis	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. domestica	.	.	.	.	.	G	G	.	G	.	.	.	.	.	.	.	.	.	A	.	.
							V	a	l										T	h	r
T. parietina	.	.	G	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	C
Tetrix denticulata	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	1990	1991	1992	1993	1994 (1970)	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004 (1980)	2005	2006	2007	2008	2009	2010
D. yakuba	A	C	T	G	T	A	A	T	T	A	A	T	A	T	A	C	G	A	T	C	A
							I	l	e												
T. atrica	A	C	T	G	T	T	T	T	G	A	A	T	A	T	A	C	G	A	T	C	T
	T	h	r	V	a	l	L	e	u	A	s	n	M	e	t	A	r	g	S	e	r
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.
T. agrestis	.	.	.	A	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
				I	l	e	P	h	e												
T. domestica	.	.	G	A	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	.
				I	l	e															
T. parietina	.	.	A	A	.	.	A	.	T	.	.	.	.	.	.	.	.	T	.	.	G
				I	l	e	I	l	e												
Tetrix denticulata	.	.	.	A	.	.	.	.	A	.	.	.	.	.	G	.	.	T	.	.	A
				I	l	e															

Continued →

Table 5.2. Continued

				1990										2000							
	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131
D. yakuba	**A**	**C**	**T**	**G**	**G**	**A**	**A**	**T**	**T**	**A**	**C**	**A**	**T**	**T**	**A**	**G**	**A**	**C**	**C**	**G**	**A**
	T	*h*	*r*	*G*	*l*	*y*	*I*	*l*	*e*				*L*	*e*	*u*	*A*	*s*	*p*	*A*	*r*	*g*
T. atrica	**T**	**T**	**A**	**A**	**G**	**A**	**A**	**T**	**A**	**A**	**C**	**T**	**A**	**T**	**A**	**G**	**A**	**G**	**A**	**A**	**A**
	L	*e*	*u*	*S*	*e*	*r*	*M*	*e*	*t*	*T*	*h*	*r*	*M*	*e*	*t*	*G*	*l*	*u*	*L*	*y*	*s*
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.
T. gigantea	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.
T. agrestis	**G**	.	**T**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**G**
	V	*a*	*l*																		
T. domestica	**G**	.	.	**G**	.	**G**	.	.	.	.	**G**	**A**	.	.	.	.	.	**T**	.	**G**	**G**
	V	*a*	*l*	*G*	*l*	*y*				*S*	*e*	*r*				*A*	*s*	*p*	*S*	*e*	*r*
T. parietina	**A**	.	**C**	**G**	.	.	.	.	.	.	**G**	**A**	.	.	.	.	.	.	.	.	**G**
	I	*l*	*e*	*G*	*l*	*y*				*S*	*e*	*r*									
Tetrix denticulata	**A**	.	**T**	**G**	.	**G**	.	.	.	.	**A**	**G**	.	.	.	.	.	.	.	**G**	**G**
	I	*l*	*e*	*G*	*l*	*y*				*L*	*y*	*s*							*S*	*e*	*r*

			2010										2020								
	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252
D. yakuba	**A**	**T**	**A**	**C**	**C**	**T**	**T**	**T**	**A**	**T**	**T**	**T**	**G**	**T**	**A**	**T**	**G**	**A**	**T**	**C**	**A**
	M	*e*	*t*																		
T. atrica	**G**	**T**	**T**	**C**	**C**	**T**	**T**	**T**	**G**	**T**	**T**	**T**	**G**	**T**	**G**	**T**	**G**	**A**	**T**	**C**	**T**
	V	*a*	*l*	*P*	*r*	*o*	*L*	*e*	*u*	*P*	*h*	*e*	*V*	*a*	*l*	*T*	*r*	*p*	*S*	*e*	*r*
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	.
T. gigantea	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	.
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	.
T. agrestis	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	**A**	.	.	.	.	.	.
T. domestica	.	.	.	**T**	.	.	.	.	.	.	.	.	.	.	**T**	.	.	**G**	.	.	**G**
				S	*e*	*r*															
T. parietina	.	.	.	.	.	.	.	.	**A**	.	.	.	.	.	**T**	.	.	.	.	.	.
Tetrix denticulata	.	.	.	.	.	**G**	.	.	.	.	.	.	.	.	**T**	.	.	.	.	.	**A**

Continued →

Table 5.2. Continued

		2030										2040									
	2553	2554	2555	2556	2557	2558	2559	2560	2561	2562	2563	2564	2565	2566	2567	2568	2569	2570	2571	2572	2573
D. yakuba	G	T	A	G	T	T	A	T	T	A	C	T	G	C	T	T	T	A	T	T	A
				V	*a*	*l*										*L*	*e*	*u*			
T. atrica	G	T	T	T	T	A	A	T	T	A	C	T	G	C	T	G	T	T	T	T	A
	V	*a*	*l*	*L*	*e*	*u*	*I*	*l*	*e*	*T*	*h*	*r*	*A*	*l*	*a*	*V*	*a*	*l*	*L*	*e*	*u*
T. saeva	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.
																I	*l*	*e*			
T. gigantea	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
																I	*l*	*e*			
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
																I	*l*	*e*			
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
																I	*l*	*e*			
T. agrestis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
T. domestica	.	.	A	.	.	G	G	.	.	.	.	C	.	.	.	A	.	.	.	.	.
							V	*a*	*l*							*I*	*l*	*e*			
T. parietina	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	G
																I	*l*	*e*			
Tetrix denticulata	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
																I	*l*	*e*			

	2050										2060										2070
	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294
D. yakuba	C	T	T	T	T	A	C	T	A	T	C	T	T	T	A	C	C	A	G	T	T
T. atrica	T	T	A	T	T	A	T	T	G	T	C	T	T	T	A	C	C	T	G	T	A
	L	*e*	*u*	*L*	*e*	*u*	*L*	*e*	*u*	*S*	*e*	*r*	*L*	*e*	*u*	*P*	*r*	*o*	*V*	*a*	*l*
T. saeva	.	.	G	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea	.	.	G	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 1	.	.	G	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 2	.	.	G	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
T. agrestis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. domestica	.	.	G	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.
T. parietina	.	.	.	.	.	.	.	.	A	.	.	A	C	.	T	.	.	A	.	.	T
Tetrix denticulata	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.

Continued →

Table 5.2. Continued

	2995	2996	2997	2998	2999	3000	3001	3002	3003	3004	3005	3006	3007	3008	3009	3010	3011	3012	3013	3014	3015
										2080										2090	
D. yakuba	C	T	T	G	C	C	G	G	A	G	C	T	A	T	T	A	C	T	A	T	A
T. atrica	T	T	A	G	C	A	G	G	T	G	C	T	A	T	T	A	C	A	A	T	A
	L	*e*	*u*	*A*	*l*	*a*	*G*	*l*	*y*	*A*	*l*	*a*	*I*	*l*	*e*	*T*	*h*	*r*	*M*	*e*	*t*
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
T. gigantea + T. saeva 2	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
T. agrestis	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
T. domestica	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
T. parietina	.	.	G	.	.	T	.	.	A	.	.	G	.	.	.	.	.	G	.	.	G
Tetrix denticulata	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.

	3316	3317	3318	3319	3320	3321	3322	3323	3324	3325
									2100	2101
D. yakuba	T	T	A	T	T	A	A	C	A	G
T. atrica	T	T	G	T	T	A	A	C	T	G
	L	*e*	*u*	*L*	*e*	*u*	*T*	*h*	*r*	
T. saeva	.	.	.	.	.	.	.	.	.	.
T. gigantea	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 1	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva 2	.	.	.	.	.	.	.	.	.	.
T. agrestis	.	.	A	.	.	.	.	.	.	.
T. domestica	.	.	.	.	.	G	.	.	G	.
T. parietina	.	.	.	.	.	G	.	.	.	.
Tetrix denticulata	.	.	.	.	.	.	.	.	.	.

T. atrica acts as a consensus sequence for both nucleotide and amino acid sequence. Nucleotide identity is shown by a dot. All *Tegenaria* spp. are included along with the outgroup *Tetrix denticulata*. The *Drosphila yakuba* sequence (Clary and Wolstenholme, 1985) is given first for comparison (amino acids identical to *T. atrica* are not shown). Boxed numbers above the alignment refer to the equivalent position in the *D. yakuba* mtDNA molecule. Nucleotides are numbered consecutively from 1 to 325.

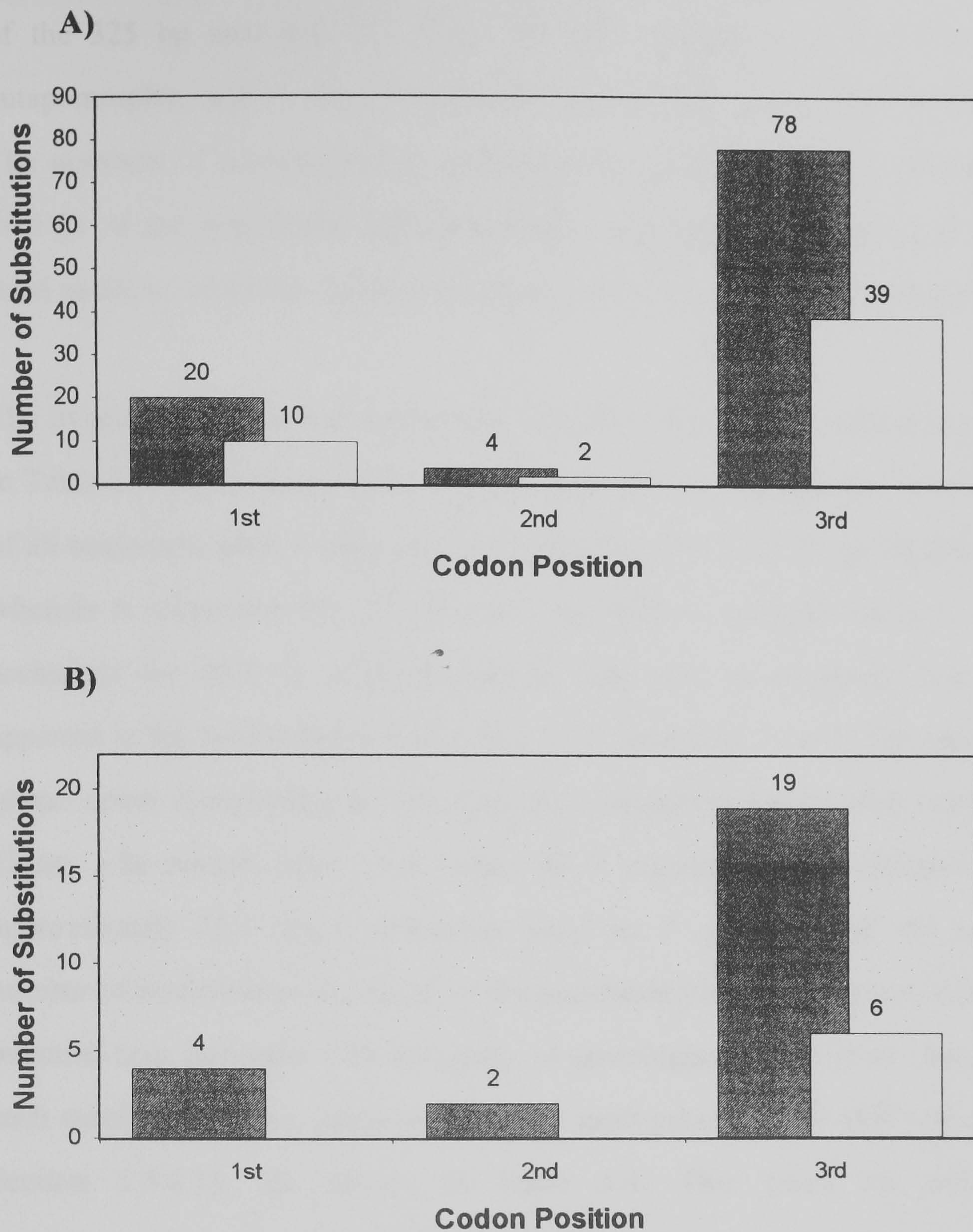


Figure 5.7 Distribution of substitutions among the three codon positions in the mt CO1 fragment. A) All *Tegenaria* species plus the outgroup *Tetrix denticulata*. B) Members of the *T. atrica* group only. Shaded boxes = total substitutions. Unshaded boxes = potentially phylogenetically informative substitutions.

For all nine haplotypes (all *Tegenaria* species plus the outgroup *T. denticulata*), of the 325 bp analysed, 223 (68.6 %) were invariant, 51 (15.7 %) were autapomorphic, and 51 (15.7 %) were potentially phylogenetically informative. The numbers of substitutions at each codon position are given in Figure 5.7. 76.5 % of the potentially phylogenetically informative sites occurred in the third position, 19.6 % in the first position, and 3.9 % in the second position.

The frequencies of the four nucleotides in each of the codon positions are given in Table 5.3. C was particularly under-represented, accounting for only 13.8 % of all nucleotide sites. G also only accounted for 17.7 % of all nucleotide sites, whereas A accounted for 21.7 % and T for 46.8 %. Overall, the A + T bias accounted for 68.5 % of all nucleotide sites and, as expected, was most apparent in the third position where 84.6 % of sites were A or T. The empirical (direct count disregarding polymorphic sites) transition:transversion ratios (see Figure 5.8) ranged from 1.9:1, when all 9 sequences were considered, to approximately 25:1 when considering only the *T. atrica* group: the relative number of tranversions increased as the maximum pairwise divergence among included taxa increased. The frequency of unambiguous nucleotide changes in each substitution class, estimated from the most parsimonious (MP) trees (see Section 5.3.4.1), are shown in Table 5.4. This gives an estimated transition:transversion ratio of 1.4:1 over all taxa. It can also be seen that A $\rightarrow$ G and T $\rightarrow$ C transitions occurred 2.4 times more frequently than their reciprocal substitutions, and that A $\leftrightarrow$ T accounted for approximately 58% of all transversions (with T $\rightarrow$ G accounting for approximately another 30%).

Figure 5.9 shows the two dimensional structural model of CO1 for a wide taxonomic range of insects as described by Lunt *et al.* (1996). Superimposed onto this figure are the variable amino acid positions for the spider species

Table 5.3. Nucleotide frequencies at each codon position for the CO1 fragment.

	Codon Position		
	First	Second	Third
A	27.12	8.33	29.53
C	8.77	31.07	1.65
G	26.91	12.35	13.79
T	37.21	48.25	55.04

The overall frequency of each nucleotide at each codon position is given for the 9 haplotypes and 325 bp of CO1 analysed.

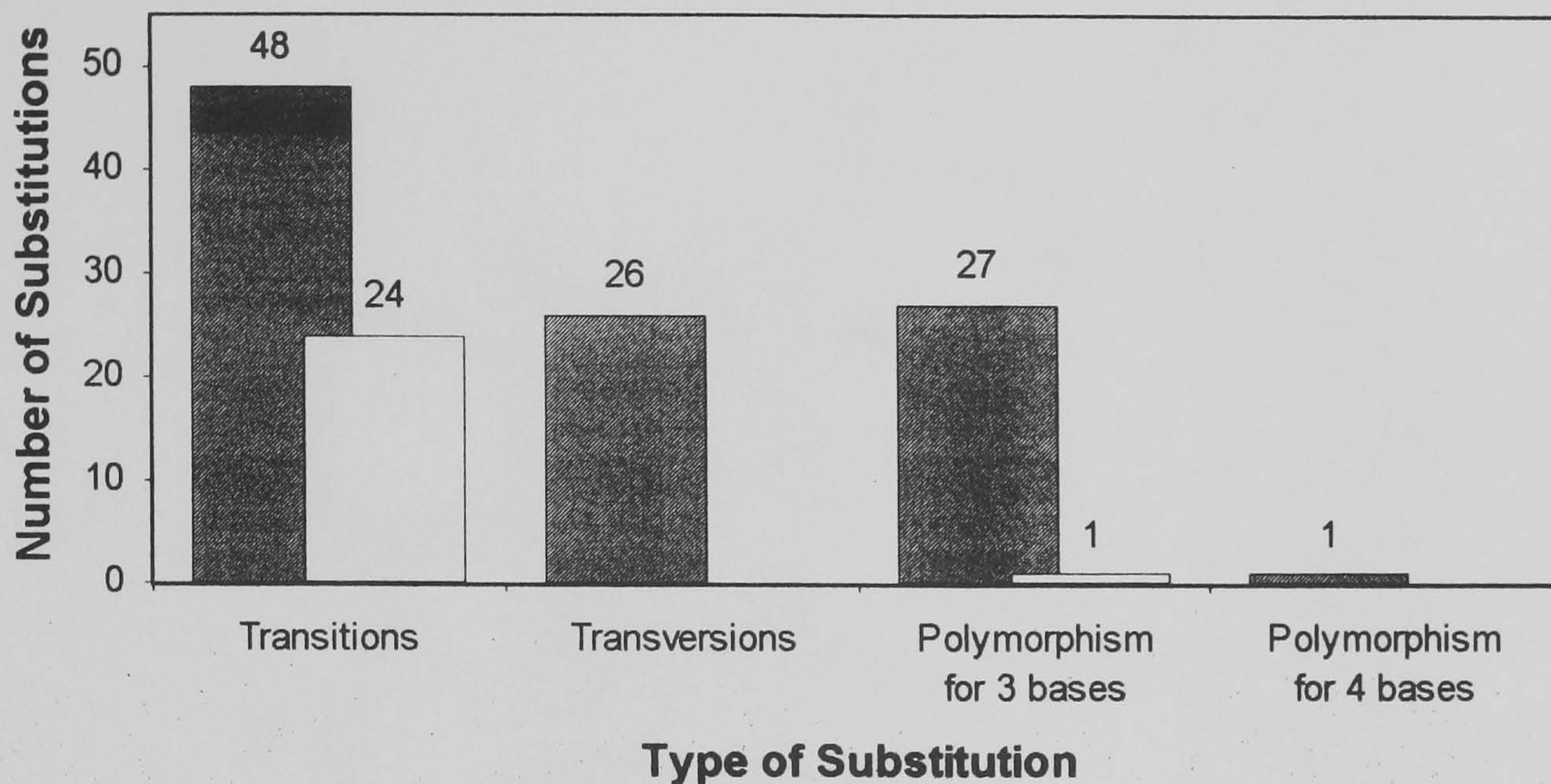


Figure 5.8. Number of substitutions of each class in the mt CO1 fragment. Shaded boxes = all *Tegenaria* species plus the outgroup *Tetrix denticulata.* Unshaded boxes = members of the *T. atrica* group only.

Table 5.4. The frequency of unambiguous nucleotide changes in the single MP tree for CO1.

	To			
From	A	C	G	T
A	-	0.000	0.327	0.115
C	0.010	-	0.000	0.029
G	0.144	0.010	-	0.029
T	0.125	0.087	0.125	-

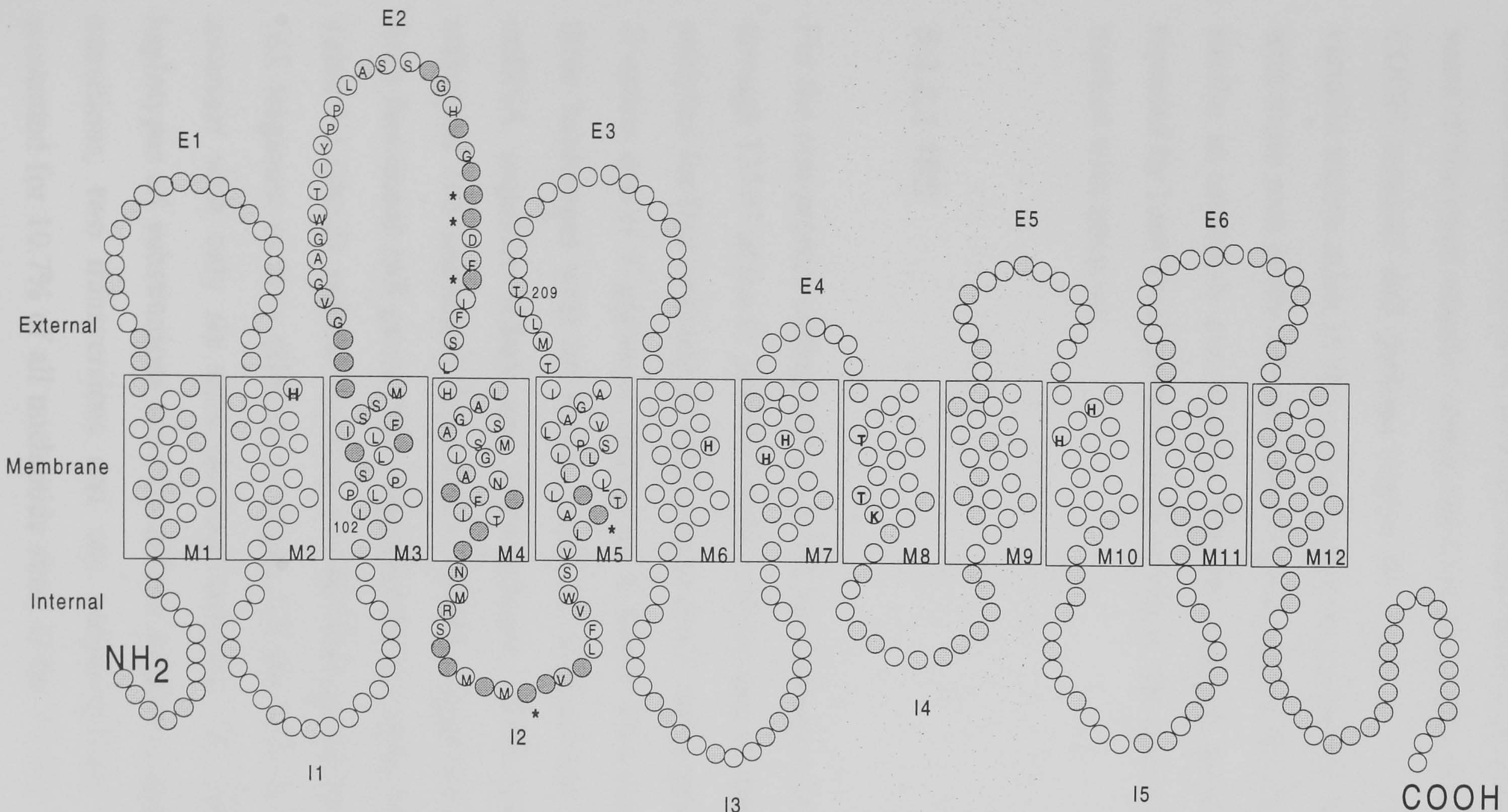


Figure 5.8. A two-dimensional model of the invertebrate (insect) CO1 gene. (Redrawn from Lunt *et al.*, 1996). The 25 structural domains are shown and the conserved amino acid sequence, from position 102 through 209 (marked), for all the *Tegenaria* species examined in this study plus *Tetrix denticulata* (see Table 5.2). The insect sequence has not been shown with the exception of those residues outside the range 102 through 209, which are of functional significance. The lightly shaded residues were those found to be variable across insect taxa by Lunt *et al.* (1996). The heavily shaded residues were those found to be variable in the spiders examined in this study. All residues variable in spiders were also variable in insects apart from those marked *.

presented in the alignment in Table 5.2 (six species of *Tegenaria* plus *Tetrix denticulata*). The 325 bp of CO1 sequence examined correspond to amino acid positions 102 through 209 of this structure. This region covers three membrane spanning domains (M3, M4 and M5), one large external loop (E2), part of another external loop (E3), and one entire internal loop (I2). These regions are some of the most variable within the CO1 molecule; with the exception of the COOH terminal and perhaps region I4 (Lunt *et al.*, 1996). The pattern of variable amino acids in the spider sequences presented here corresponds well with those seen in the insects suggesting that the evolutionary pattern of CO1 is similar in arachnids and insects. The few variable amino acid positions not reported by Lunt *et al.* (1996) but observed in the spider sequences have been marked with asterixes.

5.3.2.2 16S

For the non-protein coding 16S gene 438 bp, corresponding to positions 12888 through 13397 of the *D. yakuba* mtDNA (Clary and Wolstenhome, 1985), were analysed for four individuals belonging to the *T. atrica* group: one specimen of *T. atrica*, one of *T. gigantea*, and two of *T. saeva*. This limited sample revealed three haplotypes with no length variation. Comparison with the *D. yakuba* mtDNA sequence (Clary and Wolstenhome, 1985) and comparison with arthropod 16S sequences deposited in EMBL suggested the sequences were from functional 16S genes. The alignment for the three haplotypes is given in Table 5.5 (the *D. yakuba* sequence has not been given for comparison for the 16S fragment as it is difficult to align). Of the 438 bp analysed, 432 were invariant with only six sites showing variation. As there were only three haplotypes all substitutions were by default autapomorphic. There were three transitions, two transversions and one polymorphism for three bases. C accounted for 10.7% of all nucleotide sites, G for 11.0%, T for 32.3%, and A

Table 5.5. Alignment for the 438 bp of 16S in the 5'→3' direction for the N strand.

	13397																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
T. atrica	T	T	T	T	A	A	T	A	G	A	A	A	T	A	A	T	T	A	T	T	A	G	T	A	A
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea +*T. saeva*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
T. atrica	A	A	T	C	T	G	C	T	C	A	A	T	G	A	A	T	A	A	T	A	A	T	T	C	A
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + *T. saeva*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
T. atrica	A	T	A	G	C	C	G	C	A	A	T	T	A	A	T	T	T	G	T	G	C	T	A	A	G
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + *T. saeva*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
T. atrica	G	T	A	G	C	A	T	A	A	T	C	A	T	T	T	G	T	C	T	T	T	T	A	A	T
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + *T. saeva*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125
T. atrica	T	A	A	A	G	A	C	T	A	G	A	A	C	G	A	A	A	G	A	T	T	T	A	A	C
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + *T. saeva*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
T. atrica	A	A	T	T	A	A	A	T	T	A	C	T	A	T	T	T	A	T	A	A	T	A	T	A	A
T. saeva	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + *T. saeva*	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Continued →

Table 5.5. Continued

	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
T. atrica	A	A	A	T	T	A	T	A	A	T	T	T	A	T	C	C	T	A	A	A	T	A	T	A	A
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
T. atrica	A	A	A	G	A	T	A	T	T	T	A	T	A	T	A	T	A	A	A	A	A	A	G	A	C
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.

	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225
T. atrica	G	A	T	A	A	G	A	C	C	C	T	A	T	T	G	A	A	C	T	T	A	A	C	T	T
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250
T. atrica	A	A	G	T	T	T	A	A	C	T	G	G	G	G	A	A	G	T	T	A	A	T	C	A	A
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275
T. atrica	A	C	A	A	A	A	T	T	T	T	A	A	T	T	A	A	A	A	T	A	A	T	A	A	A
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.

	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300
T. atrica	A	A	A	A	A	A	G	A	T	C	T	A	A	T	A	A	A	A	T	T	A	A	T	T	A
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325
T. atrica	A	A	T	G	A	T	C	A	A	G	T	T	A	C	C	A	T	A	G	G	G	A	T	A	A
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Continued →

Table 5.5. Continued

	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350
T. atrica	C	A	G	C	G	C	A	A	T	A	A	T	T	T	C	T	T	A	A	A	G	A	T	C	T
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.

	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375
T. atrica	T	A	T	T	A	A	A	G	A	A	A	A	A	G	A	T	T	A	C	G	A	C	C	T	C
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400
T. atrica	G	A	T	G	T	T	G	A	A	T	T	A	A	T	A	A	C	C	T	A	T	T	A	A	A
	L	*R*	-	*N*	-	*1*	*2*	*9*	*4*	*5*	→														
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425
T. atrica	T	G	C	A	A	T	A	A	T	T	T	A	A	A	A	A	G	G	A	A	G	T	C	T	G
T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

													12888
	426	427	428	429	430	431	432	433	434	435	436	437	438
T. atrica	T	T	C	G	A	C	T	T	A	C	A	A	T
T. saeva	.	.	.	.	.	.	.	.	.	T	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	T	.	.	.

T. atrica acts as a consensus sequence for nucleotide sequence. Nucleotide identity is shown by a dot. Boxed numbers above the alignment refer to the equivalent position in the *D. yakuba* mtDNA molecule (Clary and Wolstenholme, 1985). Nucleotides are numbered consecutively from 1 to 438. The position of primer LR-N-12945, overlapping with the ND1 fragment, is marked.

for 46.0%, giving an overall A + T bias accounting for 78.3% of all nucleotide sites. The empirical transition:transversion ratio was 1.5:1.

5.3.2.3 ND1

An ND1 sequence of 543 bp, corresponding to positions 12295 through 12913 of the *D. yakuba* mtDNA (Clary and Wolstenholme, 1985) was analysed for individuals of the three species in the *T. atrica* group. The 12 individuals sequenced revealed six haplotypes with no length variation. The sequences consisted of approximately 182 bp of 16S gene and tRNA *leu* (CUN) with the remainder coding for ND1. No stop or missense codons were detected in the coding region when the sequences were translated using an invertebrate mitochondrial codon translation table. Alignment with the *D. yakuba* mtDNA sequence (Clary and Wolstenhome, 1985) and comparison with arthropod ND1 sequences deposited in EMBL suggested the sequences were from functional ND1 genes. The alignment for the six haplotypes is given in Table 5.6. The *D. yakuba* sequence has been included in this alignment for comparison. Relative to *D. yakuba* the spiders have large deletions flanking both sides of the tRNA leu^{CUN} gene. This results in approximately the first 30 bp of the ND1 gene in *D. yakuba* being absent in the spiders and a putative start codon for the spiders 30 bp downstream relative to *D. yakuba* (see Table 5.6).

For the six haplotypes of the *T. atrica* group, of the 543 bp analysed, 518 (95.4%) were invariant, 15 (2.8%) were autapomorphic, and 10 (1.8%) were potentially phylogenetically informative. The numbers of substitutions at each codon position (for the protein coding region), and of each class (across the entire fragment) and are given in Figure 5.10 and Figure 5.11 respectively. Within the protein coding region, 66.7% of potentially phylogenetically informative sites occurred in the third position, 22.2% in the second position,

Table 5.6. Alignment for the 543 bp of ND1 in the 5'→3' direction for the N strand.

	12913 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
D. yakuba	T	A	A	G	T	C	T	G	T	T	C	G	A	C	T	T	T	T	A	A	A	T	T	C	T
T. atrica	G	A	A	G	T	C	T	G	T	T	C	G	A	C	T	T	A	C	A	A	T	A	A	A	T
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.

	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
D. yakuba	T	A	C	A	T	G	A	T	C	T	G	A	G	T	T	C	A	A	A	C	C	G	G	T	G
T. atrica	T	A	C	A	T	G	A	T	T	T	G	A	G	T	T	C	A	G	A	C	C	G	G	T	A
		← L R - J - 1 2 8 8 7																							
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
D. yakuba	T	A	A	G	C	C	A	G	G	T	T	G	G	T	T	T	C	T	A	T	C	T	T	T	A
T. atrica	T	A	A	G	C	C	A	G	G	T	C	G	G	C	T	T	C	T	A	T	C	T	T	T	T
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Continued →

Table 5.6. Continued

	7	7	7			7	8	8	8	8	8	8	8	8	8	8	9	9	9	9	9	9	9	9	9
	6	7	8			9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8
D. yakuba	A	A	A	A	A	T	T	A	T	A	A	T	A	T	T	T	T	A	G	T	A	C	G	A	A
T. atrica	A	A	A	-	-	T	C	A	A	A	T	C	T	T	T	T	C	A	G	T	A	C	G	A	A
T. saeva 1	.	.	.	-	-	.	G	.	.	T	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	-	-	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	-	-	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	-	-	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	-	-	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.

	9	1	1	1	1	1	1	1	1																
	9	0	0	0	0	0	0	0	0																
		0	1	2	3	4	5	6	7																
D. yakuba	A	G	G	A	C	C	A	A	A	T	A	T	T	A	A	A	A	T	A	A	T	T	A	T	A
T. atrica	A	G	G	A	C	C	A	A	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. saeva 1	.	.	.	.	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. saeva 2	.	.	.	.	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. gigantea 1	.	.	.	.	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. gigantea 2	.	.	.	.	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

																								1	1
																								0	0
																								8	9
D. yakuba	T	T	T	T	T	A	T	A	T	A	A	G	A	A	T	A	T	T	A	T	T	A	A	T	A
T. atrica	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	G
T. saeva 1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.	.
T. saeva 2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.	.
T. gigantea 1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.	.
T. gigantea 2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.	.
T. gigantea + T. saeva	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.	.

	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3
	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
D. yakuba	T	A	T	A	A	T	A	A	A	C	T	A	T	T	T	T	G	G	C	A	G	A	T	T	A
T. atrica	A	A	A	T	T	T	T	A	A	T	A	A	A	T	T	T	A	G	C	A	A	A	T	A	A
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.

Continued →

Table 5.6. Continued

	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156
D. yakuba	G	T	G	C	A	A	T	A	A	A	T	T	T	A	G	A	A	T	T	T	A	T
T. atrica	A	T	G	C	A	T	T	A	G	A	A	T	T	A	G	A	A	T	C	T	A	A
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177
D. yakuba	A	T	A	T	G	T	A	A	T	T	T	T	T	A	T	T	A	C	A	A	A
T. atrica	A	A	A	T	A	T	C	A	A	A	A	T	A	T	T	A	T	C	A	A	T
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
												←	*t*	*R*	*N*	*A*		*L* *C*	*e* *U*	*u* *N*	

	1178	1179	1180	1181	1182											12680					
D. yakuba	T	A	G	T	A	C	T	T	G	T	T	T	T	A	T	A	T	A	G	A	A
																M	*e*	*t*	*G*	*l*	*u*
T. atrica	A	A	A	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. saeva 1	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. saeva 2	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. gigantea 1	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. gigantea 2	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. gigantea + T. saeva	.	.	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
																N	*D*	*1*	→		
												D	*r*	*o*	*s*	*o*	*p*	*h*	*i*	*l*	*a*

Continued →

Table 5.6. Continued

D. yakuba	T	T	T	A	T	T	T	T	A	T	C	A	T	T	A	A	T	T	G	G	A
	P	*h*	*e*	*I*	*I*	*e*	*L*	*e*	*u*	*S*	*e*	*r*	*L*	*e*	*u*	*I*	*I*	*e*	*G*	*l*	*y*
T. atrica	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. saeva 1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. saeva 2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. gigantea 1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. gigantea 2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T. gigantea + T. saeva	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

		183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202
D. yakuba	A	G	T	T	T	A	T	T	A	T	T	A	A	T	T	A	T	T	T	G	T
	S	*e*	*r*	*L*	*e*	*u*	*L*	*e*	*u*	*L*	*e*	*u*	*I*	*I*	*e*				*C*	*y*	*s*
T. atrica	-	T	T	A	T	C	A	T	A	T	C	C	T	C	A	A	T	T	A	G	T
			I	*I*	*e*	*M*	*e*	*t*	*S*	*e*	*r*	*S*	*e*	*r*	*I*	*I*	*e*	*S*	*e*	*r*	
T. saeva 1	-	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	-	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	-	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	-	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	-	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
			?	*N*	*D*	*1*	→														
			S	*p*	*i*	*d*	*e*	*r*													

	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
D. yakuba	G	T	A	T	T	A	G	T	A	A	G	T	G	T	A	G	C	T	T	T	T
	V	*a*	*l*				*V*	*a*	*l*												
T. atrica	A	T	A	T	T	A	A	T	T	A	G	A	G	T	A	G	C	T	T	T	T
	M	*e*	*t*	*L*	*e*	*u*	*I*	*I*	*e*	*S*	*e*	*r*	*V*	*a*	*l*	*A*	*l*	*a*	*P*	*h*	*e*
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244
D. yakuba	T	T	A	A	C	T	T	T	A	T	T	A	G	A	A	C	G	T	A	A	A
	L	*e*	*u*				*L*	*e*	*u*	*L*	*e*	*u*									
T. atrica	T	A	T	A	C	A	A	T	T	A	T	A	G	A	A	C	G	A	A	A	A
	T	*y*	*r*	*T*	*h*	*r*	*I*	*I*	*e*	*M*	*e*	*t*	*G*	*l*	*u*	*A*	*r*	*g*	*L*	*y*	*s*
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
										L	*e*	*u*									
T. gigantea 1	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
										L	*e*	*u*									
T. gigantea 2	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
										L	*e*	*u*									
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
										L	*e*	*u*									

Continued →

Table 5.6. Continued

	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265
D. yakuba	G	T	T	T	T	A	G	G	G	T	A	T	A	T	T	C	A	A	A	T	T
	V	*a*	*l*										*I*	*l*	*e*						
T. atrica	A	T	T	C	T	A	G	G	A	T	A	T	A	T	A	C	A	A	A	T	T
	I	*l*	*e*	*L*	*e*	*u*	*G*	*l*	*y*	*T*	*y*	*r*	*M*	*e*	*t*	*G*	*l*	*n*	*I*	*l*	*e*
T. saeva 1	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286
D. yakuba	C	G	T	A	A	A	G	G	A	C	C	T	A	A	T	A	A	A	G	T	T
T. atrica	C	G	A	A	A	A	G	G	A	C	C	T	A	A	T	A	A	A	G	T	A
	A	*r*	*g*	*L*	*y*	*s*	*G*	*l*	*y*	*P*	*r*	*o*	*A*	*s*	*n*	*L*	*y*	*s*	*V*	*a*	*l*
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307
D. yakuba	G	G	T	T	T	A	A	T	A	G	G	A	A	T	T	C	C	T	C	A	A
				L	*e*	*u*	*M*	*e*	*t*							*P*	*r*	*o*			
T. atrica	G	G	T	A	T	T	T	T	A	G	G	T	A	T	T	T	T	A	C	A	A
	G	*l*	*y*	*I*	*l*	*e*	*L*	*e*	*u*	*G*	*l*	*y*	*I*	*l*	*e*	*L*	*e*	*u*	*G*	*l*	*n*
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.

	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328
D. yakuba	C	C	T	T	T	T	T	G	T	G	A	T	G	C	A	A	T	T	A	A	A
							C	*y*	*s*												
T. atrica	C	C	A	T	T	T	T	C	A	G	A	C	G	C	A	A	T	T	A	A	A
	P	*r*	*o*	*P*	*h*	*e*	*S*	*e*	*r*	*A*	*s*	*p*	*A*	*l*	*a*	*I*	*l*	*e*	*L*	*y*	*s*
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Continued →

Table 5.6. Continued

	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349
D. yakuba	T	T	A	T	T	T	A	C	A	A	A	A	G	A	A	C	A	A	A	C	T
							T	h	r				G	l	u	G	l	n	T	h	r
T. atrica	T	T	A	T	T	T	A	A	T	A	A	A	A	A	T	T	T	A	A	T	A
	L	e	u	P	h	e	A	s	n	L	y	s	A	s	n	L	e	u	M	e	t
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370
D. yakuba	T	A	T	C	C	A	T	T	A	T	T	A	T	C	A	A	A	T	T	A	T
	T	y	r	P	r	o	L	e	u	L	e	u	S	e	r				T	y	r
T. atrica	A	C	A	T	C	A	G	A	A	A	C	A	A	T	A	A	A	T	T	T	C
	T	h	r	S	e	r	G	l	u	T	h	r	M	e	t	A	s	n	P	h	e
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391
D. yakuba	T	T	A	A	G	A	T	A	T	T	A	T	A	T	T	T	C	T	C	C	T
	L	e	u	S	e	r	T	y	r				I	l	e	S	e	r			
T. atrica	T	C	A	A	T	A	A	T	A	T	A	T	T	T	A	A	C	A	C	C	C
	S	e	r	M	e	t	M	e	t	T	y	r	L	e	u	T	h	r	P	r	o
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412
D. yakuba	A	T	T	T	T	T	T	C	T	T	T	A	T	-	-	-	T	T	T	T	A
	I	l	e	P	h	e							P				h	e	L	e	u
T. atrica	G	C	A	T	T	A	T	C	T	T	T	A	T	C	A	A	T	T	T	C	A
	A	l	a	L	e	u	S	e	r	L	e	u	S	e	r	I	l	e	S	e	r
T. saeva 1	.	T	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.
	V	a	l																		
T. saeva 2	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.
	V	a	l																		
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.

Continued →

Table 5.6. Continued

	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433
D. yakuba	T	C	T	T	T	A	T	T	T	G	T	T	T	G	A	A	T	A	T	G	T
	S	*e*	*r*	*L*	*e*	*u*	*P*	*h*	*e*	*V*	*a*	*l*	*T*	*r*	*p*	*M*	*e*	*t*	*C*	*y*	*s*
T. atrica	A	T	T	A	T	A	A	T	A	A	T	C	T	C	A	A	T	T	A	T	C
	I	*l*	*e*	*M*	*e*	*t*	*M*	*e*	*t*	*I*	*l*	*e*	*S*	*e*	*r*	*I*	*l*	*e*	*I*	*l*	*e*
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454
D. yakuba	A	T	A	C	C	T	T	T	T	T	T	T	G	T	T	A	A	A	T	T	A
	M	*e*	*t*	*P*	*r*	*o*	*P*	*h*	*e*	*P*	*h*	*e*	*V*	*a*	*l*	*L*	*y*	*s*	*L*	*e*	*u*
T. atrica	A	C	A	T	T	T	A	A	T	A	T	A	T	A	T	C	C	T	A	T	A
	T	*h*	*r*	*P*	*h*	*e*	*A*	*s*	*n*	*M*	*e*	*t*	*T*	*y*	*r*	*P*	*r*	*o*	*M*	*e*	*t*
T. saeva 1	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 1	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475
D. yakuba	T	A	C	T	C	T	T	T	T	A	A	T	T	T	A	G	G	T	G	G	A
	T	*y*	*r*	*S*	*e*	*r*	*P*	*h*	*e*	*A*	*s*	*n*	*L*	*e*	*u*	*G*	*l*	*y*	*G*	*l*	*y*
T. atrica	T	T	C	G	A	T	A	A	T	A	A	A	C	A	T	T	C	A	A	T	T
	P	*h*	*e*	*A*	*s*	*p*	*A*	*s*	*n*	*L*	*y*	*s*	*H*	*i*	*s*	*S*	*e*	*r*	*I*	*l*	*e*
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C

	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496
D. yakuba	T	T	A	T	T	T	T	T	T	T	T	A	T	G	T	T	G	T	A	C	A
				P	*h*	*e*				*L*	*e*	*u*	*C*	*y*	*s*	*C*	*y*	*s*	*T*	*h*	*r*
T. atrica	T	T	A	T	T	A	T	T	T	T	T	T	A	T	T	C	T	A	T	C	A
	L	*e*	*u*	*L*	*e*	*u*	*P*	*h*	*e*	*P*	*h*	*e*	*I*	*l*	*e*	*L*	*e*	*u*	*S*	*e*	*r*
T. saeva 1	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. saeva 2	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	G	.	.	.
T. gigantea 1	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.
T. gigantea 2	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	G	.	.	.
T. gigantea + T. saeva	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	G	.	.	.

Continued →

Table 5.6. Continued

	4 9 7	4 9 8	4 9 9	5 0 0	5 0 1	5 0 2	5 0 3	5 0 4	5 0 5	5 0 6	5 0 7	5 0 8	5 0 9	5 1 0	5 1 1	5 1 2	5 1 3	5 1 4	5 1 5	5 1 6	5 1 7
D. yakuba	A	G	A	T	T	A	G	G	A	G	T	T	T	A	T	A	C	A	G	T	T
				L	*e*	*u*	*G*	*l*	*y*							*T*	*h*	*r*	*V*	*a*	*l*
T. atrica	T	C	A	A	T	A	T	C	A	G	T	A	T	A	T	A	T	T	A	T	T
	S	*e*	*r*	*M*	*e*	*t*	*S*	*e*	*r*	*V*	*a*	*l*	*T*	*y*	*r*	*I*	*l*	*e*	*I*	*l*	*e*
T. saeva 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
T. saeva 2	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
				T	*h*	*r*															
T. gigantea 1	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
				T	*h*	*r*															
T. gigantea 2	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
				T	*h*	*r*															
T. gigantea + T. saeva	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
				T	*h*	*r*															

	5 1 8	5 1 9	5 2 0	5 2 1	5 2 2	5 2 3	5 2 4	5 2 5	5 2 6	5 2 7	5 2 8	5 2 9	5 3 0	5 3 1	5 3 2	5 3 3	5 3 4	5 3 5	5 3 6	5 3 7	5 3 8	5 3 9	5 4 0	5 4 1	5 4 2	5 4 3 (12295)
D. yakuba	A	T	A	G	T	A	G	C	T	G	G	C	T	G	A	T	C	T	T	C	T	A	A	T	T	C
	M	*e*	*t*	*V*	*a*	*l*	*A*	*l*	*a*							*S*	*e*	*r*	*S*	*e*	*r*					
T. atrica	C	T	A	T	T	A	A	T	T	G	G	A	T	G	A	A	T	T	A	C	C	A	A	C	T	C
	L	*e*	*u*	*L*	*e*	*u*	*I*	*l*	*e*	*G*	*l*	*y*	*T*	*r*	*p*	*I*	*l*	*e*	*T*	*h*	*r*	*A*	*s*	*n*		
T. saeva 1	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	T	.	.	.	.	.
T. saeva 2	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.
T. gigantea 1	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.
T. gigantea 2	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.
T. gigantea + T. saeva	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.

T. atrica acts as a consensus sequence for both nucleotide and amino acid sequence (shown from position 185 for the spiders). Nucleotide identity is shown by a dot, deletions are shown by a dash. Haplotypes for the *T. atrica* group are included along with the *Drosphila yakuba* sequence (Clary and Wolstenholme, 1985) for comparison (amino acids identical to *T. atrica* are not shown). Boxed numbers above the alignment refer to the equivalent position in the *D. yakuba* mtDNA molecule. Nucleotides are numbered consecutively from 1 to 543. The position of primer LR-J-12881, overlapping with the 16S fragment is shown. The putative start position of ND1 transcription is shown. There were large deletions surrounding the tRNA *leu* ^{CUN} gene relative to *D. yakuba* making alignment difficult in this region. The known and possible extent of the tRNA *leu* ^{CUN} gene is shown by solid and broken boxes respectively. The anti-codon is underlined.

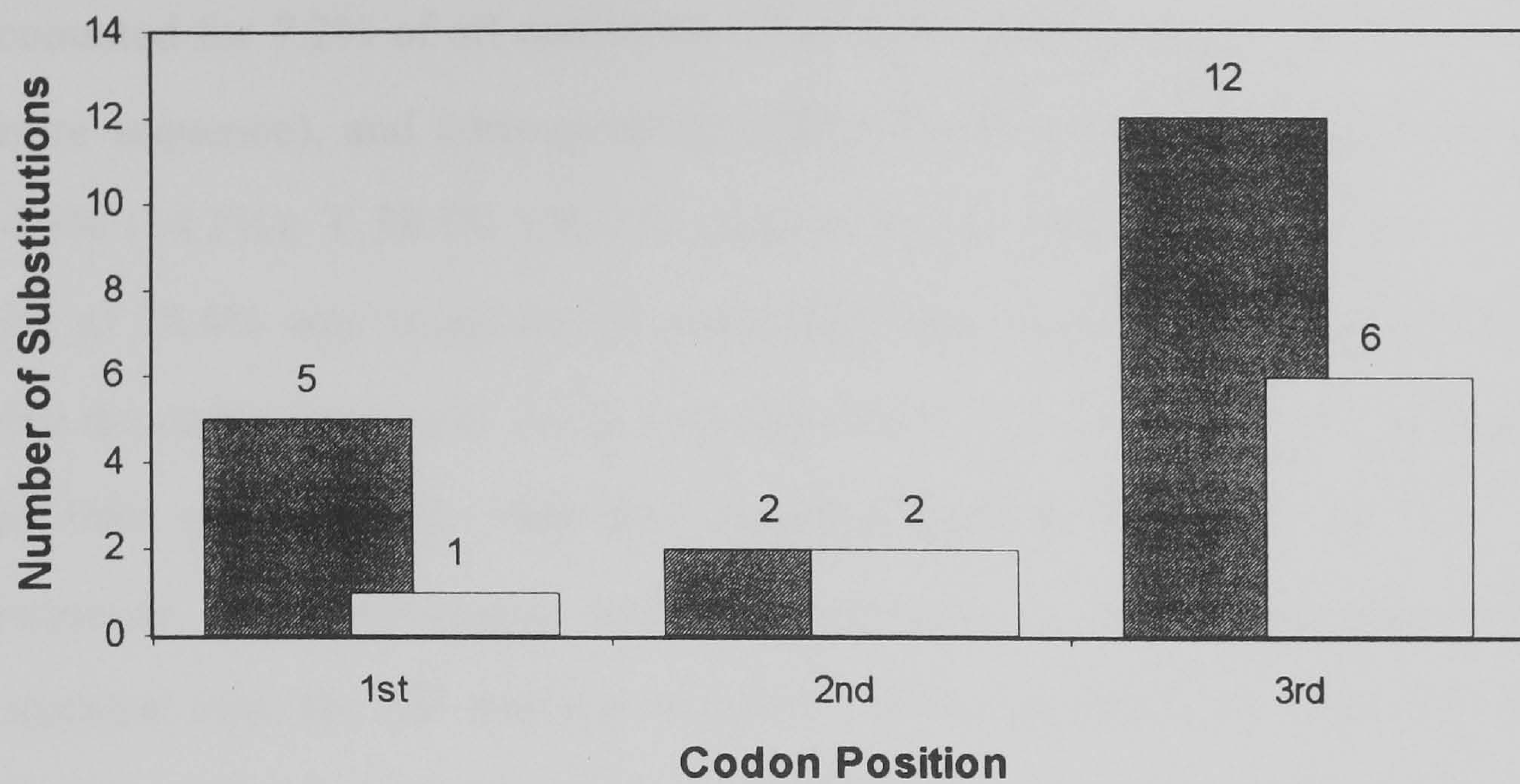


Figure 5.10. Distribution of substitutions among the three codon positions in the mt ND1 fragment. Shaded boxes = total substitutions. Unshaded boxes = potentially phylogenetically informative substitutions.

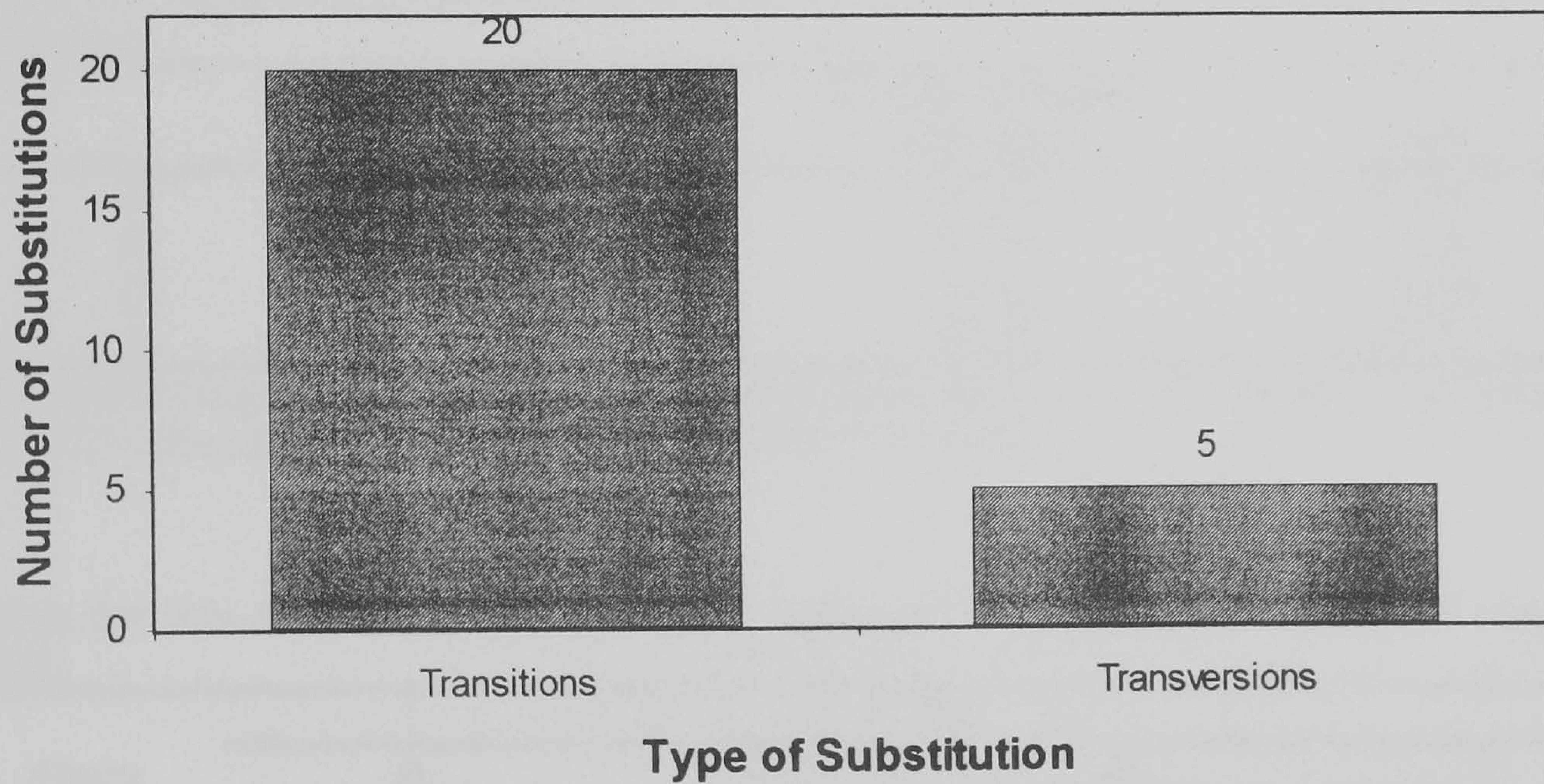


Figure 5.11. The number of substitutions of each class in the mt ND1 fragment in members of the *T. atrica* group.

and 11.1% in the first position. The frequencies of the four nucleotides in each of the codon positions for the coding region are given in Table 5.7. G accounted for 7.2% of all nucleotide sites in the coding region (9.2% over the entire sequence), and corresponding values for the other nucleotides were C, 14.5% (14.2%); T 38.1% (36.3%); and A 40.3% (40.2%). An overall A + T bias of 78.4% was found for all nucleotide sites in the coding region (76.6% over the entire sequence). As is expected, the A + T bias was most apparent in the third position. The empirical transition:transversion ratio was 4:1. The frequency of unambiguous nucleotide changes in each substitution class, estimated from the MP tree (see Section 5.3.4.4), are shown in Table 5.9. This gives an estimated transition:transversion ratio of 5.3:1. It can also be seen that A → G and T → C transitions occurred 1.7 times more frequently than their reciprocal substitutions, and that A ↔ T accounted for approximately 66.5% of all transversions (all other transversions being C → G).

Table 5.7. Nucleotide frequencies at each codon position for the coding region of the ND1 fragment

	Codon Position		
	First	Second	Third
A	44.44	20.83	55.60
C	9.44	22.78	11.20
G	12.5	8.33	0.006
T	33.61	48.06	32.63

The overall frequency of each nucleotide at each codon position is given for the 6 haplotypes and 359 coding bp (positions 185 through 543) of ND1 analysed.

Table 5.8. The frequency of unambiguous nucleotide changes in the single MP tree for ND1.

	To			
From	A	C	G	T
A	-	0.000	0.105	0.105
C	0.000	-	0.053	0.316
G	0.000	0.000	-	0.000
T	0.000	0.421	0.000	-

5.3.2.4 Combined CO1 and ND1 Fragments

Both the CO1 and ND1 fragments were sequenced in eight individuals of the *T. atrica* group. The combination of these two fragments produced a sequence of 868 bp revealing six haplotypes. Of the 868 bp analysed, 818 (94.2%) were invariant, 31 (3.6%) were autapomorphic, and 19 (2.2%) were potentially phylogentically informative. The empirical transition:transversion ratio was 9:1. The frequency of unambiguous nucleotide changes in each substitution class, estimated from the MP tree (see Section 5.3.4.6), are shown in Table 5.9. This gives an estimated transition:transversion ratio of 8.7:1. The table shows that A → G and T → C transitions occurred 1.9 times more frequently than their reciprocal substitutions, and that T → A accounted for approximately 67% of all transversions (all other transversions being G → C).

Table 5.9. The frequency of unambiguous nucleotide changes in the MP tree for the combined CO1 and ND1 data.

	To			
From	A	C	G	T
A	-	0.000	0.483	0.000
C	0.000	-	0.000	0.034
G	0.276	0.034	-	0.000
T	0.069	0.103	0.000	-

5.3.3 mtDNA Sequence Variation

5.3.3.1 CO1

Within-species CO1 sequence variation appeared to be low in the genus *Tegenaria*. The three specimens of *T. domestica*, two from different sites in Devon and one from York, all exhibited the same haplotype. The three

specimens of *T. parietina*, two from Pavia in northern Italy and one from Kent, also shared one haplotype. The three specimens of *T. atrica*, one from southern France, one from County Dublin in Eire, and one from York, again showed the same haplotype. Although the sample sizes were small the geographic separation of the specimens used was certainly large. Only one individual of *T. agrestis* and one individual of the outgroup species *Tetrix denticulata* were sequenced. Seven individuals of *T. saeva* were analysed; four exhibited a haplotype unique to this species (designated '*T. saeva*'), two revealed a haplotype shared with *T. gigantea* ('*T. gigantea* + *T. saeva 1*'), and one revealed a further haplotype shared with *T. gigantea* ('*T. gigantea* + *T. saeva 2*'). Of the 10 individuals of *T. gigantea* sequenced, only one showed a haplotype unique to this species (designated '*T. gigantea*'): six exhibited the '*T. gigantea* + *T. saeva 1*' haplotype, and three exhibited the '*T. gigantea* + *T. saeva 2*' haplotype. One putative *T. gigantea/saeva* hybrid from York was sequenced and showed the haplotype '*T. gigantea* + *T. saeva 1*'.

Table 5.10. Pairwise distances among the nine mt CO1 haplotypes analysed.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) *T. atrica*	-	0.0538	0.0605	0.0506	0.0637	0.0960	0.2255	0.2478	0.1660
(2) *T. saeva*	17	-	0.0314	0.0281	0.0345	0.1030	0.2172	0.2308	0.1660
(3) *T. gigantea*	19	10	-	0.0155	0.0062	0.0993	0.2177	0.2229	0.1549
(4) *T. gigantea + T. saeva 1*	16	9	5	-	0.0187	0.1030	0.2172	0.2224	0.1583
(5) *T. gigantea + T. saeva 2*	20	11	2	6	-	0.0995	0.2130	0.2182	0.1545
(6) *T. agrestis*	29	31	30	31	30	-	0.2188	0.1993	0.1580
(7) *T. domestica*	61	59	59	59	58	59	-	0.2478	0.2014
(8) *T. parietina*	66	62	60	60	59	55	66	-	0.1524
(9) *T. denticulata*	47	47	44	45	44	45	56	44	-

Above the diagonal: pairwise nucleotide sequence divergence estimated using the Kimura two-parameter model. Below diagonal: absolute number of nucleotide differences between sequences.

Examination of the pairwise distances in Table 5.10, and the alignment in Table 5.4, shows that the three haplotypes occurring in *T. gigantea* ('*T. gigantea*', '*T. gigantea* + *T. saeva 1*', and '*T. gigantea* + *T. saeva 2*') were similar with pairwise differences between them ranging from 0.62 % to 1.87 % (2 to 6 substitutions). The haplotype '*T. gigantea* + *T. saeva 1*' was the most

divergent having one amino acid change (methionine to valine) at nucleotide position 49. In all other respects however these haplotypes were very similar. The one *T. saeva* haplotype (not shared with *T. gigantea*) was divergent from the three *T. gigantea* haplotypes by 2.81 % to 3.45 % (9 to 11 substitutions). There were two amino acid substitutions not shared with the three *T. gigantea* haplotypes: one (valine to isoleucine) at nucleotide position 43, and one (alanine to valine) at nucleotide position 127. The *T. atrica* haplotype was divergent from the *T. gigantea* and *T. saeva* haplotypes by 5.06 % to 6.37 % (17 to 20 substitutions). There were two amino acids not shared with the other members of the *T. atrica* group: one (serine to lysine) at nucleotide position 112, and one (isoleucine to valine) at nucleotide position 268. *Tegenaria agrestis* differed from the members of the *T. atrica* group by 9.60 % to 10.30 % (29 to 31 substitutions). *Tegenaria parietina* and *T. domestica* differed from all other species haplotypes, and from each other, by similar amounts: 19.93 % to 24.78 % (55 to 66 substitutions). The distances for *T. parietina* and *T. domestica* may not be accurate, however, given the very high third position bias and pairwise distances in excess of 20 %, and that saturation and reversal mutations at the third position cannot be ruled out. This implies that the distances for these two taxa may be underestimated and that phylogenetic reconstruction of the relationships between these taxa should take account of these uncertainties. A similar case could of course be made for the outgroup species *Tetrix denticulata*, which shows unexpectedly low distances (perhaps because of reversals?). (The postion of this species as an outgroup was however supported in the parsimony analyses – section 5.3.4.)

5.3.3.2 16S

The 16S fragment, although not particularly variable and only sequenced for four individuals of the *T. atrica* group, revealed a similar pattern to the CO1

fragment. One haplotype ('*T. gigantea* + *T. saeva*') was common to an individual of *T. gigantea* from Kent and an individual of *T. saeva* from Cornwall. There was one '*T. saeva*' haplotype from an individual from Nancy, France, which diverged from the '*T. gigantea* + *T. saeva*' haplotype by 0.46% (2 substitutions). The '*T. atrica*' haplotype, from County Dublin, Eire, differed from both of these haplotypes by 1.15% to 1.39% (5 to 6 substitutions). The pairwise distances are given in Table 5.11, below.

Table 5.11. Pairwise distances among the three mt 16S haplotypes analysed.

	(1)	(2)	(3)
(1) *T. atrica*	-	0.0139	0.0115
(2) *T. saeva*	6	-	0.0046
(3) *T. gigantea + T. saeva*	5	2	-

Above the diagonal: pairwise nucleotide sequence divergence estimated using the Kimura two-parameter model. Below diagonal: absolute number of nucleotide differences between sequences.

5.3.3.3 ND1

The ND1 fragment, for 12 individuals of the *T. atrica* group, revealed six haplotypes. The two specimens of *T. atrica*, one from southern France and one from County Dublin, Eire, shared the same haplotype. Five individuals of *T. saeva* were sequenced. Three of these individuals shared one haplotype unique to this species (designated '*T. saeva 1*'), one individual showed a further haplotype unique to this species ('*T. saeva 2*'), and the final individual exhibited a haplotype in common with one individual of *T. gigantea* (designated '*T. gigantea* + *T. saeva*'). Five individuals of *T. gigantea* were sequenced, one of which exhibited the '*T. gigantea* + *T. saeva*' haplotype. Three individuals shared a haplotype unique to this species ('*T. gigantea 1*'), and one individual revealed a further unique haplotype ('*T. gigantea 2*').

Examination of the pairwise distances in Table 5.12, and the alignment in Table 5.6, revealed that the three *T. gigantea* haplotypes ('*T. gigantea 1*', '*T. gigantea 2*', '*T. gigantea + saeva*') and the '*T. saeva 2*' haplotype were similar

with pairwise distances between them ranging from 0.18% to 0.74% (1 to 4 substitutions). All substitutions between these haplotypes were silent with the exception of one amino acid change in '*T. gigantea 2*' at position 392 (alanine to valine), a change shared with the '*T. saeva 1*' haplotype. The '*T. saeva 1*' haplotype differed from the above haplotypes by 2.62% to 3.20% (14 to 17 substitutions). There was one amino acid difference relative to the above haplotypes (with the exception of position 392 already mentioned) at position 233 (leucine to methionine). The *T. atrica* haplotype differed from the '*T. saeva 1*' haplotype by 3.20% (17 substitutions), and from the *T. gigantea* group of haplotypes by 2.81% to 3.01% (15 to 16 substitutions). There were no unique amino acid substitutions.

Table 5.12. Pairwise distances among the six mt ND1 haplotypes analysed.

	(1)	(2)	(3)	(4)	(5)	(6)
(1) *T. atrica*	-	0.0320	0.0301	0.0281	0.0301	0.0281
(2) *T. saeva 1*	17	-	0.0320	0.0262	0.0281	0.0301
(3) *T. saeva 2*	16	17	-	0.0055	0.0074	0.0055
(4) *T. gigantea 1*	15	14	3	-	0.0055	0.0037
(5) *T. gigantea 2*	16	15	4	3	-	0.0018
(6) *T. gigantea + T. saeva*	15	16	3	2	1	-

Above the diagonal: pairwise nucleotide sequence divergence estimated using the Kimura two-parameter model. Below diagonal: absolute number of nucleotide differences between sequences.

5.3.3.4 Combined CO1 and ND1 Fragments

The combination of the CO1 and ND1 fragments for eight individuals of the *T. atrica* group revealed six haplotypes. There was one *T. atrica* haplotype, one haplotype unambiguously assignable to *T. saeva*, and four haplotypes that could be attributed to both *T. saeva* and *T. gigantea* (designated '*T. gigantea + T. saeva 1*', '*T. gigantea + T. saeva 2*', '*T. gigantea + T. saeva 3*', and '*T. gigantea + T. saeva 4*'). Examination of the pairwise distances in Table 5.13 below, reveals essentially the same pattern as when the fragments are considered separately. The '*T. gigantea + T. saeva 3*' and '*T. gigantea + T. saeva 4*' haplotypes only differed by 0.11% (1 substitution), and the '*T.*

gigantea + T. saeva 1' and '*T. gigantea + T. saeva 2'* haplotypes differed by only 0.35% (3 substitutions). Overall, among these four haplotypes, attributable to both *T. gigantea* and *T. saeva,* pairwise differences ranged from 0.11% to 1.16% (1 to 10 substitutions). The '*T. saeva'* haplotype differed from the above haplotypes by 2.81% to 3.29% (24 to 28 substitutions). The '*T. atrica'* haplotype differed from the '*T. saeva'* haplotype by 4.01% (34 substitutions) and from the four '*T. gigantea + T. saeva'* haplotypes by a similar amount: 3.65% to 4.25% (31 to 36 substitutions).

Table 5.13. Pairwise distances among the six combined mt CO1 and ND1 haplotypes analysed.

	(1)	(2)	(3)	(4)	(5)	(6)
(1) *T. atrica*	-	0.0401	0.0425	0.0413	0.0377	0.0365
(2) *T. saeva*	34	-	0.0329	0.0293	0.0281	0.0293
(3) *T. gigantea + T. saeva 1*	36	28	-	0.0035	0.0116	0.0104
(4) *T. gigantea + T. saeva 2*	35	25	3	-	0.0104	0.0093
(5) *T. gigantea + T. saeva 3*	32	24	10	9	-	0.0011
(6) *T. gigantea + T. saeva 4*	31	25	9	8	1	-

Above the diagonal: pairwise nucleotide sequence divergence estimated using the Kimura two-parameter model. Below diagonal: absolute number of nucleotide differences between sequences.

5.3.4 Phylogeny Reconstruction

5.3.4.1 CO1 - Parsimony Analysis

Parsimony analysis was performed on the 325 nucleotides of mitochondrial CO1 gene for the eight *Tegenaria* haplotypes plus *Tetrix denticulata* as an outgroup.

The data were first assessed to see if they contained a strong phylogenetic signal by examining the tree length distribution of 10,000 unrooted, randomly generated trees. A data set containing random variation (noise) will tend to generate a tree length distribution that is approximately normal and

symmetrical, whereas a data set containing a phylogenetic signal will tend to be left-skewed and hence most random trees will tend to be longer than the most parsimonious trees. Under these circumstances the probability of a parsimony analysis finding the correct tree is extremely high (Hillis, 1991; Hillis and Huelsenbeck, 1992). Figure 5.12 shows the tree length distribution for the 10,000 random trees generated for the CO1 data. The distribution was significantly left-skewed as measured by the *g*1 statistic ($g1 = -0.899$, $p < 0.01$, Hillis and Huelsenbeck, 1992), suggesting a strong phylogenetic signal in the data.

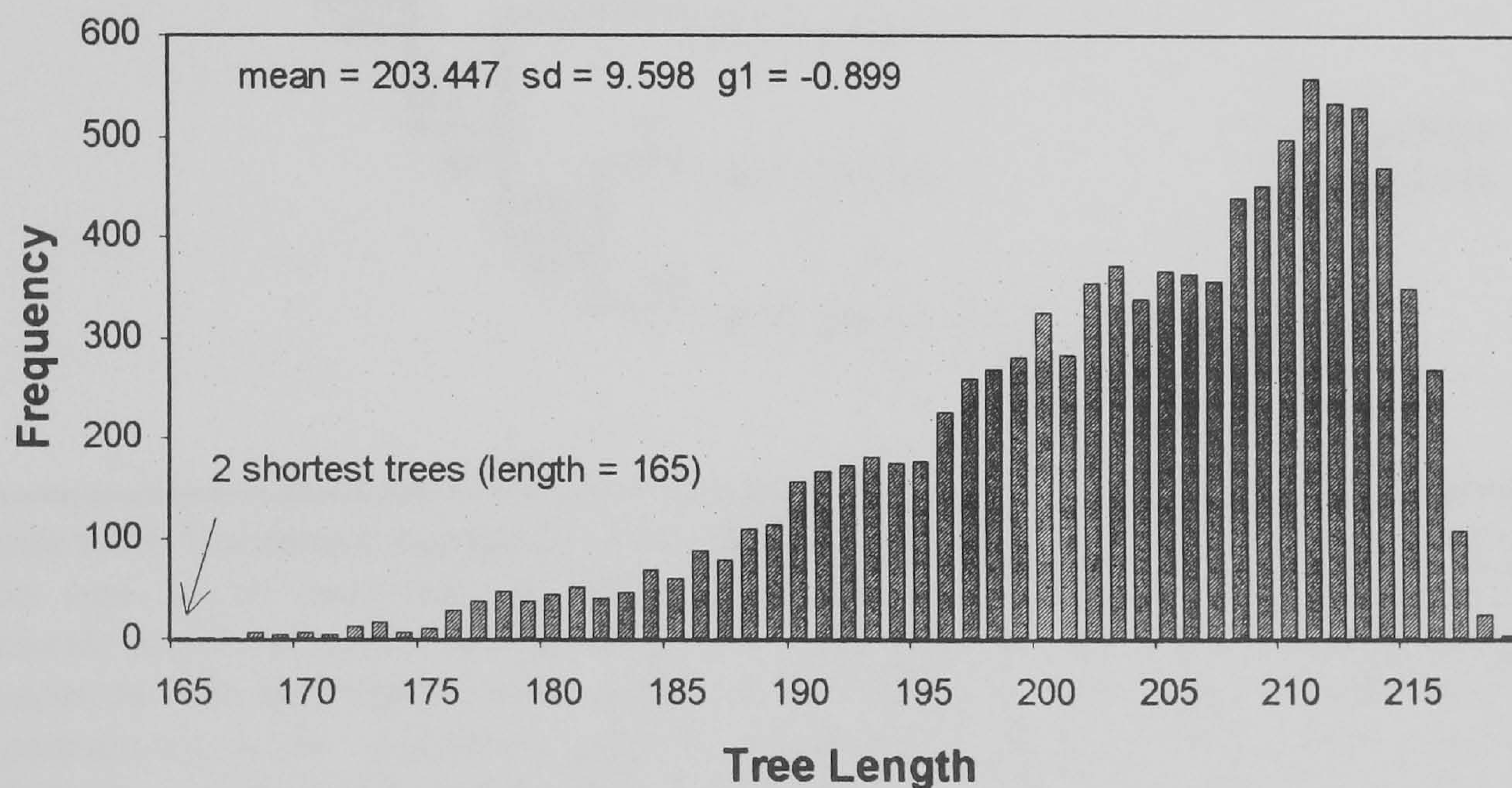


Figure 5.12. Frequency distribution of the lengths of 10 000 random trees for the CO1 data. Generated using the eight *Tegenaria* haplotypes and the outgroup *Tetrix denticulata*.

An exhaustive search in which all nucleotide sites were equally weighted quickly returned a single MP tree of length 160 - five steps shorter than the shortest random tree. This tree is shown in Figure 5.13. The consistency index (CI) was 0.819, the homoplasy index (HI) was 0.181, the retention index (RI) was 0.670, and the rescaled consistency index (RC) was 0.549. These indices measure the "fit" of characters to particular trees; hence they are useful not only in comparing characters on single trees but across multiple trees to see

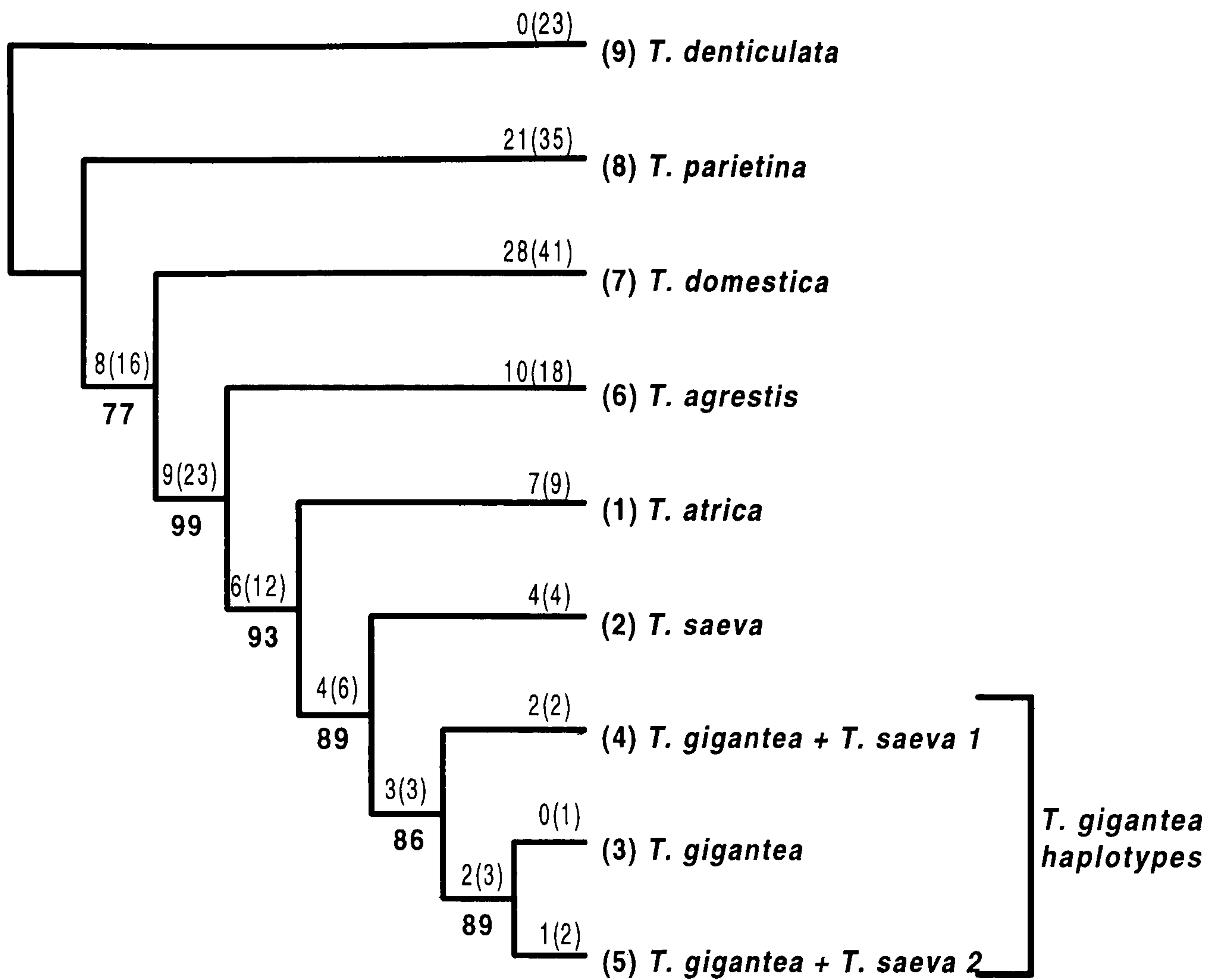


Figure 5.13. Maximum parsimony (MP) tree for the CO1 data. Bold numerals below the nodes indicate the percentage of 1000 bootstrap replicate samples recovering each clade. Numerals above the nodes and at branch tips indicate synapomorphies and autapomorphies respectively. The first number is the number of unambiguous changes, and number following, in parentheses, is the maximum possible number of changes allowed under parsimony. Haplotypes are numbered (before their names) in accordance with Table 5.10 (page 206). Haplotypes forming a distinct *T. gigantea* grouping have been bracketed.

which characters support which topology hypothesis (Swofford, 1993). They are defined in terms of three parameters as described by Swofford (1993):

> s = length (number of steps) required by the character on the tree
>
> m = minimum amount of change that the character may show on any conceivable tree
>
> g = maximum possible amount of change that a character could possibly require on any conceivable tree, i.e. the length on a completely unresolved bush.

CI = m/s (Kluge and Farris, 1969). For a single character and a tree that describes the data as well as any tree could then CI = 1. Unfortunately, CI does not range from 0 to 1 - the lower bound is a function of the character-state distribution in the data matrix.
RI = $(g-s)/(g-m)$ (Farris, 1989). When a character fits the tree as poorly as possible, RI = 0.
RC = RI × CI (Farris, 1989). RI is used to scale CI between 0 and 1.
HI = 1-CI, (Kluge and Farris, 1969; Swofford, 1993). A measure of homoplasy, the number of characters evolving more than once across the tree, and therefore not directly attributable to common ancestry.

The index values given are actually overall indices for all characters across the trees, whereby, for instance, CI = M/S, where M and S are the sums over all characters in the suite of the individual m and s values. The other indices are calculated analogously.

Exhaustive searches were also performed under various character weighting combinations. The transition:transversion ratio was varied so as to reflect the

empirical ratio (1.85:1) and the optimised ratio estimated from the MP tree (1.42:1) (see section 5.3.2.1). The transition:tranversion ratio was also weighted 3:1, 5:1, and 10:1. Codon positions were weighted 4-8-1. Weighting had no effect on the topology of the tree obtained and showed no general improvement in bootstrap values or goodness of fit measures.

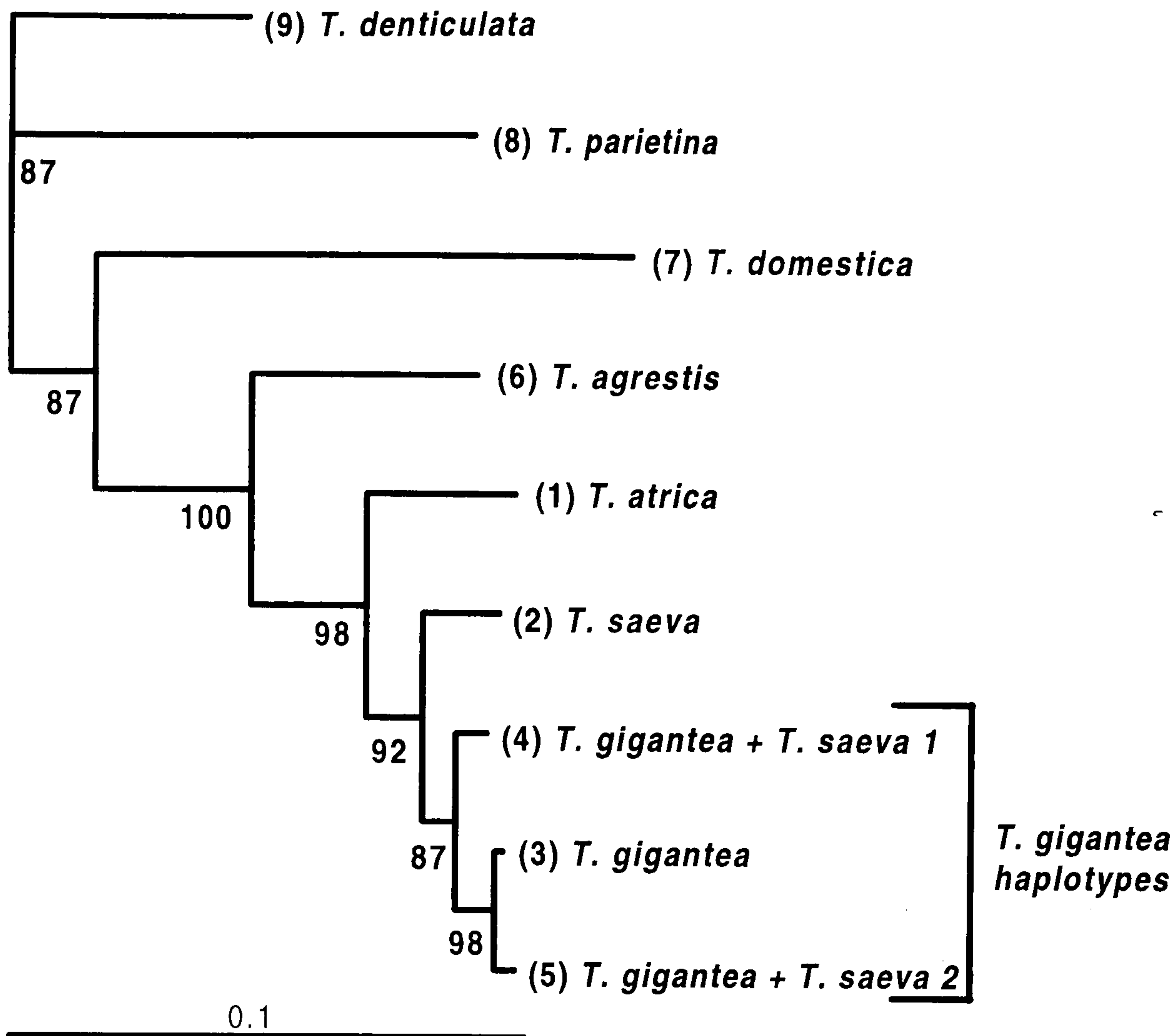
Examination of the MP tree in Figure 5.13, shows that three haplotypes, '*T. gigantea + T. saeva 1*', '*T. gigantea + T. saeva 2*', and '*T. gigantea*' grouped together. Indeed, '*T. gigantea*' and '*T. gigantea + T. saeva 2*' were only separated by a single unambiguous autapomorphy in '*T. gigantea + T. saeva 2*', and these two haplotypes were separated from '*T. gigantea + T. saeva 1*' by two unambiguous synapomorphies (and the possession of two autapomorphies by '*T. gigantea + T. saeva 1*'). This group, referred to as the "*T. gigantea* haplotypes", was separated from the '*T. saeva*' haplotype by three unambiguous synapomorphies, with '*T. saeva*' in addition possessing four unambiguous autapomorphies. '*T. atrica*' was separated from the clade formed by the above species by four unambiguous synapomorphies and in addition possessed seven unambiguous autapomorphies. *T. agrestis* was the next most closely related species to the *T. atrica* group, being separated from the *T. atrica* group by six unambiguous synapomorphies and possessing 10 unambiguous autapomorphies. *T. domestica* was separated from the preceding set of species by nine unambiguous synapomorphies with 28 unambiguous autapomorphies. *T. parietina* was the most distant species in the tree, and was separated by eight unambiguous synapomorphies and possessed 21 unambiguous autapomorphies. The marked increase in the number of ambiguous changes relative to unambiguous changes (maximum number of changes possible less the unambiguous changes), as one moves away from the *T. atrica* group towards the outgroup, reflects the greater divergences separating the more distant species and is due to uncertain/polymorphic states in, mainly, the outgroup, *T.*

parietina, and *T. domestica* (Hedin, 1997). All groupings were strongly supported by high bootstrap values.

5.3.4.2 CO1 - Distance Analysis

Figure 5.14 (A and B) shows neighbor-joining trees (Saitou and Nei, 1987) generated by CLUSTALW version 1.7 (Thompson *et al.*, 1994), from the Kimura two-parameter distance matrix shown in Table 5.10, for the 325 nucleotides of mitochondrial CO1 gene and the eight *Tegenaria* haplotypes plus *T. denticulata* as an outgroup. The topology of the tree was identical to that obtained by parsimony methods. Neighbor-joining trees were also constructed from distance matrices generated using the different transition:transversion ratios as described in the parsimony section above. Changing the ratio had no effect on the topology, little effect on branch length, and did not significantly improve the bootstrap values (which were already high). UPGMA, maximum likelihood, and Fitch-Margoliash trees were also generated using PHYLIP (Felsenstein, 1995) with no change in tree topology.

All groupings were strongly supported by high bootstrap values. As in the parsimony tree, the *T. atrica* group clusters together with the three *T. gigantea* haplotypes (*'T. gigantea'*, *'T. gigantea + T. saeva 1'*, and *'T. gigantea + T. saeva 2'*) forming a discrete group separate from *T. saeva* and *T. atrica*. This is most clearly seen in the unrooted, radial version of the tree (Figure 5.14 B).



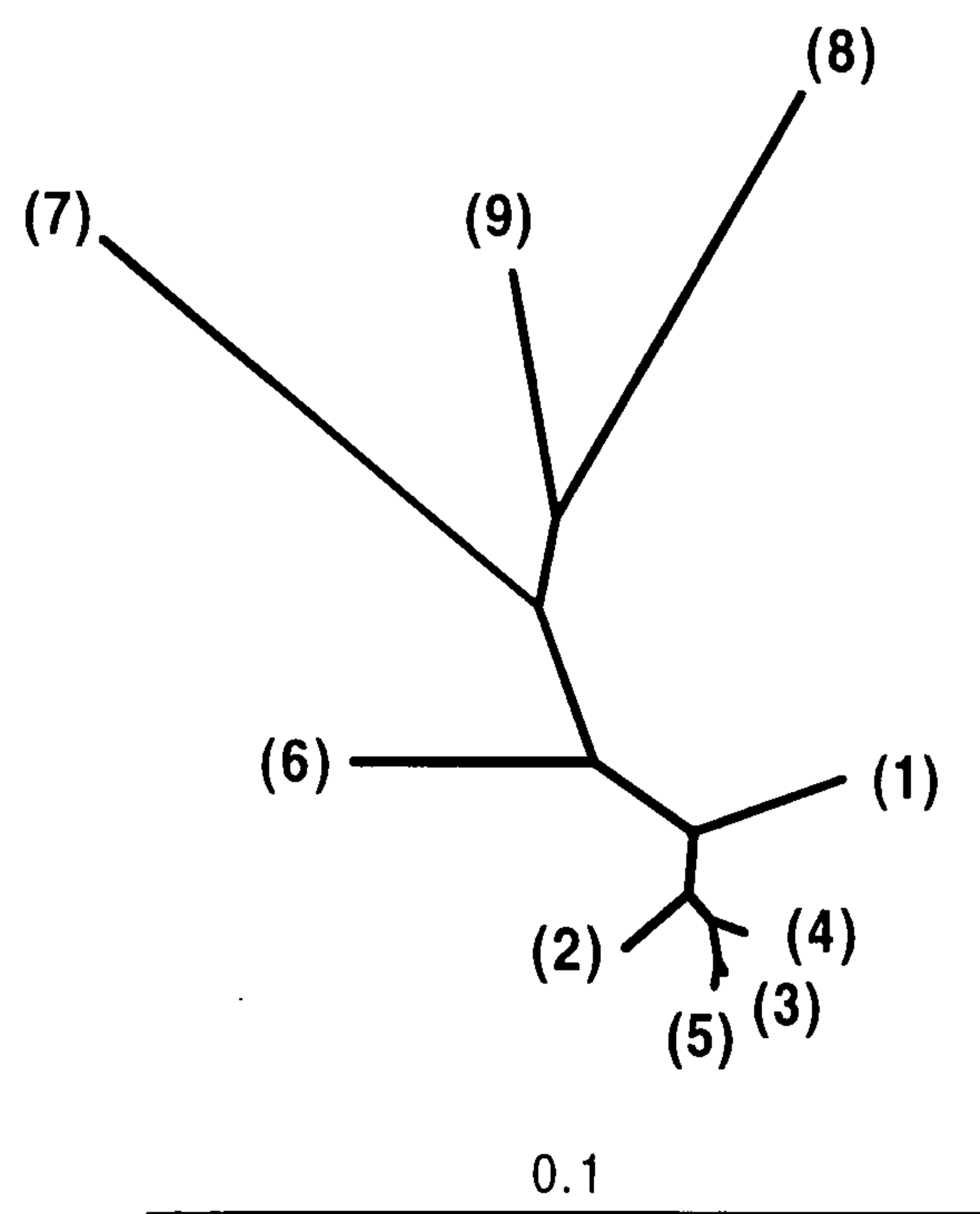


Figure 5.14. Neighbor-joining tree for the CO1 data. A) A phenogram rooted through the outgroup *T. denticulata.* Bold numerals below the nodes indicate the percentage of 1000 bootstrap replicate samples recovering each clade. Haplotypes are numbered (before their names) in accordance with Table 5.10 (page 206). Haplotypes forming a distinct *T. gigantea* grouping have been bracketed. B) An unrooted, radial representation of A to show the distances and groupings between haplotypes more clearly.

5.3.4.3 Relative Rate Tests of CO1 Evolution.

Examination of the neighbor-joining tree in Figure 5.14, shows that *T. domestica* has a slightly longer branch than may have been expected, suggesting that *T. domestica* may have a faster rate of CO1 evolution than the other taxa. Given this, and the concerns that this gene region may be approaching saturation/suffering reversal mutations with respect to the more divergent taxa (*T. parietina*, *T. domestica* and the outgroup *T. denticulata*) a relative rate test was performed (see Table 5.14). None of the comparisons was significant, however it is worth noting that *T. domestica* always experienced more substitutions relative to the outgroup than did the other ingroup taxa. Consequently the associated probability values for *T. domestica* were consistently lower than in other comparisons. It is conceivable that *T. domestica* CO1 has experienced a higher rate of evolution than CO1 in the other taxa but that this is being partly obscured by saturation/reversal mutations. Overall, however, there seems to be a constant rate of CO1 evolution.

Table 5.14. Relative rate test for CO1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) *T. atrica*	-	9/9	12/9	9/7	12/9	18/16	31/40	39/36
(2) *T. saeva*	1.0000	-	7/4	6/4	7/4	18/17	31/40	37/34
(3) *T. gigantea*	0.6636	0.5488	-	3/4	1/1	16/17	30/42	34/34
(4) *T. gigantea/saeva 1*	0.8036	0.7539	1.0000	-	4/3	33/33	29/41	35/34
(5) *T. gigantea/saeva 2*	0.6636	0.5488	1.0000	1.0000	-	16/17	29/41	34/34
(6) *T. agrestis*	0.8642	1.0000	1.0000	1.0000	1.0000	-	30/41	32/31
(7) *T. domestica*	0.3425	0.3425	0.1945	0.1882	0.1882	0.2351	-	44/32
(8) *T. parietina*	0.8176	0.8126	1.0000	1.0000	1.0000	1.000	0.2067	-

The numbers above the diagonal are the number of unique substitions in pairs of haplotypes with respect to the outgroup. The first number refers to the haplotypes in the first column, the second number to the haplotypes in the top row. The numbers below the diagonal are the binomial probabilities.

5.3.4.4 ND1- Parsimony Analysis

Parsimony analysis was performed on the 543 bp of mitochondrial ND1 sequence for the six haplotypes of the *T. atrica* group. As no other outgroup was available the trees were rooted through *T. atrica.* This was the most distant sequence and fell consistently as an outgroup to the *T. saeva* and *T. gigantea* haplotypes in the previously described CO1 analyses.

The tree-length distribution for 10 000 random trees was significantly left-skewed ($g1$ = -1.761, p < 0.01, Hillis and Huelsenbeck, 1992), suggesting a strong phylogenetic signal in the data (see Figure 5.15).

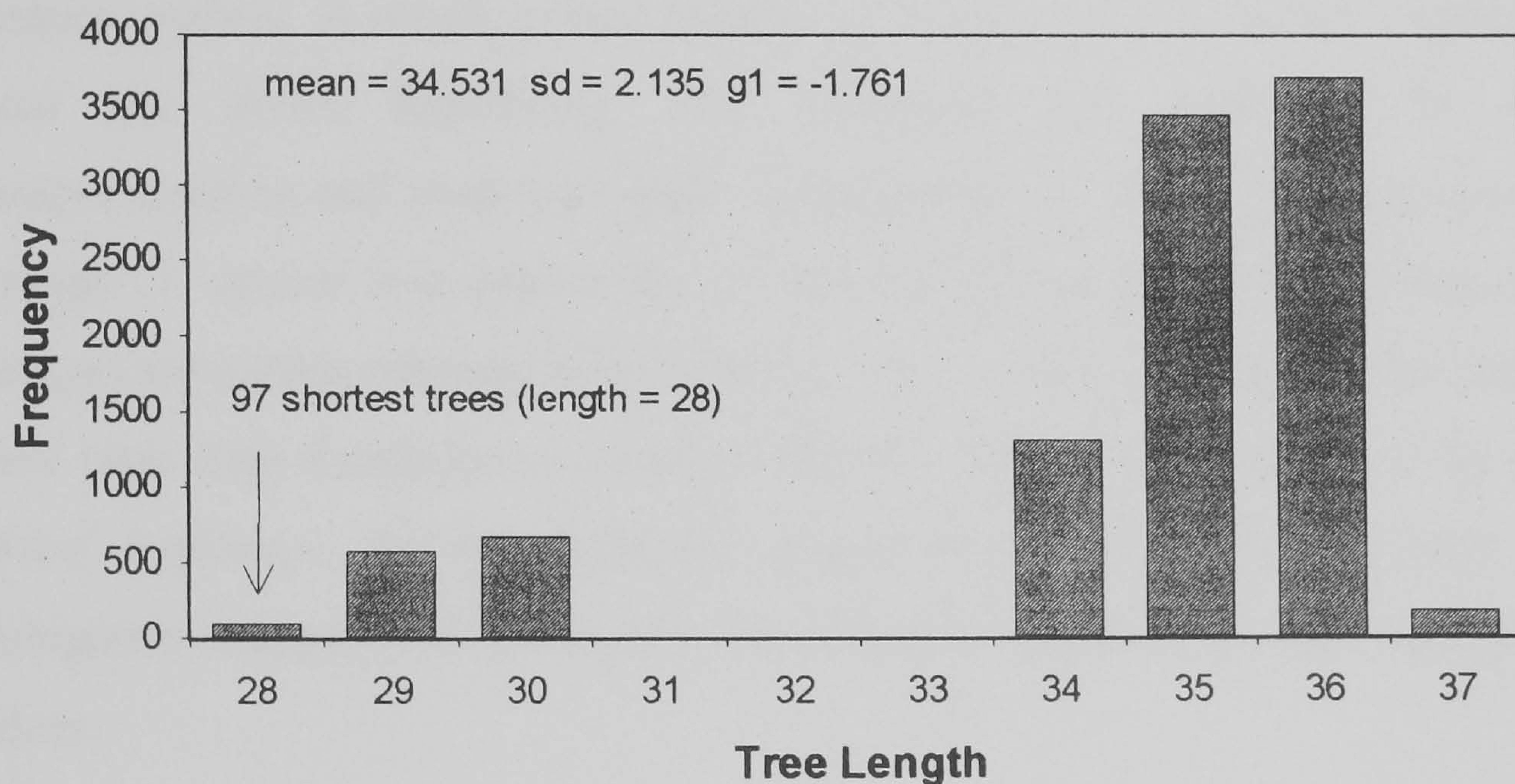


Figure 5.15. Frequency distribution of the lengths of 10 000 random trees for the ND1 data. Generated using the six haplotypes of the *T. atrica* group.

An exhaustive search, in which all nucleotide sites were equally weighted, quickly returned a single MP tree of length 28 (CI = 0.893, HI = 0.107, RI = 0.750, RC = 0.670). As for the CO1 fragment, exhaustive searches were also performed under various character weighting combinations. The transition:transversion ratio was varied so as to reflect the empirical ratio (4:1) and the optimised ratio estimated from the MP tree (5.33:1) (see section

5.3.2.3). The transition:transversion ratio was also weighted 3:1, and 10:1. Codon positions were weighted 4-8-1. Weighting had no effect on the topology of the tree obtained and showed no general improvement in bootstrap values or goodness of fit measures.

Examination of the MP tree in Figure 5.16, shows that four haplotypes grouped together. These will be referred to as the "*T. gigantea* haplotypes". '*T. gigantea 2*' and '*T. gigantea* + *T. saeva*' were only separated by a single autapomorphy in '*T. gigantea 2*'. '*T. saeva 2*' was separated from these two haplotypes by one synapomorphy, and possessed two autapomorphies. '*T. gigantea 1*' was separated from the preceding haplotypes by a single synapomorphy with no autapomorphies. A much greater number of steps makes '*T. saeva 1*' distinct from the above haplotypes; this haplotype was separated by six synapomorphies and possessed eight autapomorphies. As the tree was rooted through '*T. atrica*' it is impossible to calculate the number of unambiguous changes separating this haplotype from the rest; all that can be said is that there were nine steps ('ambiguous autapomorphies'). With the exception of the '*T. atrica*' haplotype, for the technical reason just described, there were no ambiguous changes. All groupings were strongly supported by high bootstrap values.

5.3.4.5 ND1 - Distance Analysis

Figure 5.17 (A and B) shows neighbor-joining trees (Saitou and Nei, 1987) generated by CLUSTALW version 1.7 (Thompson *et al.*, 1994), from the Kimura two-parameter distance matrix shown in Table 5.12, for the 543 nucleotides of mitochondrial ND1 gene for the six *T. atrica* group haplotypes. The topology of the tree was identical to that obtained by parsimony methods.

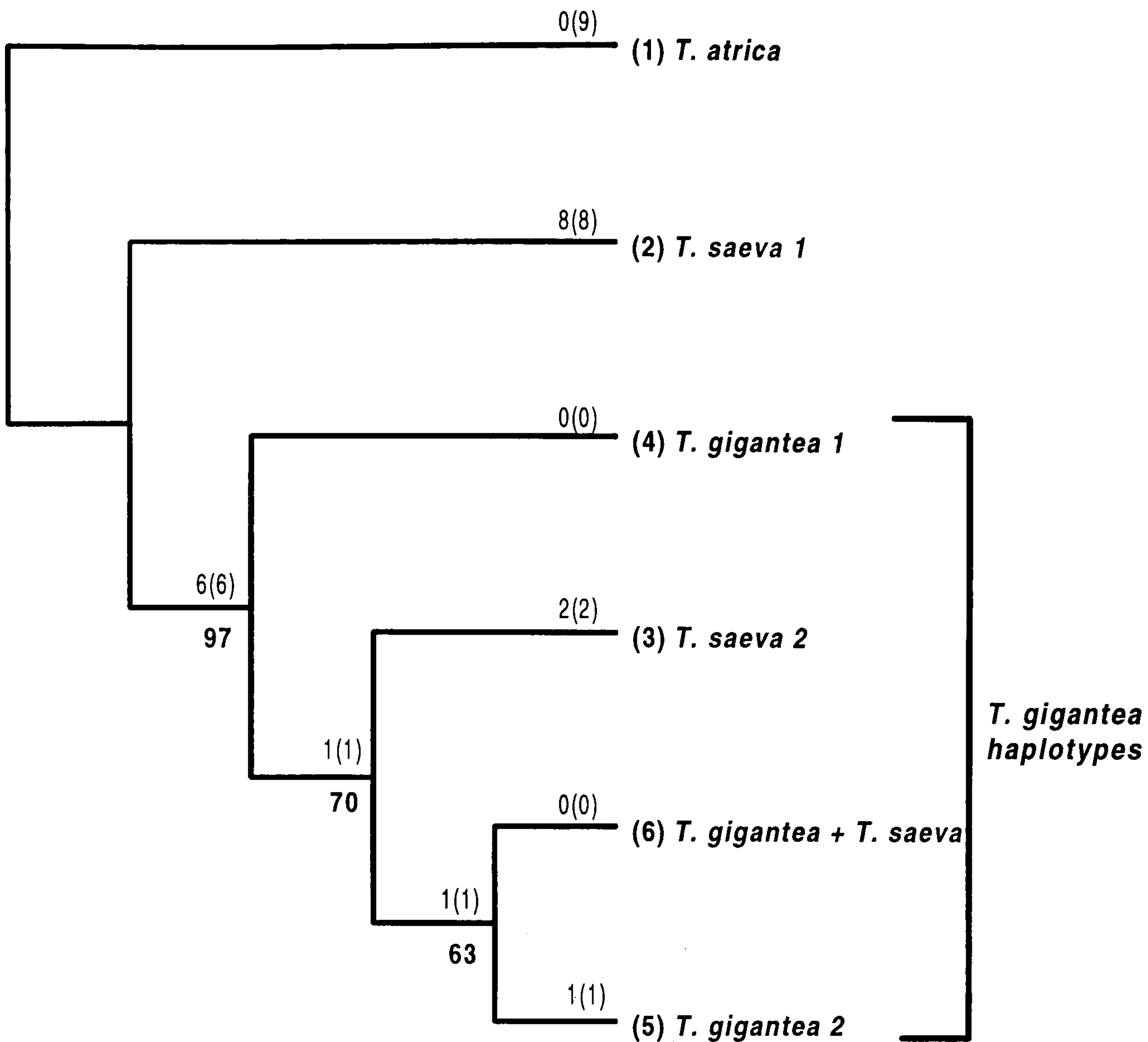


Figure 5.16. Maximum parsimony (MP) tree for the ND1 data. Bold numerals below the nodes indicate the percentage of 1000 bootstrap replicate samples recovering each clade. Numerals above the nodes and at branch tips indicate synapomorphies and autapomorphies respectively. The first number is the number of unambiguous changes, and number following, in parentheses, is the maximum possible number of changes allowed under parsimony. Haplotypes are numbered (before their names) in accordance with Table 5.12 (page 209). Haplotypes forming a distinct *T. gigantea* grouping have been bracketed.

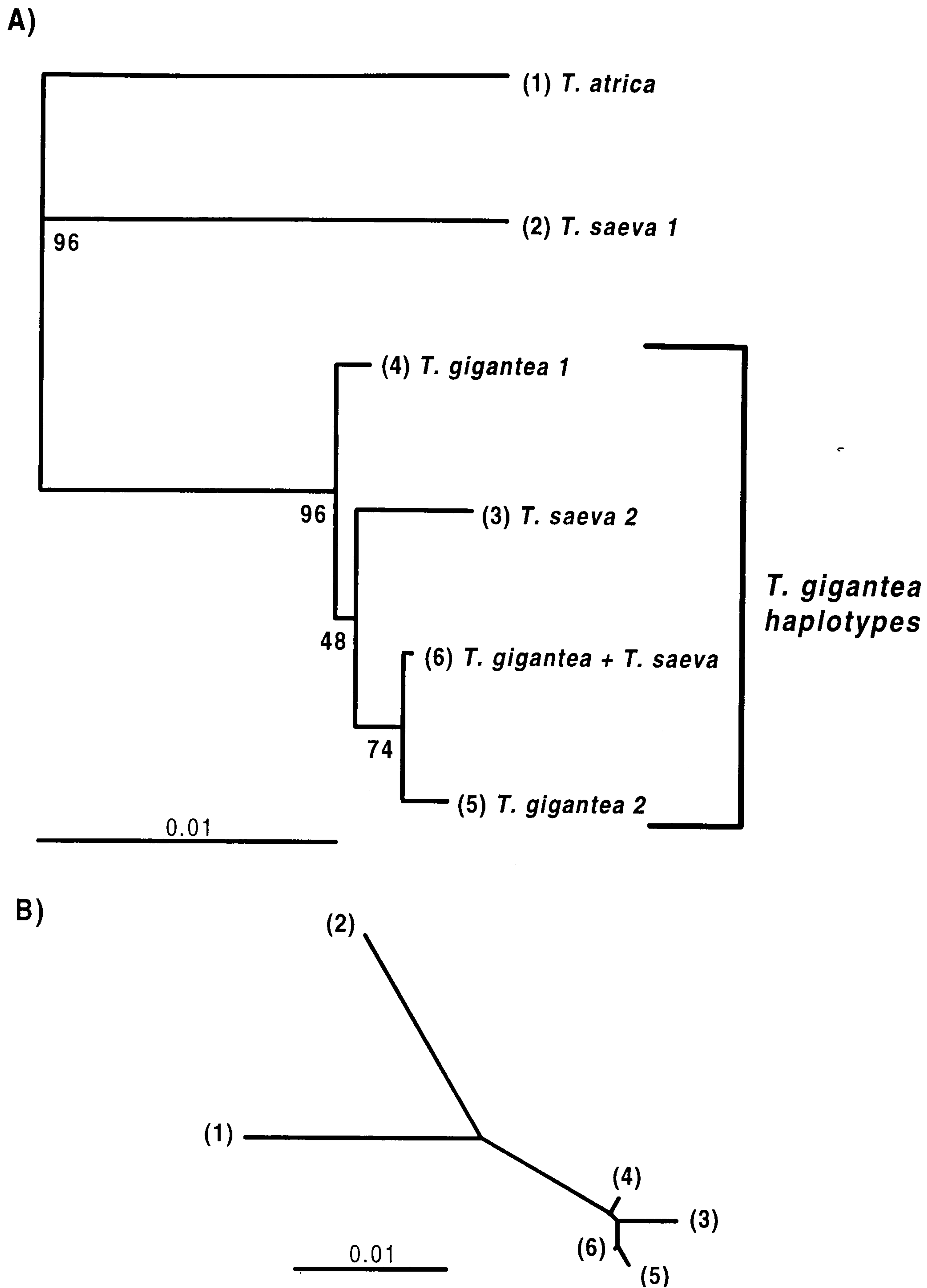


Figure 5.17. Neighbor-joining tree for the ND1 data. A) A phenogram rooted through the outgroup *T. denticulata*. Bold numerals below the nodes indicate the percentage of 1000 bootstrap replicate samples recovering each clade. Haplotypes are numbered (before their names) in accordance with Table 5.12 (page 209). Haplotypes forming a distinct *T. gigantea* grouping have been bracketed. B) An unrooted, radial representation of A to show the distances and groupings between haplotypes more clearly.

Neighbor-joining trees were also constructed from distance matrices generated using the different transition:transversion ratios as described in the parsimony section above. Changing the ratio had no effect on the topology, little effect on branch length, and did not significantly improve the bootstrap values (which were already high). UPGMA, maximum likelihood, and Fitch-Margoliash trees were also generated using PHYLIP (Felsenstein, 1995) with no change in tree topology.

All groupings were strongly supported by high bootstrap values with the exception of the node joining '*T. gigantea 1*' to the other three *T. gigantea* haplotypes, which is perhaps not surprising as these sequences were so similar. As in the parsimony tree, the four *T. gigantea* haplotypes ('*T. gigantea 1*', '*T. gigantea 2*', '*T. gigantea + T. saeva*', and '*T. saeva 2*') form a discrete group separate from '*T. saeva*' and '*T. atrica*'. This is most clearly seen in the unrooted, radial version of the tree (Figure 5.17 B).

5.3.4.6 Combined CO1 and ND1 Fragments - Parsimony Analysis

Parsimony analysis was performed on the 868 bp of combined CO1 and ND1 sequence for the six combined haplotypes of the *T. atrica* group. As for the ND1 analyses the trees were rooted through *T. atrica*.

The tree-length distribution for 10 000 random trees was significantly left-skewed ($g1$ = -1.470, $p < 0.01$, Hillis and Huelsenbeck, 1992), suggesting a strong phylogenetic signal in the data (see Figure 5.18).

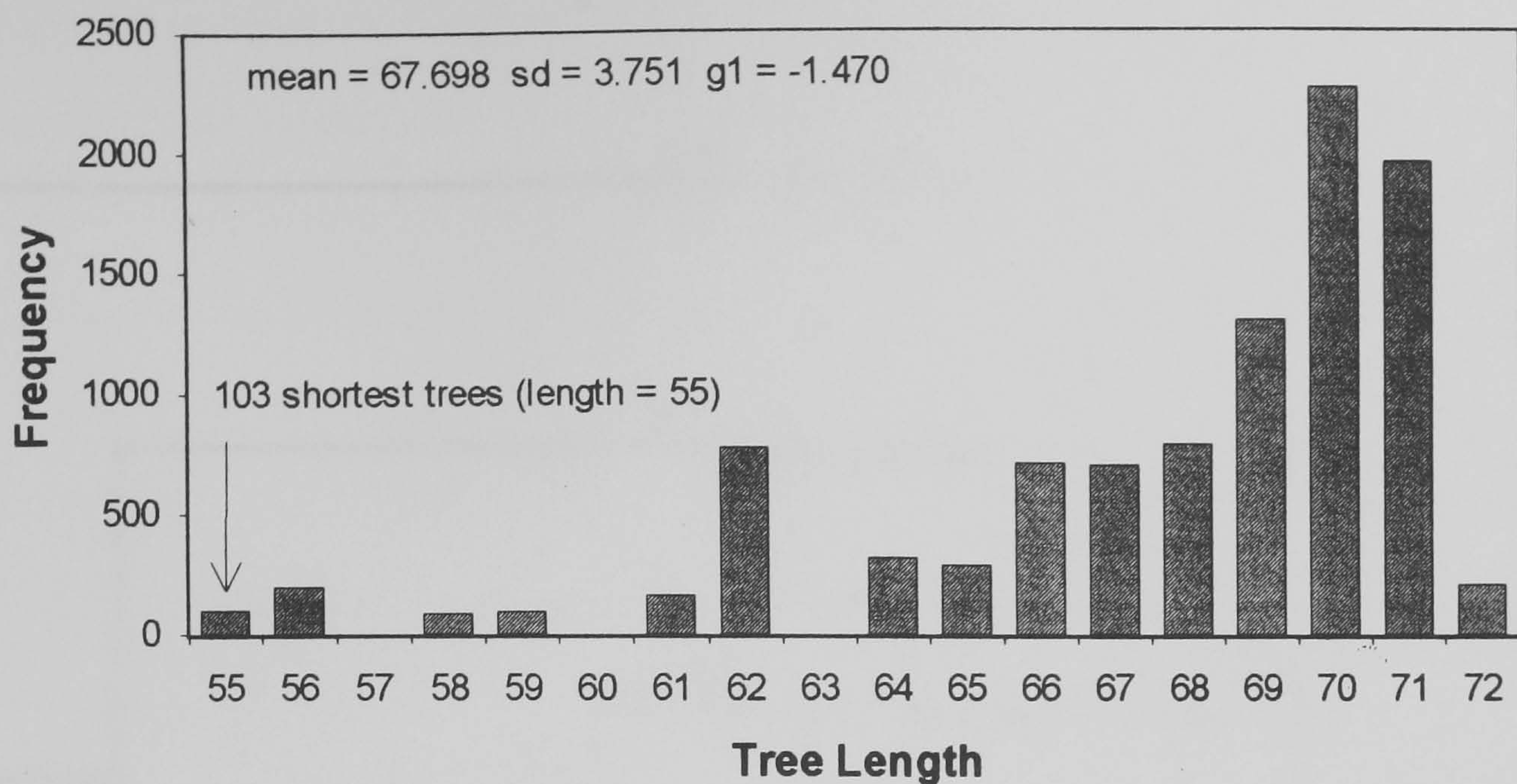


Figure 5.18. Frequency distribution of the lengths of 10 000 random trees for the combined CO1 and ND1 data. Generated using the six haplotypes of the *T. atrica* group.

An exhaustive search, in which all nucleotide sites were equally weighted, quickly returned a single MP tree of length 55 (CI = 0.909, HI = 0.091, RI = 0.773, RC = 0.702). The goodness of fit indices indicated a stronger fit than when the ND1 or CO1 data were analysed separately (although of course different haplotypes were involved, and hence different trees, so this comparison is not strictly valid). Groupings were supported by extremely high bootstrap values, suggesting a robust result; given this, the complications in combining fragments with different evolutionary patterns, and the lack of effect in the previous analyses, different character-weighting regimes were not attempted.

Examination of the MP tree in Figure 5.19, shows that the four haplotypes shared between *T. gigantea* and *T. saeva* (the "*T. gigantea* haplotypes") formed two clades. '*T. gigantea* + *T. saeva 1*' and '*T. gigantea* + *T. saeva 2*' were separated by two unambiguous autapomorphies in '*T. gigantea* + *T. saeva 1*'. '*T. gigantea* + *T. saeva 3*' and '*T. gigantea* + *T. saeva 4*' were separated by a

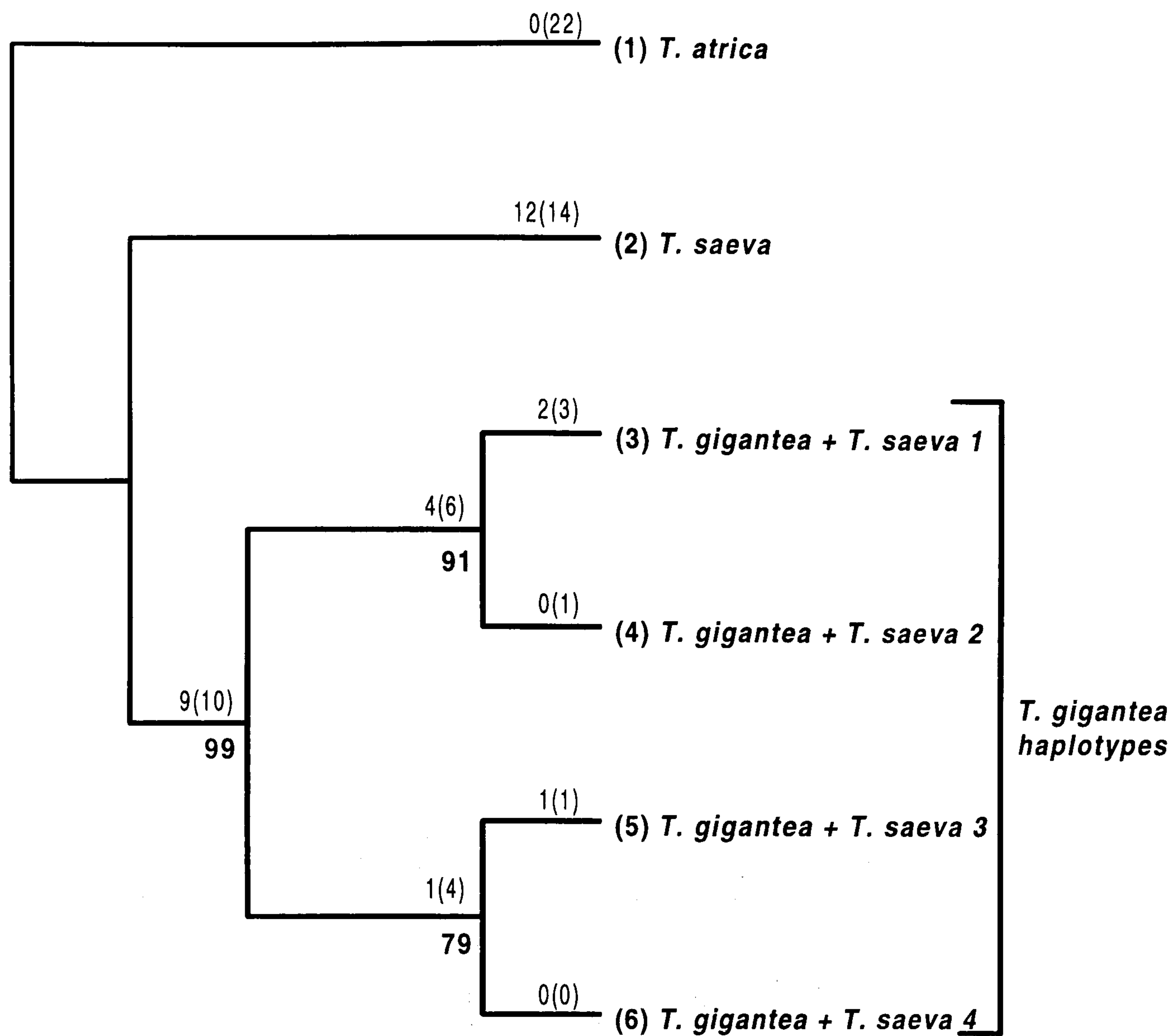


Figure 5.19. Maximum parsimony (MP) tree for the combined CO1 and ND1 data. Bold numerals below the nodes indicate the percentage of 1000 bootstrap replicate samples recovering each clade. Numerals above the nodes and at branch tips indicate synapomorphies and autapomorphies respectively. The first number is the number of unambiguous changes, and number following, in parentheses, is the maximum possible number of changes allowed under parsimony. Haplotypes are numbered (before their names) in accordance with Table 5.13 (page 210). Haplotypes forming a distinct *T. gigantea* grouping have been bracketed.

single unambiguous autapomorphy in '*T. gigantea* + *T. saeva 3*'. In total, five unambiguous synapomorphies separated these two clades from each other. However it is clear that these two clades group together to form a *T. gigantea* haplotypes clade which is separated from '*T. saeva*' by nine unambiguous synapomorphies, with '*T. saeva*' possessing 12 unambiguous autapomorphies. As the tree was rooted through '*T. atrica*' it is impossible to calculate the number of unambiguous changes separating this haplotype from the rest, all that can be said is that there were 22 steps ('ambiguous autapomorphies'). With the exception of the '*T. atrica*' haplotype, for the technical reason just described, the number of ambiguous changes was generally low.

5.3.4.7 Combined CO1 and ND1 Fragments - Distance Analysis

Figure 5.20 (A and B) shows neighbor-joining trees (Saitou and Nei, 1987) generated by CLUSTALW version 1.7 (Thompson *et al.*, 1994), from the Kimura two-parameter distance matrix shown in Table 5.13, for the 868 bp of combined CO1 and ND1 sequence for the six combined haplotypes of the *T. atrica* group. The topology of the tree was identical to that obtained by parsimony methods. As for the parsimony method, different weighting regimes were not implemented. Neighbor-joining, UPGMA, maximum likelihood, and Fitch-Margoliash trees were also generated using PHYLIP (Felsenstein, 1995) with no change in tree topology.

All groupings were strongly supported by extremely high bootstrap values. As in the parsimony tree, the four *T. gigantea* haplotypes ('*T. gigantea* + *T. saeva 1*', '*T. gigantea* + *T. saeva 2*', '*T. gigantea* + *T. saeva 3*', and '*T. gigantea* + *T. saeva 4*') form a discrete group separate from '*T. saeva*' and '*T. atrica*', but with '*T. gigantea* + *T. saeva 1*' and '*T. gigantea* + *T. saeva 2*', and '*T. gigantea* + *T.*

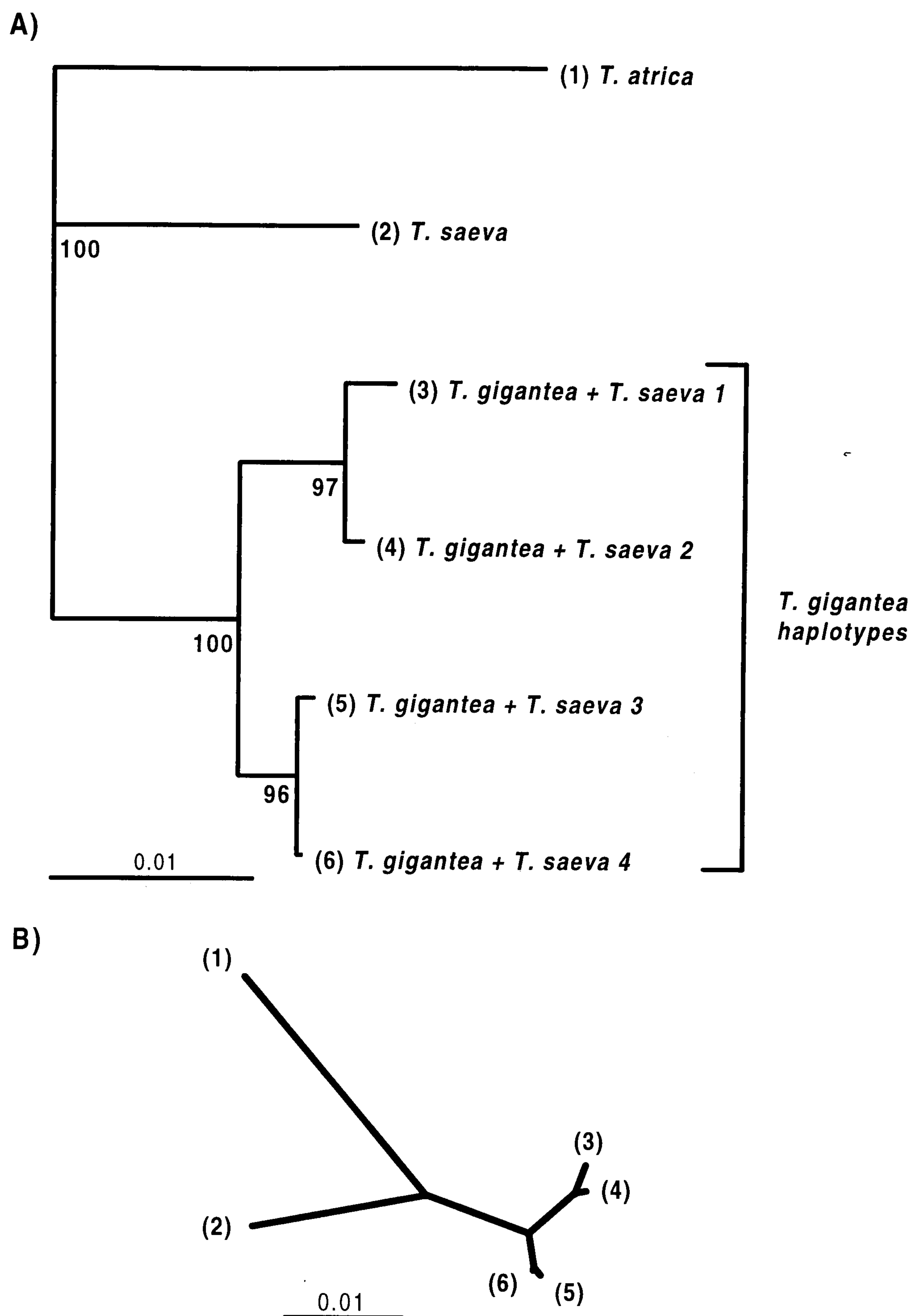


Figure 5.20. Neighbor-joining tree for the combined CO1 and ND1 data. A) A phenogram rooted through the outgroup *T. denticulata.* Bold numerals below the nodes indicate the percentage of 1000 bootstrap replicate samples recovering each clade. Haplotypes are numbered (before their names) in accordance with Table 5.13 (page 210). Haplotypes forming a distinct *T. gigantea* grouping have been bracketed. B) An unrooted, radial representation of A to show the distances and groupings between haplotypes more clearly.

saeva 3' and '*T. gigantea + T. saeva 4'*, forming two very distinct clades within this group. This is very clearly seen in the unrooted, radial version of the tree (Figure 5.20 B).

5.4 Results B: A putative pseudogene of CO1 in *Tegenaria parietina*

5.4.1 Sequence Characteristics

The PCR amplification of four specimens of *T. parietina* using the primers C1-J-1718 and C1-N-2191 (Nancy) to generate a 472 bp long fragment of the CO1 gene, resulted in two bands, one of which was significantly smaller than expected. Two of these animals were from different sites in southern England and two were from northern Italy. Direct manual sequencing of this mixed PCR product failed. Isolation of the bands, followed by reamplification and direct automated sequencing of three specimens, revealed that one specimen from England and the two specimens from Italy shared the same mitochondrial haplotype (as previously described). Another, shorter (349 bp), sequence was obtained from the fast (lower molecular weight) band. By careful isolation of the bands it was possible to determine that all three animals also shared an identical sequence for this fragment. Both fragments were short enough that the entire length, from primer to primer, was able to be sequenced in both directions and hence verified. The mitochondrial sequence was also verified by cloning (S. Mascheretti, pers. comm.). Unfortunately the anomalous sequence could not be readily cloned.

Alignment of the anomalous sequence to the entire mitochondrial CO1 fragment for *T. parietina* (corresponding to positions 1719 through 2190 of the *Drosophila yakuba* mtDNA (Clary and Wolstenholme, 1985)), and comparison with sequences in the EMBL database, showed that the anomalous fragment had great homology with the mitochondrial CO1 fragment. The full alignment with the *T. parietina* mitochondrial fragment is given in Table 5.15 (the numbering follows that of the functional mitochondrial fragment). The anomalous sequence showed a number of interesting features: 1) 53 bp from

Table 5.15. Alignment of the putative *T. parietina* CO1 pseudogene against the entire *T. parietina* CO1 fragment.

	1719																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
T. parietina CO1	T	T	T	A	A	T	G	T	T	A	G	G	G	G	C	G	C	C	T	G	A	T	A	T	A
T. parietina PsCO1	.	.	.	.	.	.	.	.	.	.	T	.	T	.	.	T	.	.	.	.	.	.	.	.	G

	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
T. parietina CO1	G	C	T	T	T	T	C	C	G	C	G	A	A	T	A	A	A	T	A	A	T	T	T	G	A
T. parietina PsCO1	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.

	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68							
T. parietina CO1	G	T	T	T	T	T	G	A	T	T	A	T	T	A	C	C	A	C							
T. parietina PsCO1	.	G	.	.	.	.	.	.	.	.	G	.	.	.	.	.	T	.	T	C	G	T	G	A	T
																			R	E	P	E	A	T	→

T. parietina CO1																									
T. parietina PsCO1	A	T	G	G	C	T	T	T	T	C	C	T	C	G	G	A	T	A	A	A	T	A	A	T	T

																			69	70	71	72	73	74	75
T. parietina CO1																			C	T	T	C	T	T	T
T. parietina PsCO1	T	G	A	G	G	T	T	T	T	G	A	T	T	G	T	T	A	C	.	.	C	.	.	.	.

	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
T. parietina CO1	G	T	T	T	A	T	G	C	T	A	T	T	C	A	T	T	T	C	T	T	C	T	A	T	G
T. parietina PsCO1	A	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.	.	.	.	.	.	.	-	.	A

	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125
T. parietina CO1	G	T	G	G	A	T	A	T	A	G	G	A	G	T	T	G	G	A	G	C	G	G	G	A	T
T. parietina PsCO1	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	A	.	.

Continued →

Table 5.15. Continued

	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
T. parietina CO1	G	G	A	C	T	A	T	T	T	A	T	C	C	T	C	C	T	T	T	G	G	C	G	T	C
T. parietina PsCO1	.	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	A	.	T	T	.	.

	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
T. parietina CO1	T	T	C	T	C	T	T	G	G	G	C	A	C	A	T	A	G	G	A	A	G	T	T	C	T
T. parietina PsCO1	.	.	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	A	G	.	A	A	.	.	G

	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
T. parietina CO1	A	T	G	G	A	T	T	T	T	G	C	G	A	T	T	T	T	T	T	C	T	T	T	A	C
T. parietina PsCO1	.	.	.	A	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	G	.

	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225
T. parietina CO1	A	T	T	T	A	G	C	T	G	G	G	G	C	T	T	C	T	T	C	A	A	T	T	A	T
T. parietina PsCO1	.	.	.	.	.	.	A	.	.	.	A	.	A	A	.	.	.	.	.	T	.	.	.	T	.

	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250
T. parietina CO1	A	G	G	G	G	C	T	A	T	T	A	A	C	T	T	T	A	T	T	T	C	C	A	C	A
T. parietina PsCO1	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	T	.	.	.

	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275
T. parietina CO1	A	T	T	A	T	T	A	A	T	A	T	A	C	G	T	T	C	G	A	T	C	G	G	A	A
T. parietina PsCO1	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	M	-	-	-	-	-	-	-	-

	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300
T. parietina CO1	T	A	A	G	A	A	T	A	G	A	G	A	A	G	G	T	T	C	C	T	T	T	A	T	T
T. parietina PsCO1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325
T. parietina CO1	T	G	T	T	T	G	A	T	C	T	G	T	T	T	T	A	A	T	T	A	C	T	G	C	T
T. parietina PsCO1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350
T. parietina CO1	A	T	T	T	T	G	T	T	A	T	T	A	T	T	A	T	C	A	C	T	T	C	C	A	G
T. parietina PsCO1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Continued →

Table 5.15. Continued

	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375
T. parietina CO1	T	T	T	T	G	G	C	T	G	G	A	G	C	G	A	T	T	A	C	G	A	T	G	T	T
T. parietina PsCO1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400
T. parietina CO1	G	T	T	G	A	C	T	G	A	T	C	G	A	A	A	C	T	T	T	A	A	T	A	C	T
T. parietina PsCO1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425
T. parietina CO1	T	C	A	T	T	T	T	T	T	G	A	C	C	C	T	G	C	T	G	G	A	G	G	A	G
T. parietina PsCO1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450
T. parietina CO1	G	G	G	A	T	C	C	T	A	T	T	T	T	A	T	T	T	C	A	A	C	A	T	T	T
T. parietina PsCO1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	.	G	.	.	.	.	.

																						2190
	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472
T. parietina CO1	A	T	T	T	T	G	G	T	T	T	T	T	T	G	G	T	C	A	T	C	C	G
T. parietina PsCO1	-	.	.	.	.	.	A	.	.	.	.	.	.	.	.	G	.	.	.	.	.	C

Nucleotide identity is shown by a dot, deletions are shown by dashes. Boxed numbers above the alignment refer to the equivalent position in the *D. yakuba* mtDNA molecule (Clary and Wolstenholme, 1995). Nucleotides for the mitochondrial (functional) molecule are numbered consecutively from 1 to 472. The repeated region and its repeat are shown by a solid and broken box respectively.

position 16 through 68 were immediately repeated - as a 53 bp insertion between position 68 and 69; 2) there was the deletion of an A at position 98; 3) there was a 175 bp deletion from position 268 through 442; 4) translation with the invertebrate mtDNA genetic code and the nuclear genetic code both failed to identify an open reading frame, suggesting that the anomalous sequence was non-functional.

Over the entire length of the alignment (ignoring insertions and deletions) there were 50 substitutions. 29 of these were transitions and 21 were transversions giving a transition:transversion ratio of 1.38:1. Considering only the more conserved (easier to align) region from position 69 through 266 there were 35 substitutions. 22 of these were transitions and 13 were transversions, giving a transition:transversion ratio of 1.69:1. For this region, 25 (71.4%) of the substitutions occurred in what would have been the third position for the mitochondrial fragment, 6 (17.1%) occurred in the second position, and 4 (11.4%) occurred in the first position. The pairwise distances for all the *Tegenaria* species plus the outgroup *T. denticulata* , together with the anomalous sequence (*T. parietina Ps*) were calculated for the region corresponding to positions 69 through 266 of the alignment in Table 5.15. These distances are shown in Table 5.16, where it can clearly be seen that, in terms of simple sequence divergence, the anomalous sequence differs from all the *Tegenaria* haplotypes, *including T. parietina*, by a large degree (19.04% to 21.62%). This suggests either an ancient origin for the anomalous sequence or a very different rate of evolution.

In terms of sequence divergence, the 53 bp repeat of positions 16 through 68 was interesting. In Table 5.17, positions 16 through 68 of the mitochondrial fragment have been aligned against the same region for the anomalous fragment, and against the repeat of this region. It can be clearly seen that although the anomalous fragment differed from the mitochondrial fragment by

seven substitutions (four transversions and three transitions), the repeat shows great homology to the same region of the anomalous fragment. Indeed, the repeat differs by only a single transversion. This suggests a recent origin for the repeat; an ancient origin is highly unlikely as it is improbable that two tandem copies of the same region of sequence would follow an identical pattern of substitutions independently.

Table 5.16. Pairwise distances among the nine mt and one pseudogene CO1 haplotypes analysed. Calculated from nucleotide positions 69 through 266 of the pseudogene alignment (Table 5.15). Above the diagonal: pairwise nucleotide sequence divergence estimated using the Kimura two-parameter model. Below diagonal: absolute number of nucleotide differences between sequences.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) *T. atrica*	-	0.0573	0.0683	0.0573	0.0681	0.1034	0.2285	0.2456	0.1954	0.1682
(2) *T. saeva*	11	-	0.0414	0.0308	0.0413	0.0919	0.2146	0.2112	0.2020	0.1619
(3) *T. gigantea*	13	8	-	0.0205	0.0051	0.0860	0.2155	0.2120	0.2162	0.1437
(4) *T. gigantea + T. saeva 1*	11	6	4	-	0.0204	0.0863	0.2146	0.2112	0.2020	0.1493
(5) *T. gigantea + T. saeva 2*	13	8	1	4	-	0.0863	0.2146	0.2045	0.2154	0.1370
(6) *T. agrestis*	19	17	16	16	16	-	0.1942	0.1896	0.1904	0.1431
(7) *T. domestica*	37	35	35	35	35	32	-	0.2112	0.2020	0.1970
(8) *T. parietina*	40	35	35	35	34	32	35	-	0.2054	0.1095
(9) *T. parietina Ps*	33	34	36	34	36	32	34	35	-	0.1663
(10) *T. denticulata*	29	28	25	26	24	25	33	20	29	-

Table 5.17. Comparison of the 53 repeated base pairs in the CO1 pseudogene.

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
T. parietina CO1	G	C	C	T	G	A	T	A	T	A	G	C	T	T	T	T	C	C	G	C	G	A	A	T	A
T. parietina PsCO1	T	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.
PsCO1 Repeat	T	.	G	.	.	.	.	.		G	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65
T. parietina CO1	A	A	T	A	A	T	T	T	G	A	G	T	T	T	T	T	G	A	T	T	A	T	T	A	C
T. parietina PsCO1	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	G	.	.	.	.
PsCO1 Repeat	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	G	.	.	.	.

	66	67	68
T. parietina CO1	C	A	C
T. parietina PsCO1	.	T	.
PsCO1 Repeat	.	T	.

The table shows great sequence homology of the repeat to itself but not to the functional mitochondrial gene - suggesting a recent duplication event post pseudogene genesis.

5.4.2 Relative rate tests of CO1 pseudogene evolution.

In order to determine whether the perceived divergence of the anomalous sequence was due to time of divergence or a different rate of evolution, a relative rate test (Mindell and Honeycutt, 1990) was performed for the region corresponding to positions 69 through 266 of the alignment in Table 5.15. None of the comparisons was significant suggesting a similar rate of evolution for the CO1 fragment across species (as already shown in Section 5.3.4.3), and a similar rate of evolution for the anomalous fragment. Note however, as in Section 5.3.4.3, the probability values for *T. domestica* were in general lower than the others suggesting a possibly slightly higher rate of evolution for this species.

Table 5.18. Relative rate test for the CO1 pseudogene.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) *T. atrica*	-	6/5	9/5	7/4	9/4	13/9	19/23	26/17	18/17
(2) *T. saeva*	1.0000	-	6/3	4/2	6/2	13/8	17/23	23/15	18/20
(3) *T. gigantea*	0.4240	0.5078	-	2/3	1/0	9/9	17/25	21/14	18/20
(4) *T. gigantea + T. saeva 1*	0.5488	0.6875	1.0000	-	3/1	10/9	17/24	22/16	16/20
(5) *T. gigantea + T. saeva 2*	0.2668	0.2891	1.0000	0.6250	-	8/9	16/25	21/16	16/20
(6) *T. agrestis*	0.5235	0.3833	1.0000	1.0000	1.0000	-	16/24	19/15	16/19
(7) *T. domestica*	0.6440	0.4296	0.2800	0.3489	0.2110	0.2682	-	26/13	22/18
(8) *T. parietina*	0.2221	0.2559	0.3105	0.4177	0.5114	0.6076	0.0533	-	15/24
(9) *T. parietina Ps*	1.0000	0.8714	0.8714	0.6177	0.6177	0.7359	0.6358	0.1996	-

The numbers above the diagonal are the number of unique substitions in pairs of haplotypes with respect to the outgroup. The first number refers to the haplotypes in the first column, the second number to the haplotypes in the first row. The numbers below the diagonal are the binomial probabilities.

5.4.3 Phylogenetic analysis

Parsimony analysis was performed on the 198 nucleotides corresponding to positions 69 through 266 of the alignment in Table 5.15, for the anomalous fragment, the eight *Tegenaria* haplotypes, plus *T. denticulata* as an outgroup.

The tree-length distribution for 10,000 random trees was significantly left-skewed ($g1 = -0.666$, $p < 0.01$, Hillis and Huelsenbeck, 1992), suggesting a strong phylogenetic signal in the data (see Figure 5.21).

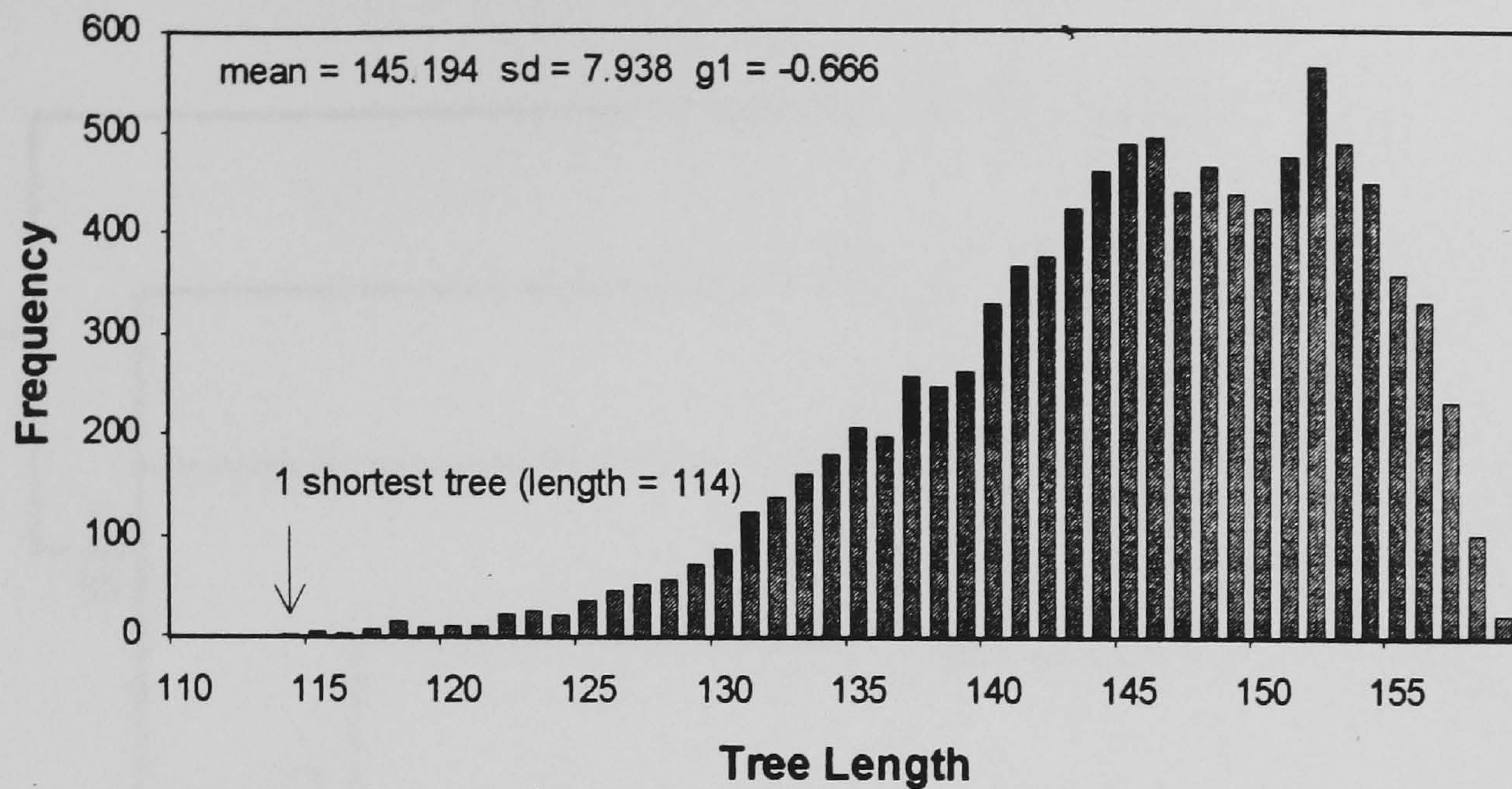


Figure 5.21. Frequency distribution of the lengths of 10,000 random trees for the CO1 pseudogene data. Generated using the eight *Tegenaria* haplotype, the outgroup *Tetrix denticulata*, and the putative pseudogene.

An exhaustive search, in which all nucleotide sites were equally weighted, returned two MP trees. Both trees were identical in terms of number of steps (length = 109, 5 steps shorter than the shortest random tree), and both were identical in terms of goodness of fit measures (CI = 0.789, HI = 0.211, RI = 0.689, RC = 0.544). In both the trees, *T. parietina Ps* (the anomalous sequence) grouped with *T. domestica*, although the trees disagreed whether *T. parietina Ps* and *T. domestica* should form a separate clade of their own, or whether *T. parietina Ps* should fall between *T. domestica* and *T. agrestis*. The tree in Figure 5.22 shows the consensus tree for these two MP trees. The position of *T. parietina Ps* has been left as an unresolved polytomy. Reweighting regimes, as previously described, did not alter the result appreciably. The relationships were not resolved by including more distant taxa (*Drosophila yakuba* and *Tetragnatha quasimodo*) in the analyses (data not shown). The inclusion of *T. parietina Ps* lowered the bootstrap support for the tree, relative to the CO1 tree presented earlier, however all groupings were the same. Because of the polytomy it was impossible to calculate the maximum possible numbers of

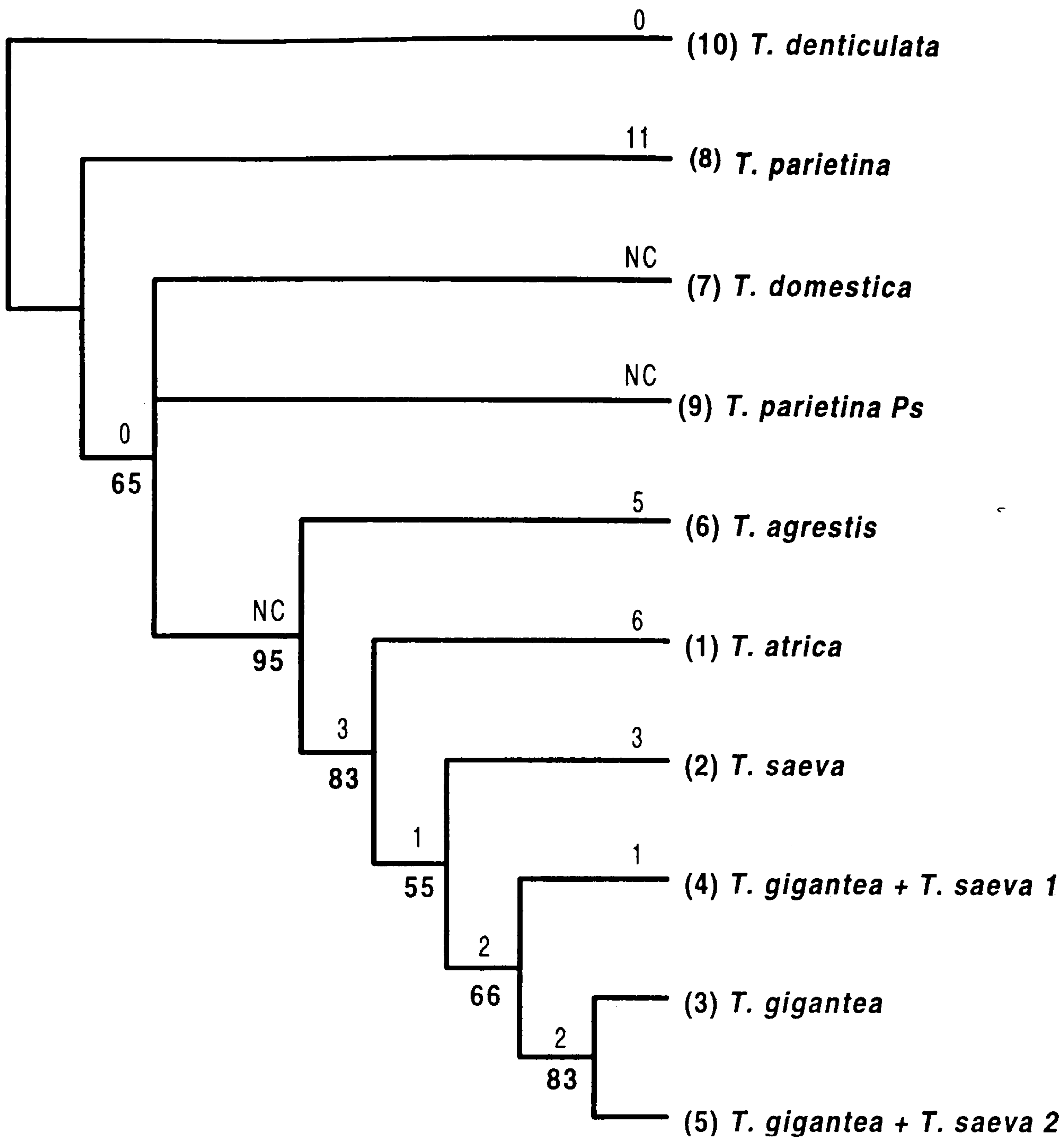


Figure 5.22. Maximum parsimony (MP) tree for the CO1 pseudogene data. Bold numerals below the nodes indicate the percentage of 1000 bootstrap replicate samples recovering each clade. Numerals above the nodes and at branch tips indicate synapomorphies and autapomorphies respectively. Only the number of unambiguous changes could be calculated because of the polytomy. N.C. = unambiguous changes not calculated across the polytomy. Haplotypes are numbered (before their names) in accordance with Table 5.16 (page 233).

changes on each branch, nor was it possible to calculate the number of unambiguous changes within the polytomy itself. However it was shown, in section 5.3.4.1, that the number of ambiguous changes in the outgroup, *T. parietina*, and *T. domestica* was large: reflecting uncertainties/polymorphisms (homoplasies) due to the distance of these taxa, possible saturation and reversals. Without more intermediate taxa the position of *T. parietina Ps*, and whether it should group with *T. domestica* or *T. parietina*, can not be clearly stated. Distance and maximum likelihood based approaches (not shown) also supported the placement of *T. parietina Ps* with *T. domestica*, although bootstrap support for this was always low.

5.5 Discussion

5.5.1 Phylogenetic Inferences and Introgression

Within Europe, including European Russia and Turkey, there are around 18 genera and at least 228 described species belonging to the family Agelenidae. Around 97 (approximately 40%) of these species belong to the genus *Tegenaria* (figures calculated from Maurer, 1992). These figures are only approximate as the phylogenetic relationships and taxonomic placements of not only the species but the genera are still very uncertain. Indeed there are probably many more species yet to be discovered, names to be synonymized, and genera to be moved to different families within Europe alone. (See Chapters 1.0 and 2.0 for more discussion). Clearly there is a great need for a full-scale examination of the Agelenidae in terms of a traditional cladistic, morphological analysis. The results presented here, though making little impact on the problem, suggest that molecular data gained from the CO1, 16S and ND1 gene fragments could greatly augment morphological work on the relationships of these spiders. Wilcox *et al.* (1997) in their work on morphologically cryptic species-complexes of pseudoscorpions suggest that mtDNA CO1 sequence data "may give some valuable clues to what the higher categories (subfamilies, tribes, etc.) will include". The data presented here suggest that CO1 is indeed useful for elucidating the relationships between species within genera. However, relationships between genera and families may well be better served by the less variable 16S fragment. This , in concert with the contiguous ND1 fragment, could prove to be a powerful tool in the phylogeny reconstruction of spiders from species to family level. The primers for all of these fragments appear to be widely conserved and of general utility. The ND1 fragment has now been successfully amplified in such diverse spider families as the Agelenidae (this work), the Nesticidae and Salticidae (Hedin, 1997), and the Atypidae (J. Johannesen, pers. comm.). The 16S fragment has

also been employed in the Nesticidae and Salticidae (M. Hedin, pers. comm.), and many other invertebrate and vertebrate taxa (Simon *et al.*, 1994). The CO1 sequence has been amplified in a wide-range of invertebrates including pseudoscorpions (Wilcox *et al.*, 1997) and the spider genera *Tetragnatha* (R. Gillespie, pers. comm.) and *Dysdera* (M. Arnedo and C. Ribera, pers. comm.). The primers are not very different from those available for vertebrates (Simon *et al.*, 1994). The presence of a putative pseudogene of CO1 in *T. parietina* might suggest caution in using this gene for phylogeny reconstruction and highlights the need to be aware of the possible existence of pseudogenes in *all* mtDNA studies.

The CO1 data presented here, within the limits of the species sampled, strongly support the existence of a *T. atrica* group clade and thus ratify the traditional morphologically based grouping of these species as sister-species. Both the phylogenetic trees for CO1 in Figures 5.13 and 5.14 , and the charts showing the hierarchical levels of CO1 divergence in Figure 5.11, illustrate this clearly. The limited number (three) of other species sampled permits little detailed discussion of their relationships. These three species represent different and divergent morphological species groups and it would interesting to add more species to the trees in order to explore these groupings. The bootstrap values and branch lengths (i.e. distances) indicate well supported and deep divisions within the genus *Tegenaria*.

All the analyses - the CO1, ND1, and combined CO1 and ND1 data - show a clear and well supported distinction between *T. atrica* and its sister-species. However, between *T. gigantea* and *T. saeva* the distinction is not so immediately clear. All the analyses reveal one haplotype corresponding to *T. saeva* that is distinct and divergent (in terms of evolutionary distances and numbers of unambiguous changes on the MP tree) from all other haplotypes involving these two species. The remaining haplotypes are closely related and

can be regarded as *T. gigantea* haplotypes. Three such haplotypes were identified in the CO1 data, two of which were common to both *T. gigantea* and *T. saeva*. For the ND1 data, four similar haplotypes were identified, two of which were only found in *T. gigantea*, one in both *T. gigantea* and *T. saeva*, and one only in *T. saeva* (though different from the 'real' *T. saeva* haplotype by some 17 unambiguous changes). The combination of the CO1 and ND1 data revealed four similar haplotypes all of which were common to both species. The overall similarity of these haplotypes and their divergences from the *T. saeva* and *T. atrica* haplotypes strongly suggests that they are of a common *T. gigantea* origin and result from introgression of the *T. gigantea* mitochondrial DNA into *T. saeva* populations. The identification of a highly divergent *T. saeva* haplotype (and lack of this haplotype within *T. gigantea* populations) argues against the alternative explanation that this represents the maintenance of an ancestral polymorphism within *T. saeva*. No individuals of *T. gigantea* exhibited a haplotype that could be attributed to *T. saeva*, though sample sizes were small and absence of evidence is not evidence of absence. Of the CO1 haplotypes, five out of nine individuals identified as *T. saeva* exhibited a *T. gigantea* haplotype (approximately 56%, excluding the two individuals from York, where levels of hybridization are known to be high, reduces this figure to 43%). For the ND1 data, two out of five individuals identified as *T. saeva* exhibited a *T. gigantea* haplotype (40%). If these values are representative then introgressed *T. gigantea* mtDNA in *T. saeva* populations occurs frequently. It is not possible to say whether this pattern represents past or current introgression as the rate of evolution is not fast enough to elucidate such fine time scales. Recombinant individuals came from a broad geographic area (see Table 5.1) and were found far away from the area of parapatry in Dorset (with one such individual from Trewen, Cornwall some 175 km west of the hybrid zone). These individuals are unlikely to be the product of *very* recent hybridization. However, the fact that the haplotypes of *T. gigantea* origin found in *T. saeva* individuals are also generally found, unmodified, in current *T.*

gigantea populations suggests against these representing very ancient events. A more extensive sequencing survey from southern Europe, where one might expect to find a greater haplotype diversity (Hewitt, 1996), may be able to distinguish between recent and ancient introgression events.

5.5.2 Divergence

Brower (1994), in an assessment of mitochondrial CO1 divergence in a variety of invertebrate taxa, inferred a rate of 2.3% per million years. This appears to be a fairly robust estimate and agrees with that inferred previously in *Dolichopoda* cave crickets by Venanzetti *et al.* (1993). This rate was also applied to arachnids by Wilcox *et al.* (1997). Accepting that this is only an estimate and assuming that it is broadly applicable to the genus *Tegenaria*, approximate times of divergence may be calculated. Of course it is important to exercise caution in interpreting these estimates because they are based on fairly short amounts of sequence and few individuals/haplotypes, and because the molecular clock is inferred from data for other taxa.

Pairwise distances for the CO1 data (see Table 5.10) had a maximum divergence of 24.78% and therefore estimates should all fall within the portion of linear relationship between sequence divergence and divergence time (beyond about 30-40 % divergence the curve begins to plateau as variable sites become saturated with mutations) (Avise, 1994; Brown *et al.*, 1979; Moritz *et al.*, 1987). No pairwise comparisons failed the relative rate test, suggesting that the rate of CO1 evolution across the taxa studied was similar. There were three CO1 haplotypes grouped together in what can be considered as a *T. gigantea* group of haplotypes. These three *T. gigantea* haplotypes had a mean divergence from the *T. saeva* haplotype of 3.13% (range = 2.81% - 3.45%) giving a mean divergence time for *T. gigantea* and *T. saeva* of 1.4 million years ago (Ma) (range = 1.2 Ma - 1.5 Ma). The mean pairwise distance of the

T. gigantea/*T. saeva* clade from *T. atrica* was 5.72% (range = 5.06% - 6.37%) giving a mean divergence time of 2.5 Ma (range = 2.2 Ma - 2.8 Ma). Although one must allow for latitude in these estimates, they suggest that the radiation of the *T. atrica* group occurred around the beginning of the Pleistocene Epoch, and hence at the beginning of the Quaternary Period. The start of the Quaternary (the Pliocene-Pleistocene boundary) is marked in the fossil record by the "point in the stratigraphic column where faunal and floral elements indicate an abrupt change from warm to cold conditions" (Lowe and Walker, 1984). This point corresponds to the first major cold pulse of the Quaternary and the onset of the modern age of glacial-interglacial cycles mediated by Milankovitch processes (Hewitt, 1996; Lowe and Walker, 1984). There have been 19 such cycles in the last 700 000 years alone with as many as 21 during the entire Quaternary. No strict agreement for the exact date of the Pliocene-Pleistocene boundary has been achieved but upper estimates place it at more than 2.4 Ma (Lowe and Walker, 1984) when the ice-sheets in the northern hemisphere first began to grow large (Hewitt, 1996; Webb and Bartlein, 1992). Figure 5.23 illustrates the climatic stages of the Quaternary Period in Europe as deduced from stratigraphies in the Netherlands, together with estimates of mean summer temperatures and dates (redrawn from Lowe and Walker, 1984). The divergence times for *T. gigantea* and *T. saeva* and for *T. atrica* and the *T. gigantea/T. saeva* clade are shown. There is a remarkable correspondence between the divergence date for *T. atrica* and the *T. gigantea*/*T. saeva* clade and the first glaciation of the Pleistocene (Praetiglian), and also between the divergence of *T. gigantea* and *T. saeva* and the second major glaciation (Eburonian). This may suggest that speciation in the *T. atrica* group may have been driven by allopatric divergence and genome reorganisation in refugia during these two ice-ages with the likely formation of hybrid zones, and possible associated processes such as reinforcement during the interglacials (Hewitt, 1996). The subsequent ranges and genetic make-up of these species is likely to have been sculpted by these and subsequent climatic oscillations.

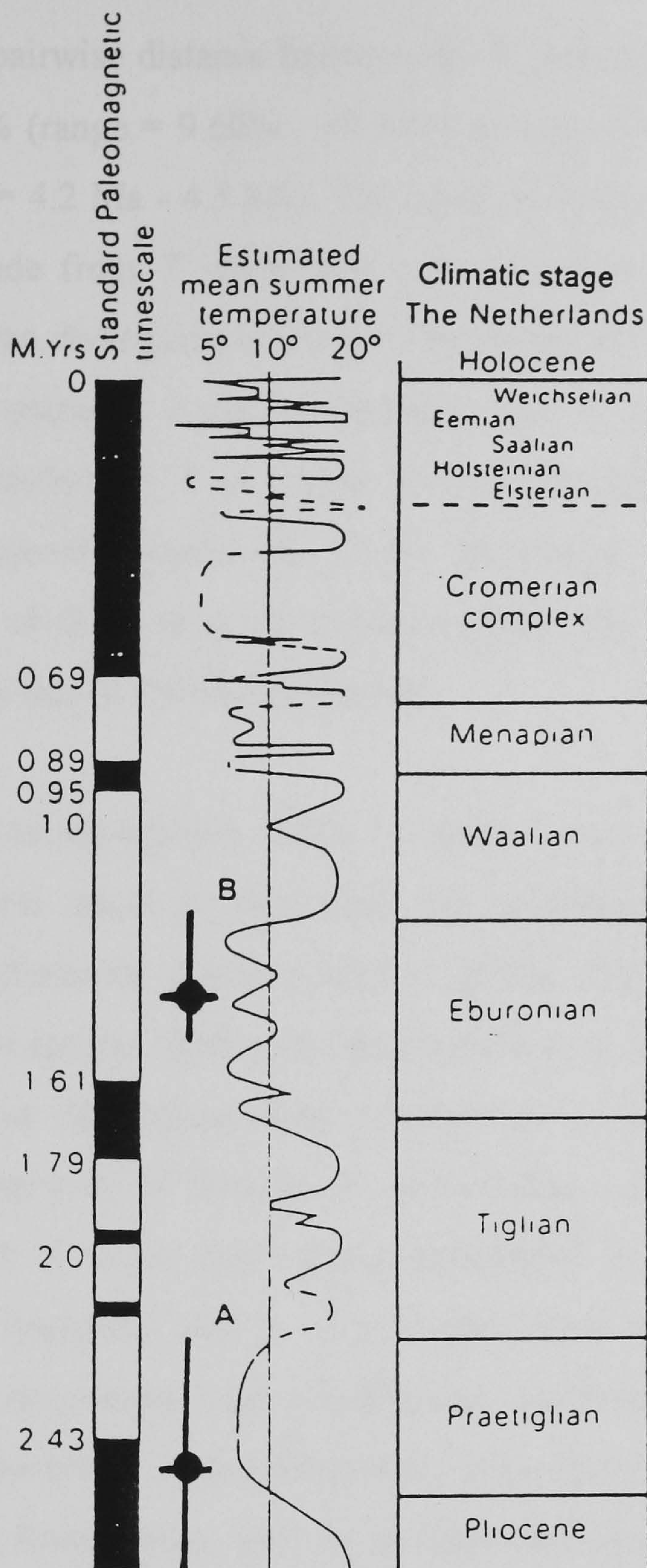


Figure 5.23. European Quaternary climatic stages and divergence dates for the *T. atrica* group from mitochondrial CO1 divergence data. Given, against the standard palaeomagnetic timescale, are the estimated mean summer temperatures and the climatic stages (glacials and interglacials) derived from stratigraphies in the Netherlands. **A**: The estimated divergence time of the *T. gigantea/T. saeva* clade from *T. atrica*. **B**: The estimated divergence time of *T. gigantea* and *T. saeva*. The vertical line represents the range of the time estimates (maximum and minimum). The intersection of the horizontal line and the vertical line indicates mean time estimate (A = 2.5 Ma = Praetiglian. B = 1.4 Ma = Eburonian). (Redrawn with modification from Lowe and Walker, 1984).

The mean pairwise distance between the *T. atrica* group clade and *T. agrestis* was 10.02% (range = 9.60% - 10.30%) giving a mean divergence time of 4.4 Ma (range = 4.2 Ma - 4.5 Ma). The mean divergence of the *T. atrica* group/*T. agrestis* clade from *T. domestica* was 21.82% (range = 21.30% - 22.55%) giving a mean divergence time of 9.5 Ma (range = 9.3 Ma - 9.8 Ma). Similarly, the mean pairwise distance between the *T. atrica* group/*T. agrestis*/*T. domestica* clade and *T. parietina* was 22.70% (range = 19.93% - 24.78%) giving a mean divergence time of 9.9 Ma (range = 8.7 Ma - 10.8 Ma). Thus placing all of these deep divergences within the *Tegenaria* in the Pliocene Epoch at the end of the Tertiary Period.

The molecular divergence of the *T. atrica* group contrasts with an apparent morphological stasis in that they are indistinguishable except for subtle genitalic differences. Another feature of the data was the lack of variation found within species. Although sample sizes were small, no sequence variation was found in the cosmopolitan synanthrope *T. domestica* (three individuals from different parts of Britain); *T. atrica* (three individuals from Britain, Eire, and France); *T. saeva* (six 'non-recombinant' individuals from France and south-west England); and in *T. parietina* (both mitochondrial and putative pseudogene sequences in three individuals; one from Kent and two from Italy). *Tegenaria parietina* is predominantly a southern European species and its presence in Britain may well be through introduction in previous centuries, perhaps from northern Italy, so perhaps the lack of sequence divergence between northern Italy and Britain is not so surprising. Curiously, *T. gigantea* did show sequence variation, the haplotypes differing by between 0.62% and 1.87%. Overall, the amount of sequence variation within the members of the genus *Tegenaria* studied here varied from 0% to 24.78%. These levels of variation agree with those reviewed by Simon *et al.* (1994) in insects where cytochrome oxidase variation ranges from generally less than 1% between individuals of the same species to 25% divergence between congenerics. Wilcox *et al.* (1997) point out that such high levels of sequence divergence far

exceed those shown by congeneric species of birds (as described by Seutin *et al.*, 1993), though perhaps this has more to do with the differing judgements of invertebrate and vertebrate sytematicists on what constitutes a genus.

5.5.3 The putative pseudogene of CO1 in T. parietina

The full characterization of a pseudogene requires extensive study and falls outside the scope of this work. Such studies should include techniques to verify the location of the pseudogene in the nuclear genome and hence may include PCR amplification from purified mitochondrial DNA and purified nuclear DNA, designing internal primers to the pseudogene and sequencing the flanking regions (perhaps by inverse PCR), and fluorescent *in situ* hybridization to metaphase preparations using the pseudogene as a probe. Perhaps the main feature of the putative pseudogene in *T. parietina* of relevance to this work is its presence. The presence of nuclear copies of mitochondrial genes requires recognition if they are not to confound phylogenies based on these mitochondrial genes. Even when their presence is acknowledged they can still present enormous difficulties in PCR, as they are often more ancestral in sequence and may therefore even amplify in preference to the desired product when using conserved primers - making cloning of PCR products a prerequisite. This is particularly likely to be a problem in vertebrate studies where the rate of nuclear evolution may be up to ten times slower than in the mitochondria

There are three possible explanations for this anomalous sequence. First, it may indeed be a pseudogene located in the nucleus; secondly, it could represent an alternative allele of the same gene on an alternative mitochondrial molecule (heteroplasmy); or thirdly, it could be a non-functional intra-mitochondrial duplication. No open reading frame was found when the anomalous sequence

was translated with either the invertebrate mitochondrial or nuclear genetic codes, yet the sequence showed great homology to the mitochondrial CO1 gene; this strongly suggests that it is a non-functional copy of the mitochondrial gene. This non-functionality makes heteroplasmy unlikely as an explanation. A mitochondrion lacking CO1 activity would be non-functional and deleterious to the carrier. Given the rapid fixation of mitochondrial haplotypes (Avise *et al.*, 1994) the retention of a non-functional genome seems unlikely. The third possibility seems implausible as mitochondria are generally poor in 'junk' DNA and are constrainted in the amount of DNA that they can tolerate to maintain efficient replication (Mirol, 1996; Moritz *et al.*, 1987). An inefficiently replicating mitochondrion would probably be eliminated relatively quickly.

For brevity I shall not discuss the possible mechanisms by which mitochondrial genes may find their way into the nuclear DNA but simply state that such pseudogenes have been described in a wide variety of taxa (see Zhang and Hewitt, 1996, for a review) and provide support for the serial endosymbiosis theory of Margulis (1981) whereby mitochondria, chloroplasts, and possibly other organelles were formed by the inclusion of free-living prokaryotes into the eukaryotic cell and subsequent (on-going) removal of genetic control to the nucleus.

The putative pseudogene did not fail the relative rate test, suggesting that it has been evolving at a similar rate to the functional CO1 genes. This result was not unexpected since it is known that the invertebrate mitochondrial DNA molecule does not evolve particularly more rapidly than the nuclear DNA (Crozier, 1993; Harrison, 1989; Powell *et al.*, 1986). The pseudogene was approximately as equally divergent from all the *Tegenaria* as it was from *T. parietina* with pairwise distances (see Table 5.16) ranging from 19.04% to 21.62% (mean = 20.36%), giving a divergence time between 8.3 Ma and 9.4 Ma (mean = 8.9 Ma). Clearly the pseudogene is ancient and from the

phylogenetic reconstruction it is not possible to say when it formed in relation to the other *Tegenaria*. It may be present in all or many of the species examined, and may not amplify simply because the primer recognition sites have been obliterated by mutation, or - and perhaps more plausibly - it only occurred in the lineage leading to *T. parietina* and examination of other species in the *T. ferruginea* group (to which *T. parietina* belongs) would reveal the anomalous sequence. Perhaps one of the most interesting features of the anomalous sequence was the repeat. The lack of variation between the two sections of the repeat suggests that it, unlike the anomalous sequence as a whole, is very recent. It is therefore intriguing to speculate whether species closely related to *T. parietina* would possess the anomalous sequence but lack the repeat.

6 Conspecific and Heterospecific Crosses: Comparative Analyses of Courtship Behaviour and Reproductive Success in *T. saeva* and *T. gigantea*.

6.1 Introduction

6.1.1 Spider Mating Behaviour

The mating behaviour of spiders can be divided into three phases: sperm induction (filling the palps via a 'sperm-web' onto which sperm has been deposited from the genital opening), courtship and copulation (Platnick, 1971). Sperm induction is not a strict precondition for courtship (Platnick, 1971; Foelix, 1996), and is not of particular interest here. This chapter focuses on courtship and, to a lesser extent, the mechanics of copulation. In animals that repeatedly copulate or show post-copulatory mate-guarding, as is frequent in spiders (see for instance, Fahey and Elgar, 1997; or reviews in Elgar, 1995; 1997), post-copulatory courtship may be observed. However, the focus here is on pre-copulatory courtship and its possible role in prezygotic reproductive isolation. Behavioural barriers to hybridization should act before copulation occurs (though the possibility of post-copulatory manipulation of sperm by the female, by behavioural or physiological means, cannot be excluded (Eberhard, 1985)). Further, one would expect mechanisms of mate discrimination to act as early as possible if costs such as 'time out of the mating pool' and injurious encounters are to be minimized.

6.1.2 Courtship

Courtship can be defined as those behavioural patterns which are preparatory to mating (Foelix, 1996). The functions of courtship in spiders have been defined

by Platnick (1971), using Tinbergen's (1954; in Platnick, 1971) analysis of courtship as a two-way releaser system (display in A releases response in B releases response in A, and so on). This two-way interaction renders the interactants into 'units of a super-individual order'. The functions of spider courtship are:

1) To synchronize mating activities.
2) To orient the individuals.
3) To suppress non-sexual tendencies (aggression and cannibalism).
4) To ensure species-specific mating (and test the partner's quality).

The importance of these functions is generally believed to be great enough to overcome selection against prolonged, often conspicuous, displays (Platnick, 1971).

Platnick (1971) identified three phylogenetic levels of courtship behaviour (with level I being the most ancestral) based on the prime releaser of male display - the factor or factors required for the male to initiate courtship.

1) Level I courtship is initiated through direct contact with the female and is characteristic of many mygalomorphs and haplogynes, most Clubionidae and Thomosidae, and some Lycosidae.
2) Level II courtship involves chemotactic perception of silk and distance chemoreception of pheromones, and is characteristic of most web-building families, including the Agelenidae and hence *Tegenaria* species.
3) Level III courtship requires the male to sight the female and is characteristic of the visually acute Oxyopidae, Salticidae and the closely related Lyssomanidae.

CHEMICAL COMMUNICATION:

Male spiders actively seek females, and therefore the first problem a searching male faces is in locating the female web. There has been little work carried out on this initial step, partly because of an emphasis in studies which aim to identify the releaser of male courtship (Krafft, 1982) (as illustrated by the phylogenetic groupings outlined above). Chemical communication over short distances (up to 1 m) via female pheromones has been shown to affect locomotion and orientation in *Schizocosa* species (Lycosidae), is well known in orb-weaving spiders, and acts as a secondary courtship releaser in some salticids (Tietjen and Rovner, 1982). Further, Miyashita and Hayashi (1996) have shown that males of *Nephila clavata* are attracted to a volatile chemical produced by freshly moulted adult females, although they suggest that this may not be a compound 'designed' for sexual communication but merely a component of the moulting fluid.

Such distance effects have not been shown for an agelenid, however the sexual pheromone laid down in the silk of an adult and, possibly, sub-adult female's web may well play a role in attracting the male to the web, as well as eliciting courtship behaviour when on the web. Male *Tegenaria domestica* have been shown actively to follow draglines laid-down by adult females, but not to follow those of males (Krafft, 1982). Similar behaviour has also been explored in the closely related and equally sedentary agelenid *Ceolotes terrestris* (Krafft, 1982). Such 'trail- following' behaviour has been extensively studied in wandering lycosids (Krafft, 1982; Tietjen and Rovner, 1982). In the experiments reported in this chapter the males were denied the opportunity to search for the web of the female, but instead were placed directly on to it. The silk-born sex pheromones of the female may be sufficient to release male courtship behaviour. Females may also carry cuticular chemical cues, as may males. It has also been suggested that males may lay a web pheromone analogous to that of the female. *Lactrodectus hesperus* females exhibit

characteristic courtship movements when placed upon the web of a male (Krafft, 1982), and male lycosids have been observed to secrete greater amounts of silk than usual during chemoexploration and trail following (Tietjen and Rovner, 1982). The male pheromone could act to reduce female aggression in a similar manner to the flight-arresting pheromone produced by male Lepidoptera (Tietjen and Rovner, 1982). Such a pheromone could also be important in male-male interactions (Tietjen and Rovner, 1982). Males from a wide variety of genera, including *Tegenaria*, possess special gnathocoxal glands as a secondary sexual characteristic, which may well play a role in chemical communication. The epigastric and clypeal glands, characteristic of males in some other families and presently of unknown function may also act in this way (Legendre and Lopez, 1974; Krafft, 1982).

The importance of chemical communication in spider courtship suggests that it should be a good candidate for a premating isolation mechanism. Experiments on the cabbage looper moth *Trichoplusia ni* have shown that male behaviour can succesfully evolve to track mutant forms of the female attractant pheromone (Liu and Haynes, 1994); however, wild populations of the same species show probably identical sex pheromones to the alfalfa looper moth *Autographa californica* (Tietjen and Rovner, 1982). In a similar fashion, many male spiders only respond to the sex pheromone of conspecific females whereas others will frequently attempt to court females of closely related species, despite the obvious risk of predation by the female (Tietjen and Rovner, 1982). For example, male black widow spiders of the species *Lactrodectus mactans* and *L. hesperus* court either type of female (Tietjen and Rovner, 1982), male *Araneus pallidus* will court females of *A. diadematus* (Foelix, 1996), and males of *Schizocosa ocreata* and *S. rovneri* will court females of either species (Stratton and Uetz, 1981; 1983; 1986; 1987; Uetz and Stratton, 1982). Males of *S. ocreata* readily court females of a third species, *S. stridulans*, but males of *S. stridulans* show a lowered courtship response to *S.*

ocreata (Stratton, 1997) Interestingly, females of all three species reject heterospecific males, but males court both conscious and anaesthetised females, suggesting a chemosensory trigger but lack of specificity. Stratton and Uetz (1981) demonstrated that silk pheromone was *the* most important stimulus in eliciting male courtship response in *S. ocreata* and *S. rovneri*.

ACOUSTIC AND TACTILE COMMUNICATION:

The web of a spider constitutes an essential route of information transmission to the spider's perceptual systems (Witt, 1975). The male spider must locate the female on the web and approach, entering her 'privacy sphere' without triggering aggression or flight (Krafft, 1982). He does this by signalling to the female. In Lycosidae and Salticidae these signals may take the form of visually complex manoeuvres, often in association with vibrational cues through the substrate. In web spiders, communication is almost entirely acoustic[1], through the web. The male typically transmits information by drumming on the web with his palps, pulling and shaking the threads with his legs, and by vibrating the abdomen. The female responds by replying to his signals, or even by immobility - which can be regarded as a suppression of aggression. Production of sound has been described in 26 families of spiders (Uetz and Stratton, 1982) and three categories of sound production have been identified: stridulation, percussion, and vibration of structures. Stridulation occurs when a scraper is rubbed across a file. These may be any two rigid surfaces, for instance: abdomen rubbed against prosoma, one appendage rubbed against another or against the abdomen, or two opposing joint surfaces rubbed together. Stridulation in the Agelenidae has been described as occurring through the scraping of the abdomen and the pedicel (Uetz and Stratton, 1982). Percussion is generated by shocks to the substrate. In web building spiders this is usually the web, and may involve plucking of the web, tapping or drumming of the

[1] The term 'acoustic' is used to imply the transmission of signal through the vibration of molecules, and can therefore apply to sounds (heard) or vibrations through the subsrate (felt)

palps or legs and/or the abdomen (Uetz and Stratton, 1982). Vibration of structures (for example webs) may be caused by rapid oscillations of the appendages. For instance, *Heteropoda venatoria* may produce audible sounds on hard substrates in this manner (Uetz and Stratton, 1982).

On contacting the female there is often much tactile interplay with the legs or palps. The exchange of tactile signals, which almost certainly involve chemosensory and mechanical aspects, are important to prepare the female for copulation and may serve in mate quality assessment and species recognition (Krafft, 1982; Rovner, 1982). In *Lycosa helluo*, for example, tactile exchanges are essential for mating (Nappi, 1965).

6.1.3 Copulation

Mating position in the *Tegenaria atrica* group is Type 3 (of the categories defined by Foelix (1996)), with the male approaching the female from the front and passing to one side to insert a palp, and then repeating the procedure on the other side. Mating behaviour is described further in Section 6.3.1, and the complex structure of the copulatory organs, and their evolution and possible roles in mechanical isolation are discussed in Chapter 3 (section 3.1.1) and 7 (section 7.4).

6.1.4 Courtship and Reproductive Isolation

The often conspicuous and stereotypical nature of courtship in animals makes this an unusually tractable area for behavioural study. It is obvious that the courtship behaviours of animals are very diverse and frequently highly idiosyncratic between species. One only needs to consider the courtship and territorial songs of orthopterans, anurans or birds to realise this. However,

many of these diverse traits will have evolved under the various modalities of sexual or natural selection and play no direct role in mate recognition.

Courtship behaviour probably plays a major role in species recognition and mate choice in spiders. Most studies of spider courtship have focused on the male (partly because females often show little overt behaviour (Jackson, 1982)), and descriptions of the visually elaborate displays of Lycosidae and Salticidae. All studies of lycosids in which closely related species have been compared (Den Hollander *et al.*, 1973; Francescoli and Costa, 1992; Orta-Ocaña *et al.* 1996; Stratton, 1997; Stratton and Uetz, 1981; 1983; 1986; 1987; Uetz and Stratton, 1982; Vlijm and Dijkstra, 1966) have revealed major differences in the sequence and type of display elements in courtship. Major differences in the frequency and structure of the acoustic signals of males have also been noted in *Schizocosa* species (Stratton, 1997; Stratton and Uetz, 1981; 1983; 1986; 1987; Uetz and Stratton, 1982). Many of the species in the above studies were first delimited from their courtship behaviour.

Courtship comparisons have also been made for sedentary species. Significant differences in courtship behaviour have been observed in species of *Dictyna* and *Mallos* (Dictynidae) (Jackson, 1979). *Amaurobius ferox*, *A. similis*, and *A. fenestralis* (Amaurobiidae) (Krafft, 1978; 1982; Krafft *et al.* 1978; Leborgne, 1984; 1989; Leborgne and Krafft, 1979) have also been studied, with particular reference to acoustic signalling. By way of example, males of the sympatric species *A. ferox* and *A. fenestralis* emit abdominal vibrations (stridulation) at mean frequencies of 45 Hz and 150 Hz, respectively. The agelenids, *Coelotes terrestris*, *Tegenaria parietina*, *T. atrica*, *T. domestica* and *T. pagana*, have been studied similarly (Boulanger *et al.*, 1986; Krafft, 1978; 1982; Krafft *et al.*, 1978; Leborgne, 1986; 1989; Leborgne and Krafft, 1979; Leborgne *et al.*, 1980; Mielle, 1978). The closely related and sympatric species, *Tegenaria*

domestica and *T. pagana*, showed essentially no male vibrational motifs in common and males rarely courted heterospecific females (Boulanger *et al.*, 1986; Leborgne *et al.*, 1980). The courtship behaviour of *T. atrica* has been described (Krafft *et al.*, 1978; Leborgne, 1989; Mielle, 1978) and this species has been shown to have a mean abdominal vibration frequency of 30 Hz (compared with the non-immediate relative *T. parietina* at 15 Hz).

Inheritance of courtship behaviour has been studied in the lycosid species, *Schizocosa ocreata* and *S. rovneri* (Stratton, 1997; Stratton and Uetz, 1981; 1983; 1987; Uetz and Stratton, 1982). These species are behaviourally isolated (ethospecies). Males will court heterospecific females but are always rejected. No post-copulatory barriers to hybridization were observed when anaesthetised females were 'forced' to copulate with conspecific males, however the F_1 hybrid males were rejected by females of both species and F_1 hybrid females tended to reject males of all classes (Stratton, 1997; Stratton and Uetz, 1981; 1983; 1987; Uetz and Stratton, 1982). Similarly, anaesthetised females were employed to study developmental patterns in *Lycosa carbonelli* and *L. thorelli* (Francescoli and Costa, 1992). The species in the two studies mentioned have not been demonstrated to hybridize in the wild. The works of Vlijm and Dijkstra (1966) and Den Hollander *et al.* (1973) on *Pardosa* species, and of Reiskind and Cushing (1996) on *Lycosa ammophila* and *L. ericeticola*, represent the only studies of reproductive and courtship behaviour in spiders that may naturally hybridize. Indeed, the survey of Reiskind and Cushing (1996), in Florida, represents the only examination of a spider hybrid zone to date. Further, no study has examined courtship behaviour in two species from areas in which they are both allopatric and sympatric.

Uetz and Stratton (1982) note that if a trait (behavioural, morphological or physiological) is under selection to facilitate species recognition and reproductive isolation between two *hybridizing* taxa then it should not be identical in parapatric or sympatric closely related species. Andersson (1994) goes further:

1) The trait should affect premating isolation, with displacement reducing the likelihood of mismatings.
2) Displacement in the area of overlap should not result from a geographical trend also present in allopatry.
3) The divergence should have arisen from interactions between the two species and not from unique aspects of the environment in the area of overlap.

Two additional points should be made:
1) It is difficult to predict *a priori* which traits, for example elements of courtship display, are likely to undergo modification or exaggeration in areas of overlap.
2) Female mate choice must exhibit a correlated response to any change in male signal. This implies that the variation in male traits in the area of overlap should become narrowed, or at least some traits exaggerated relative to others, because they are under selection. Accordingly, the female response to stimuli should also become more narrow and follow the male's signals. If the female response remains broad, and quirks of the female sensory system make them generally prefer extreme males, this can result in females actually preferring males of another species (Andersson, 1994).

6.1.5 Aims

The work reported here describes the courtship and copulatory behaviour of *Tegenaria gigantea* and *T. saeva* for the first time, with the aim of attempting to quantify and identify any differences between these two species. Further, if there are any differences between the two species, are these greater in individuals from parapatric regions than from allopatric areas? Furthermore, are any differences in courtship behaviour correlated with a greater level of female discrimination against heterospecific males in parapatry?

In addition to elucidating courtship differences, the mating trials performed here also intended to verify, in the laboratory, the possibility of successful hybridization between the two species both in mechanical terms and in the production of viable (and fertile) offspring.

6.2 Materials and Methods

6.2.1 Sample Collection, Husbandry, and Experimental Procedure

Specimens of *T. saeva* and *T. gigantea* were obtained through two field collections (coll. P. J. P. Croucher) during 1996: one in late June and one in late August and early September. These sampling periods predated the natural breeding season (approximately mid September to late November) and hence ensured the collection of a high proportion of penultimate and antepenultimate-instar animals; female virginity was a necessary prerequisite for these experiments. Samples were taken from four geographic areas, corresponding to parapatric and allopatric distributions of the two species, using the detailed distribution knowledge collected from the survey work performed in 1994 and 1995 (see Chapter 2). Parapatric *T. saeva* and *T. gigantea* were collected from the vicinity of the contact zone in Dorset (zones 3 and 4, corresponding to sites within the 40 km $\times$ 40 km area covered by Ordnance Survey map sheet 195), deeply allopatric *T. saeva* were collected from Devon and Cornwall (zone 1) more than 100 km west of the contact zone, and deeply allopatric *T. gigantea* were collected from Kent, and East Sussex (zone 6) more than 100 km east of the contact zone.

Spiders were collected by hand as described in Chapter 2, and housed in 275 ml clear plastic containers. A diet of blowflies, *Lucilia caesar* (L.) and *Calliphora vomitoria* (L.), was provided at approximately four day intervals, and water was available *ad libitum*. Animals were maintained at 20-22°C, with a 12 hour : 12 hour day : night light cycle. On moulting to maturity, individuals were restrained in a 'Spi-pot' (Roberts, 1995), and placed under a dissecting microscope, in an attempt to identify the species. (A 'Spi-pot' is a holding device consisting of two clear plastic drinking cups, one inserted into the other. The inserted cup has a wedge of expanded polystyrene glued to one end and the outer-cup has the base removed and replaced with cling-film.) Live males

were generally easy to identify using this approach but live females were more difficult to identify as a result of the hairs covering the epigyne. Mature females were then transferred to perspex mating arenas measuring 23 cm (l) × 23 cm (w) × 9.5 cm (h). The arenas, which had been thoroughly cleaned with water and ethanol to remove any silk or pheromones from previous occupants, were provided with ventilation holes and covered with cling-film to prevent escape. The provision of food and water remained as above. Females were allowed to establish webs for not less than 3 days prior to mating.

Mating trials took place in a quiet, controlled temperature room under the conditions specified above. The mating arena was placed on white paper above soft rubber-foam (to limit extraneous vibrations) and the whole assembly placed on the base of a photographic copy stand. The arena was illuminated with two diametrically opposed incandescent bulbs (40 W), at approximately 50 cm distance, and courtship episodes videotaped using a Samsung 8 mm VHS Camcorder. The recordings were made for reference and for the verification of behaviours which were recorded against a timer by hand. The behavioural components of courtship and mating behaviour had been previously characterized from trial pairings performed in the previous year. The various behaviours are described under Results (Section 6.3).

Females and males were paired in eight combinations and an attempt was made to obtain at least ten full courtship sequences ending in attempted copulation for each of these eight combinations (although as a result of the difficulty in identifying females from the parapatric area, and the difficulty in synchronizing the maturity of the appropriate males and females, this was not always possible). In total 100 pairings were performed and the combinations and numbers of each are given below:

ALLOPATRIC:

Allopatric *T. saeva* female × Allopatric *T. saeva* male (Allopatric TSf × TSm)	12
Allopatric *T. gigantea* female × Allopatric *T. gigantea* male (Allopatric TGf × TGm)	13
Allopatric *T. gigantea* female × Allopatric *T. saeva* male (Allopatric TGf × TSm)	15
Allopatric *T. saeva* female × Allopatric *T. gigantea* male (Allopatric TSf × TGm)	13

PARAPATRIC:

Parapatric *T. saeva* female × Parapatric *T. saeva* male (Parapatric TSf × TSm)	12
Parapatric *T. gigantea* female × Parapatric *T. gigantea* male (Parapatric TGf × TGm)	8
Parapatric *T. gigantea* female × Parapatric *T. saeva* male (Parapatric TSf × TSm)	10
Parapatric *T. saeva* female × Parapatric *T. gigantea* male (Parapatric TSf × TSm)	17

An adult male was gently introduced onto the web of the female and all behaviours recorded for approximately 45-60 minutes. After this period, the male and female were left together for between 24 and 48 hours, to ensure sufficient opportunity for insemination to occur (in a few pairings where the female behaved in an overtly aggressive manner the male was removed and the experiment terminated). Whenever possible males were used only once, and

when this was not possible the male was allowed at least one week before being used in another experiment. After the insemination period, the male was removed and the female transferred back to a smaller container and maintained until at least three egg-sacs had been produced. Finally, males and females were killed by freezing at -80°C, transferred to 95% ethanol and the identification verified.

Each individual egg-sac was removed from a female's enclosure shortly after it had been produced and maintained according to Gunnarsson and Andersson (1992). The egg-sac was placed on a paper support inside a 40 ml plastic tube which was stoppered with absorbent cotton wool. The stopper was moistened with a few drops of water each week to maintain humidity. On hatching, about one month later, the spiderlings were counted whilst being pooted into a 275 ml clear plastic container, with moist tissue for water. Having transferred all the spiderlings, the egg-sacs were then torn open and any undeveloped eggs counted. Spiderlings were kept together until they started to eat each other and then 3-5 from each pairing were separated and reared to maturity. Spiderling husbandry was as previously described for adults except that small spiders were given a diet of *Drosophila* spp. supplemented with diamond-backed moths *Plutella xylostella* (L.) until large enough to take blowflies.

6.2.2 Transition Matrix Analyses of Pre-copulatory Courtship

The courtship behaviour of male and female *T. saeva* and *T. gigantea* was broken down into 18 specific and stereotyped behaviours or elements (see Section 6.3 Results). Although many subtleties in these elements will have been overlooked, and some closely intertwined behaviours have been pooled, these elements were easily identifiable (stereotyped) and as such are candidates for transmitting information between the sexes. It is necessary to find a quantitative approach to distinguish real "signals" and responses from accidental sequences of events. Transition analysis considers what would happen if none of the events was a signal and applies a goodness-of-fit test to a

accidental sequences of events. Transition analysis considers what would happen if none of the events was a signal and applies a goodness-of-fit test to a two-way transition matrix. In such a case, for example, any event performed by one spider would have the same chance of being followed by each of the events available in the repertoire of the other spider (Forster, 1982). However, transition analysis not only allows the examination of inter-individual sequences but also intra-individual sequences, for instance where one male behaviour is followed immediately by another male behaviour. Transition analysis has been employed in spiders to characterize the courtship behaviour of the salticid *Trite auricoma* (Forster, 1982), to assess the role of courtship as a species barrier between the lycosids *Pardosa pullata*, *P. prativaga* and *P. sphagnicola* (Den Hollander *et al.*, 1973), and to assess differences in the courtship of the two male morphs of the salticid *Maevia inclemens* (Clark, 1994). It has also been developed, especially in the application of information theory (which is not appropriate here), in evaluating aggressive communication in mantis shrimps (Dingle, 1969), and in the extensive study by Baylis (1976) comparing courtship dynamics in two species of cichlid fishes. The approach here closely follows and extends that of Clark (1994), which in turn was largely adapted from Dingle (1969) and Baylis (1976).

The preceding and following events from the recorded behavioural sequence of each pairing, for only those pairings ending in attempted copulation, were organized into a transition probability matrix. The individual matrices were then pooled to form one matrix for each type of pairing combination (for example: Allopatric TSf × TSm). Within a matrix, each cell or dyad, represents the total acts of behaviour j following behaviour i. The transition probability (P_{ij}) - the proportion of times that behaviour j follows behaviour i - could be calculated by dividing the observed frequency for each dyad by its corresponding row total (example from Table 6.4; page 278): cell A. MLSTA / B. MLCE; $Pij = 13/27 = 0.48$, etc.). Expected values for each cell were

calculated, as for a normal contingency table, by multiplying the column frequency by the corresponding row total (example from Table 6.4): column frequency for B. MLCE = 0.145 (N.B. the value given in the table is 0.15 because all values in the tables were rounded to 2 decimal places); row total for A. MLSTA = 27; expected for cell A. MLSTA / B. MLCE = $0.145 \times 27 = 3.92$, etc.).

Clearly many of the expected frequencies in the transition matrices have values less than 1. Therefore caution must be exercised in applying the χ^2 goodness-of-fit test to the data when it involves values generated from the individual cells in the matrix: the lowest expected values will contribute disproportionately to χ^2 values. However, with the application of appropriate caution this analysis can still be very illuminating. In order to be conservative, following Clark (1994), the Yates' continuity correction was applied in estimating the χ^2 values for *each cell* in a row by employing the first term in Clark's (1994) 'equation 1' (Equation 6.1, below), and hence the total row χ^2 value (with *df* = 17) could be generated. The sum of these row χ^2 values equals the χ^2 value for an entire matrix.

For each transition matrix, the significant dyads (cells) in each row had to be determined. Again, to be conservative, only rows with χ^2 values greater than 33.41 ($P < 0.01$) were considered. A statistical value cannot be attached to an individual cell with 0 *df*, therefore a modified χ^2 value with 1 *df* was generated for each cell in an analysed row, according to the formula devised by Clark (1994):

$$\chi^2 = \frac{((|OB - EX|) - 0.5)^2}{EX} + \frac{(((|GT - OB|) - (|GT - EX|)) - 0.5)^2}{GT - EX}$$

Equation 6.1

where *OB* is the cell Observed Value, *EX* is the cell Expected Value, and *GT* is the matrix Grand Total.

Further statistical manipulations will be discussed as they become relevant.

6.2.3 Duration of Courtship and Copulatory Parameters

In addition to the transition analyses of behavioural sequences outlined above, a number of temporal features of the courtship and copulatory behaviours (the behaviours and phases of courtship are described in Section 6.3) for the different pairing combinations were analysed: (1) *Total time spent in courtship*: onset of courtship (start of chemoexploratory behaviour) to onset of first attempted copulation; (2) *Latency to chemoexploratory behaviour*: time from when the male is first placed on the female web to when he first displays chemoexploratory behaviour; (3) *Duration of Phase I courtship*: time from the onset of chemoexploratory behaviour and associated elements until first occurrence of rapid drumming and vibration of the opithosoma (intense Phase II courtship); (4) *Duration of Phase II courtship*: time from onset of rapid drumming and vibration of opithosoma until first attempted copulation; (5) *Ratio of Phase II / Phase I*: the relative duration of Phase II and Phase I courtship; (6) *Duration of chemoexploratory bouts*: average length of a chemoexploratory bout broken by a stationary phase of 10 seconds or more; (7) *Insertion duration*: the average duration of an individual palpal insertion from moment of location with epigyne and insertion of conductor to removal of palp (corresponding to one haematodochal inflation). All the above were analysed for pairings ending in attempted copulation. Comparisons were made, where possible, between pairings ending in copulation and pairings not ending in copulation for (2), (3) and (6). In addition, an attempt was made to relate insertion duration to the morphology of the male palp and the female epigyne,

to examine whether relative female : male size (genitalia) was related to copulatory success.

6.3 Results

6.3.1 Courtship and Mating Behaviour.

The courtship of *T. saeva* and *T. gigantea* could be divided broadly into three phases: Phase I consisted of male exploratory behaviour and associated components. In general Phase I behaviour would switch to Phase II behaviour after the male had made physical contact with the female. Phase II behaviour was characterized by a change to a more intense phase of courtship clearly directed towards the female. Phase II courtship would end when attempted copulation first occurred, to be replaced by Phase III (copulatory and post-copulatory) courtship. These three phases, although easily definable, were not mutually exclusive. Phase II courtship could switch back to Phase I courtship, and the copulatory period of Phase III courtship largely consisted of Phase II elements. The behavioural elements used in this analysis will be discussed in the context of the generalization of their occurrence in these three phases.

PHASE I. MALE STATIONARY (A. MLSTA): male not locomoting or performing any other behaviour. A male was regarded as stationary if he remained motionless for 10 seconds or longer. MALE CHEMOEXPLORE (B. MLCE): male engaged in apparently non-directed exploratory behaviour. On contacting the female web the male would typically adopt a stance with his eight legs widely spaced and maintain this posture throughout Phase I courtship. The male would typically take a few short steps forward whilst gently moving the palpal tarsi against the female web in a pedalling motion. The palpal motion most likely primarily serves a chemosensory role (as well as perhaps signalling to the female - although web displacement seemed minimal). Chemosensory sensilla are known to be concentrated on the distal segments of the palps and legs of spiders (Foelix, 1996) and adult male lycosids have been demonstrated to have particularly high numbers of

chemosensory sensilla on the palpal tibia (relative to females and penultimate instar males), particularly the medial and dorsal surfaces (Tietjen and Rovner, 1982). These sensilla are implicated in female drag-line 'trail' following behaviour by males (Tietjen and Rovner, 1982). The forward locomotion was generally accompanied by the spinnerets being held widely spaced, laying a broad band of silk with gentle lateral opithosomal sweeps. After a few steps the male would pause and having depressed the spinnerets towards the web stratum, move the opithosoma sharply and suddenly upwards (abdominal kick or tap) - thus vibrating the web. These abdominal kicks did not appear to represent a single smooth action and may carry complex vibrational information (see Section 6.1.2). Bouts of locomotion were generally interrupted by pauses and abdominal kicks every 2-5 seconds. No distinction was made between locomotion towards or away from the female. MALE SHAKE/FLEX WEB (C. MLSF): the male, whilst not locomoting, would gently shake the web by rhythmic tensioning and relaxation of all eight legs. The male was also occasionally seen to pull the web more forcibly. MALE BITE-WEB (D. MLBW): the male takes the female web within his chelicerae and pulls it up and releases it with an often audible snap. The male generally pressed his body against the web, taking the web in his chelicerae, followed by raising himself in an anterior-posterior cat-stretch-like motion. MALE RUNS RANDOMLY (E. MLRND): the male interrupts a behaviour with a sudden, short, non-directed sprint. MALE MOUTHS PALPS/FEET (F. MLMP): the male passes the tibia of the palps or legs through the chelicerae in a preening type behaviour, presumably to clean these chemosensory regions. FRONT LEG TOUCH (G. FLT): this is also a Phase II and Phase III behaviour and simply represents the touching of one sex by the other with the front legs (legs I). Usually this would involve the male approaching and touching the female which would generally, especially in Phase II courtship, occur from behind with the male's front legs touching the females back legs (legs IV) or opithosoma. This would generally result in a female reaction such as the female

walking away, orienting to the male, or repelling the male. Orientation to the male might result in a further FLT. Only very occasionally would this represent the result of an approach by a proactive female.

PHASE II. MALE DRUM, VIBRATE ABDOMEN (H. MLDVA): this element was characteristic of Phase II and consisted of two interlinked behaviours. The male would pause and then gently begin to vibrate the opithosoma and drum gently on the web with the palps. Over a period of up to a few seconds the intensity of these actions would increase terminating in a powerful drum-roll which would reach a peak and then abruptly cease. This behaviour might be repeated several times before the male would begin to advance slowly whilst drumming gently and rapidly vibrating the opithosoma. The male would pause in his advance every few seconds to repeat the cycle. This behaviour was generally directed towards the female and repeated cycles would often result in the male contacting the female with his front legs. MALE WALK AWAY (I. MLWA): specifically during Phase II, whilst performing MLDVA, the male might turn and retreat from the female's location. MALE ORIENTATE TO FEMALE (J. MLOF): as for MLWA but the male swivels towards the female. FACE-TO-FACE (K. FTF): a poignant moment when the male and female are facing each other and close enough to touch palps. SIDE-TO-SIDE (L. STS): can only follow on from FTF, the male moves his prosoma to one side of the females opithosoma in a position such that copulation can occur, and such that four of his legs may 'embrace' her, the female may respond by rotating her opithosoma by about 30° to facilitate copulation. ATTEMPTED COPULATION (M. ACOP): following on from STS the male attempts to copulate with the female (see below). The behaviours FLT, FTF, STS and ACOP, although Phase II behaviours, involve a direct interaction between the male and the female and will generally be classed as 'male/female' behaviours.

FEMALE SPECIFIC BEHAVIOURS. FEMALE WALK (O. FEW): female walks. This may be in response to contact by the male (walking/turning away) or an 'independent' locomotion (not directed towards the male). FEMALE ORIENTATE TO MALE (P. FEOM): female swivels towards the male, generally by taking one or two steps forward and then rotating. FEMALE APPROACH MALE (Q. FEAM): female locomotes towards the male's location. FEMALE SIGNAL (R. FESIG): the female pulls the web gently with her legs and then generally settles with the body lowered to the web and the legs pulled up close to the body. She may also rub her palps against the web in a motion reminiscent of male drumming. REPELLED (N. REP): male is repelled by an apparently aggressive action of the female. Typically the female would adopt an aggressive warning posture and then run rapidly at the male causing him to flee quickly (see Figure 6.1). REP could occur in any phase of courtship.

PHASE III. The transition analysis only includes data up to the first attempted copulation and therefore does not attempt to characterize Phase III behaviour in any detail. Copulation behaviour generally follows the same pattern as the initial copulation attempt. A typical scenario would be that the male repeatedly approaches the female (MLDVA) and touches her back legs (FLT). At this point the female might initially respond by walking away from the male (FEW). However, eventually she responds by taking a step forward and turning towards the male (FEOM) resulting in a further FLT. The two spiders may then edge forward (FTF) and the male moves to the left or right of the female (STS). The female may walk away at any point. If the female is to the right of the male (viewed from above), he then attempts to copulate with his right palp. The female rotates her opithosoma through about 30° to facilitate copulation and the male shakes the palp in an anterior-posterior arc until he engages the palpal tibial apophyses with the posterior epigynal apophysis. The sclerites of the palp then undergo a complex series of expansions and rotations such that

Figure 6.1. A female *Tegenaria saeva* displaying an aggressive warning posture to an approaching male *T. gigantea*. The female is on the left, first pair of legs raised and spread, fangs bared. The male is approaching from the bottom-right. (The object to the top-right is part of a vibration recording device).

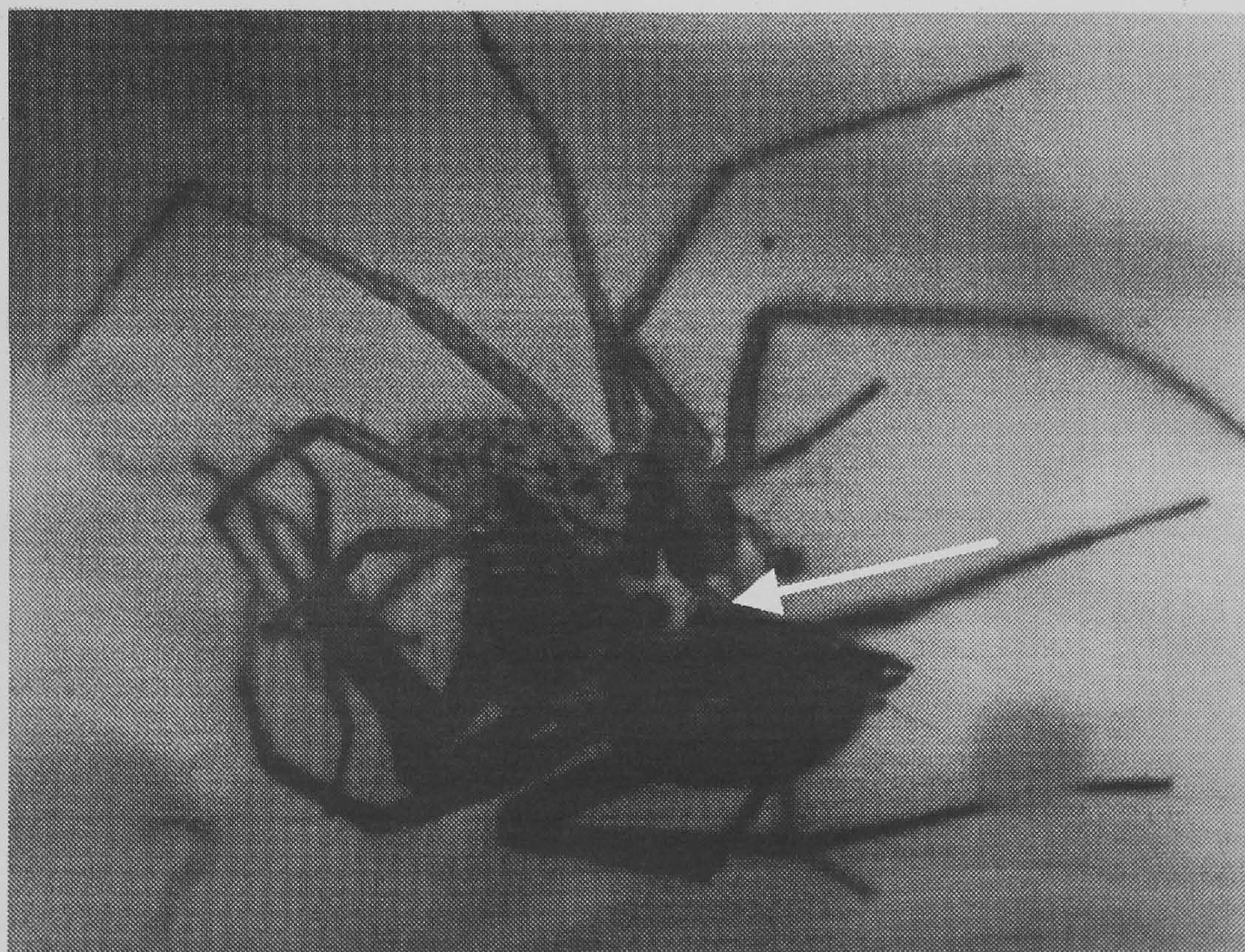


Figure 6.2. A pair of *T. saeva* '*in copula*'. Note the haematodochal inflation (arrowed).

the embolus is positioned in front of the epigynal opening (receptacle opening or copulatory duct). The final inflation of the haematodochal membranes pushes the conductor towards the opening, and the embolus into the copulatory duct (Foelix, 1996) (see Figure 6.2). As the haematodochal inflation subsides the palp is removed, shaken and re-engaged, in other words the palp is removed from the epigyne between each inflation. Copulation may be repeated many times with the same palp and is broken when the male backs up and moves to the other side of the female to use the other palp, or when the female walks away. When the sexes separate the male returns to MLDVA and the process may be repeated many times (often over many hours - personal observation). Eventually copulation ceases and the male is seen to perform abdomen kicks whenever he or the female moves.

6.3.2 Conspecific and heterospecific crosses

Overall there were 100 female/male pairings of which 84 (84%) resulted in attempted copulation (Table 6.1). Trial date (Julian day - data not shown) had no significant effect on the likelihood of pairings ending or not ending in attempted copulation (Mann-Whitney *U*-test, 2-tailed, $P = 0.08$, n.s.). In all but two of the pairings (one for the allopatric TSf × TSm combination and one for the allopatric TSf × TGm combination) males responded positively with courtship behaviour (Table 6.1). Clearly males of both *T. saeva* and *T. gigantea* respond to being placed on an adult female's web with courtship, regardless of the identity of the occupier. Whether attempted copulation would occur depended largely therefore on the female's response to the male. In pairings where females responded favourably, attempted copulation occurred. For the allopatric pairing combinations: 10 out of 12 (83%) TSf × TSm pairings; 12 out of 13 (92%) TGf × TGm pairings; 13 out of 15 (87%) TGf ×TSm pairings; and 11 out of 13 (85%) TSf × TGm pairings resulted in attempted copulation

(Table 6.1). There were no significant differences between these copulation frequencies for any pairwise comparison (Fisher's Exact Test; 2-tailed; $P > 0.5$, all comparisons) (expected values were too low for χ^2 or G-tests). For the parapatric pairing combinations: 12 out of 12 (100%) TSf × TSm pairings; 8 out of 8 (100%) TGf × TGm pairings; 8 out of 10 (80%) TGf × TSm pairings; and 10 out of 17 (59%) TSf × TGm pairings showed attempted copulation (Table 6.1). Clearly there was no difference in copulation frequency between the conspecific pairings, nor did *T. gigantea* females show a significantly less favourable response to *T. saeva* males than to their males of their own species (comparison of TGf × TGm with TGf × TSm; Fisher's Exact Test; 1-tailed; $P = 0.294$; n.s.). However, *T. saeva* females did respond less favourably to *T. gigantea* males than to males of their own species (comparison (all pairings - allopatric and parapatric) of TSf × TSm with TSf × TGm; Fisher's Exact Test; 1-tailed; $P = 0.013$). That parapatric *T. saeva* females copulate less frequently with heterospecifics than do *T. gigantea* females lacked support (comparison of TSf × TGm with TGf × TSm; Fisher's Exact Test; 1-tailed; $P = 0.244$, n.s.). Despite the fact that *T. saeva* females apparently copulated less frequently with *T. gigantea* males than with males of their own species, there was no evidence to support the notion that *T. saeva* females discriminated against *T. gigantea* males in parapatry more than in allopatry (comparison of parapatric TSf × TGm and allopatric TSf × TGm; Fisher's Exact Test; 1-tailed; $P = 0.130$; n.s.). The probability value is more suggestive of such a relationship, however, than the equivalent comparison for *T. gigantea* females (comparison of parapatric TGf × TSm with allopatric TGf × TSm; Fisher's Exact Test; 1-tailed; $P = 0.532$; n.s.). These interpretations should be treated with caution given only moderate sample sizes and repeated testing.

Overtly aggressive female behaviour towards males (Table 6.1) was only observed in TSf × TGm pairings, occurring in 2 out of 13 (15%) allopatric TSf × TGm pairings and 5 out of 17 (29%) of parapatric TSf × TGm pairings. The greater frequency of female aggression in the parapatric pairings was not significant (Fisher's Exact Test; 1-tailed; P = 0.326; n.s.). In some cases the males were removed when the female appeared to be a physical threat (see Table 6.1).

Cannibalism (Table 6.1) of the male by the female was a rare event, occurring in only 8 out of 100 (8%) pairings. Six out of 8 of these events occurred in heterospecific pairings but this was not significant (Fisher's Exact Test; 1-tailed (excluding cases where the male was removed); P = 0.210; n.s.).

Table 6.1. Male and female responses in each pairing combination.

Pairings	No. of pairs	Male Response		Female Response			Male Eaten*
Allopatric		+	-	+	-	--	
TSf × TSm	12	11	1	10	2	0	0
TGf × TGm	13	13	0	12	1	0	0
TGf × TSm	15	15	0	13a	2	0	2
TSf × TGm	13	12	1	11a	0	2(1)	3
Parapatric							
TSf × TSm	12	12	0	12	0	0	0
TGf × TGm	8	8	0	8	0	0	2
TGf × TSm	10	10	0	8	2	0	0
TSf × TGm	17	17	0	10	2	5 (2)	1

A plus sign for males indicates that courtship behaviour was shown. A plus sign for females indicates that females were receptive and that the pairing resulted in attempted copulation; a minus sign indicates that the female was unresponsive or evaded the male's approaches; a double minus sign indicates an overtly aggressive response to the male by the female. Figures in brackets indicate males that were removed to prevent the female killing them. In each of the two figures marked 'a', two of the pairings resulted in copulation just after the detailed observation period and were not included in the transition analyses, but were included in the analyses described here. *The 'Male Eaten' column indicates males that were eaten by females within the 24 hours *following* the trial.

Table 6.2. Reproductive success in each pairing combination.

Pairings	No. of pairs	No. of cops.	No. producing egg-sacs	No. producing offspring	Mean No. offspring ± S.D.	Mean No. eggs ± S.D.	Ave. viable
Allopatric							
TSf x TSm	12	10	10a	10	38.53 ± 19.64 (*n* = 10)	51.17 ± 7.13 (*n* = 10)	0.75
TGf x TGm	13	12	12	12	40.94 ± 25.46 (*n* = 12)	59.78 ± 10.00 (*n* = 12)	0.68
TGf x TSm	15	13	14a	0	none	62.62 ± 6.48 (*n* = 14)	-
TSf x TGm	13	11	12a	0	none	57.86 ± 10.22 (*n* = 12)	-
Parapatric							
TSf x TSm	12	12	11	10	42.06 ± 22.62 (*n* = 10)	59.91 ± 6.41 (*n* = 10)	0.70
TGf x TGm	8	8	8	8	33.67 ± 13.27 (*n* = 8)	63.62 ± 8.51 (*n* = 8)	0.53
TGf x TSm	10	8	10	1	32.67 (*n* = 1)	56.80 ± 6.65b (*n* = 10)	0.59
TSf x TGm	17	10	14a	0	None	60.71 ± 11.43 (*n* = 14)	-

Columns show the number of pairings attempted; the number of pairings observed to attempt copulation; the number of females producing at least one egg-sac; the number of females producing offspring; the mean number of offspring per egg-sac for each pairing combination taken as the mean of the average number of offspring from up to three egg-sacs per female (*n* = the number of females for which this figure evaluated); the mean number of eggs per egg-sac averaged across up to three egg-sacs; and the average viability - the mean proportion of offspring per egg for females producing at least one offspring (sample size as for 'Mean No. of Offspring'. a = 1 female died before producing an egg-sac. b = mean numberof eggs for the parapatric Tgf x Tsm pairings wac calculated from the egg-sacs of all 10 females producing eggs; although only one female produced offspring.

Values for some parameters of reproductive success in each of the pairing combinations are summarized in Table 6.2. Most, 40 out of 45 (89%), conspecific pairings resulted in the production of offspring, whereas only 1 out of 55 (0.02%) heterospecific pairings resulted in offspring. The single successful production of hybrids came from a parapatric TGf × TSm pairing. There was no significant difference in the mean number of eggs produced per egg-sac by females for each of the pairing combinations (see Table 6.3). The

average viability (the mean proportion of offspring per egg) per pairing comparison (Table 6.2) was subjected to a *G*-test of heterogeneity using the simultaneous test procedure of Sokal and Rohlf (1995; p. 722) (and as employed in Chapter 4). The allopatric TGf × TSm, allopatric TSf × TGm, and parapatric TSf × TGm combinations were excluded as they produced zero offspring. The test therefore included the average viability for the single successful parapatric TGf × TSm and all conspecific pairings (Table 6.2), there was no significant deviation from homogeneity (G_H = 8.33; $P > 0.05$ n.s.; tabulated χ^2, df = 4). This suggests (although there was only one hybrid progeny) that, although TGf × TSm pairings will rarely produce viable eggs, when they do the average viability may not be different from that in conspecific pairings.

Table 6.3. ANOVA: Mean number of eggs per egg-sac for each pairing combination.

Analysis of variance					
Source of variation	*df*	*SS*	*MS*	F_S	*P*
Among pairing combinations	7	1086.4	155.2	2.06	0.06 n.s.
Within pairing combinations	83	6262.8	75.5		
Total	90	7349.2			

Analysis of variance on the mean number of eggs per egg-sac for each female in each of the pairing combinations (overall mean given in Table 6.2). The ANOVA was not quite significant at the $P < 0.05$ level. Although it was close to significance, the means in Table 6.2 lack any obvious pattern between allopatric/paraptric or conspecifc/heterospecific pairings.

All undeveloped eggs from all pairings were, almost without fail, dark yellow (yolky), which suggests that they were unfertilized.

The F_1 progeny of the single successful TGf × TSm pairing contained both males and females. (Unfortunately, relatively few offspring were reared because it was initially believed that there were four successful interspecific crosses - subsequent examination under the microscope revealed that three of the females were misidentified. Consequently, many young had either been preserved or had been allowed to be eaten by their siblings.) Four males and

two females were successfully raised to sexual maturity. A further seven individuals died during moults at around the third to fourth instar. This gave an approximate laboratory mortality rate of 54% compared with an estimated mortality of 4% (5 out of 139 individuals with 3 to 5 raised from each female) in non-hybrid progeny reared under the same conditions. With so few animals raised, and only one sample, interpretation must be cautious, however the discrepancy between these two estimates is one order of magnitude and may suggest a considerable hybrid disadvantage.

A further intriguing observation was that the sex ratio of the conspecific offspring was heavily male-biased. Overall, both allopatric and parapatric TSf × TSm pairings resulted in 67 males and 11 females at adulthood and allopatric and parapatric TGf × TGm pairings resulted in 46 males and 10 females. These numbers were not significantly different (Fisher's Exact test: $P = 1.00$). Overall there were 113 males and 21 females giving a sex ratio of 5.4:1. This was highly deviant from the expected 1:1 ratio ($G = 69.4$; tabulated χ^2; $df = 1$; $P < 0.001$). The reason for such a male bias is unclear, but it may result from males cannibalizing females before the young were separated at around the second instar. Little can be deduced about the primary sex-ratio without sexing eggs.

6.3.3 Transition Matrix Analysis

Only those pairings that ended with copulation *during the observation period* (see Table 6.1) were used in the transition matrix analysis ($n = 10$ allopatric TSf × TSm; $n = 12$ allopatric TGf × TGm; $n = 11$ allopatric TGf × TSm; $n = 9$ allopatric TSf × TGm; $n = 12$ parapatric TSf × TSm; $n = 8$ parapatric TGf × TGm; $n = 8$ parapatric TGf × TSm; $n = 10$ parapatric TSf × TGm). The transition matrices for each pairing combination are given in Tables 6.4 through 6.11.

For all cross types, the courtship behavioural acts preceding were not independent of the acts following indicating non-randomness in the sequence of behaviours (allopatric TSf × TSm: χ^2 = 1187.53; allopatric TGf × TGm: χ^2 = 1680.85; allopatric TGf × TSm: χ^2 = 1577.00; allopatric TSf × TGm: χ^2 = 996.48; parapatric TSf × TSm: χ^2 = 1366.90; parapatric TGf × TGm: χ^2 = 658.04; parapatric TGf × TSm: χ^2 = 931.59; parapatric TSf × TGm: χ^2 = 1164.90; All $df = 272$, $P < 0.001$; Tables 6.4 through 6.11 respectively).

Significant dyads (cells) were extracted from each transition matrix by estimating the χ^2 value for each cell of the matrix according to Equation 6.1. This analysis was only performed for rows with a row χ^2 greater than 33.41 ($P < 0.01$). In this way a table of preceding acts that significantly facilitate (greater than expected), and inhibit (less than expected), a following act at the 0.01 level with 1 degree of freedom, could be constructed and compared for each pairing combination. These results are not readily interpretable in terms of differences between the pairing combinations, but do reveal important transitions in the behavioural sequences. Given the complexity of these tables and their limited comparative value they have been given as an appendix (Appendix A.3). Examination of these tables reveals them to be in broad agreement with the general description of courtship behaviour and its elements presented earlier. (These data must be treated with some caution, because although the tests employed have been very conservative, some dyads have very low expected values which have the effect of inflating significance. This was particularly true for the inhibitory effects: behaviours appeared to be inhibited by other behaviours purely because the apparently inhibited behaviours occurred infrequently.)

Table 6.4. Allopatric *T. saeva* female × *T. saeva* male (TSf × TSm) transition matrix (n = 10).

Allopatric TSf x TSm	Following acts																				
	Preceding acts	A. MLSTA	B. MLCE	C. MLSF	D. MLBW	E. MLRND	F. MLMP	G. FLT	H. MLDVA	I. MLWA	J. MOF	K. FTF	L. STS	M. ACOP	N. REP	O. FEW	P. FEOM	Q. FEAM	R. FESIG	Row Total	Row χ2
A.	MLSTA	0	13	5	0	1	0	0	6	0	0	0	0	0	0	2	0	0	0	27	34.65
	exp.	2.09	3.92	1.39	0.26	0.35	0.09	2.96	4.70	0.17	0.61	1.31	1.13	0.87	0.00	4.53	1.57	0.44	0.61		
B.	MLCE	8	0	3	2	1	1	14	3	0	0	0	0	0	0	10	1	4	3	50	42.85
	exp.	3.87	7.26	2.58	0.44	0.65	0.16	5.48	8.71	0.32	1.13	2.42	2.10	1.61	0.00	8.39	2.90	0.81	1.13		
C.	MLSF	6	4	0	0	0	0	0	5	0	0	0	0	0	0	1	0	0	0	16	25.14
	exp.	1.24	2.32	0.83	0.15	0.21	0.05	1.75	2.79	0.10	0.36	0.77	0.67	0.52	0.00	2.68	0.93	0.26	0.36		
D.	MLBW	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	3	67.69
	exp.	0.23	0.44	0.15	0.03	0.04	0.01	0.33	0.52	0.02	0.07	0.15	0.13	0.10	0.00	0.50	0.17	0.05	0.07		
E.	MLRND	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	79.07
	exp.	0.31	0.58	0.21	0.04	0.05	0.01	0.44	0.70	0.03	0.09	0.19	0.17	0.13	0.00	0.67	0.23	0.06	0.09		
F.	MLMP	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	108.53
	exp.	0.15	0.29	0.10	0.02	0.03	0.01	0.22	0.35	0.01	0.05	0.10	0.08	0.06	0.00	0.33	0.12	0.03	0.05		
G.	FLT	0	0	0	0	0	0	0	3	0	0	1	0	0	0	22	7	0	1	34	67.50
	exp.	2.63	4.94	1.75	0.33	0.44	0.11	3.73	5.92	0.22	0.77	1.65	1.43	1.10	0.00	5.70	1.97	0.55	0.77		
H.	MLDVA	3	5	3	1	0	0	15	0	2	3	5	0	0	0	8	7	0	2	54	35.00
	exp.	4.18	7.84	2.79	0.52	0.70	0.17	5.92	9.41	0.35	1.22	2.61	2.26	1.74	0.00	9.06	3.14	0.87	1.22		
I.	MLWA	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	103.64
	exp.	0.15	0.29	0.10	0.02	0.03	0.01	0.22	0.35	0.01	0.05	0.10	0.08	0.06	0.00	0.34	0.12	0.03	0.05		
J.	MOF	0	0	1	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0	7	28.34
	exp.	0.54	1.02	0.36	0.07	0.09	0.02	0.77	1.22	0.05	0.16	0.34	0.29	0.23	0.00	1.17	0.41	0.11	0.16		
K.	FTF	0	0	0	0	0	0	0	0	0	0	0	13	0	0	3	0	0	0	16	220.57
	exp.	1.24	2.32	0.83	0.15	0.21	0.05	1.75	2.79	0.10	0.36	0.77	0.67	0.52	0.00	2.68	0.93	0.26	0.36		
L.	STS	0	0	0	0	0	0	0	0	0	0	1	0	10	0	2	0	0	0	13	209.62
	exp.	1.01	1.89	0.67	0.13	0.17	0.04	1.43	2.26	0.08	0.29	0.63	0.55	0.42	0.00	2.18	0.75	0.21	0.29		
N.	REP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	exp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
O.	FEW	3	13	2	0	2	0	2	21	0	4	0	0	0	0	0	3	0	1	51	38.61
	exp.	3.95	7.40	2.63	0.49	0.66	0.16	5.59	8.88	0.33	1.15	2.47	2.14	1.65	0.00	8.55	2.96	0.82	1.15		
P.	FEOM	0	0	1	0	0	0	1	5	0	0	8	0	0	0	2	0	1	0	18	59.86
	exp.	1.39	2.61	0.93	0.17	0.23	0.06	1.97	3.14	0.12	0.41	0.87	0.75	0.58	0.00	3.02	1.05	0.29	0.41		
Q.	FEAM	0	3	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	5	39.72
	exp.	0.39	0.73	0.26	0.05	0.06	0.02	0.55	0.87	0.03	0.11	0.24	0.21	0.16	0.00	0.84	0.29	0.08	0.11		
R.	FESIG	0	2	1	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	8	26.73
	exp.	0.62	1.16	0.41	0.08	0.10	0.03	0.88	1.39	0.05	0.18	0.39	0.34	0.26	0.00	1.34	0.46	0.13	0.18		
Total:		24	45	16	3	4	1	34	54	2	7	15	13	10	0	52	18	5	7	310	1187.53
Frequency		0.08	0.15	0.05	0.01	0.01	0.00	0.12	0.17	0.01	0.02	0.05	0.04	0.03	0.00	0.17	0.06	0.02	0.02		

Behaviours are identified by their letter codes and acronyms (for example A = MLSTA). The top number in a row is the observed value and the bottom number the expected value. χ^2 values are given in the far right column.

Table 6.5. Allopatric *T. gigantea* female × *T. gigantea* male (TGf × TGm) transition matrix (n = 12).

Allopatric TGf x TGm		Following acts																			
	Preceding acts	A. MLSTA	B. MLCE	C. MLSF	D. MLBW	E. MLRND	F. MLMP	G. FLT	H. MLDVA	I. MLWA	J. MOF	K. FTF	L. STS	M. ACOP	N. REP	O. FEW	P. FEOM	Q. FEAM	R. FESIG	Row Total	Row χ^2
A.	MLSTA	0	12	6	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	20	90.63
	exp.	0.87	1.70	0.98	0.05	0.41	0.15	2.88	4.58	0.77	0.67	0.72	0.72	0.62	0.10	2.78	0.77	0.72	0.51		
B.	MLCE	11	0	4	0	0	1	16	1	0	0	0	0	0	0	2	1	5	1	42	79.41
	exp.	1.84	3.56	2.05	0.11	0.86	0.32	6.05	9.61	1.62	1.40	1.51	1.51	1.30	0.22	5.83	1.62	1.51	1.08		
C.	MLSF	3	8	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	1	19	36.74
	exp.	0.83	1.61	0.93	0.05	0.39	0.15	2.74	4.35	0.73	0.63	0.68	0.68	0.59	0.10	2.64	0.73	0.68	0.49		
D.	MLBW	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	250.98
	exp.	0.04	0.08	0.05	0.00	0.02	0.01	0.14	0.23	0.04	0.03	0.04	0.04	0.03	0.01	0.14	0.04	0.04	0.03		
E.	MLRND	1	3	2	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	8	30.35
	exp.	0.35	0.68	0.39	0.02	0.16	0.06	1.15	1.83	0.31	0.27	0.29	0.29	0.25	0.04	1.11	0.31	0.29	0.21		
F.	MLMP	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	3	74.33
	exp.	0.13	0.25	0.15	0.01	0.06	0.02	0.43	0.69	0.12	0.10	0.11	0.11	0.09	0.02	0.42	0.12	0.11	0.08		
G.	FLT	0	2	0	0	7	0	0	5	2	0	8	0	0	1	33	2	1	0	61	126.41
	exp.	2.67	5.17	2.98	0.16	1.25	0.47	8.78	13.96	2.35	2.04	2.20	2.20	1.88	0.31	8.47	2.35	2.20	1.57		
H.	MLDVA	0	0	1	1	0	2	36	0	10	10	3	1	0	0	9	6	5	6	90	106.73
	exp.	3.93	7.63	4.40	0.23	1.85	0.69	12.96	20.59	3.47	3.01	3.24	3.24	2.78	0.46	12.49	3.47	3.24	2.31		
I.	MLWA	1	1	0	0	0	0	0	12	0	1	0	0	0	0	0	0	0	0	15	30.90
	exp.	0.66	1.27	0.73	0.04	0.31	0.12	2.16	3.43	0.58	0.50	0.54	0.54	0.46	0.08	2.08	0.58	0.54	0.39		
J.	MOF	0	1	0	0	0	0	0	11	0	0	0	0	0	0	0	1	0	0	13	32.41
	exp.	0.57	1.10	0.63	0.03	0.27	0.10	1.87	2.97	0.50	0.43	0.47	0.47	0.40	0.07	1.80	0.50	0.47	0.33		
K.	FTF	0	0	0	0	0	0	0	1	0	0	0	13	0	0	0	0	0	0	14	299.38
	exp.	0.61	1.19	0.68	0.04	0.29	0.11	2.02	3.20	0.54	0.47	0.50	0.50	0.43	0.07	1.94	0.54	0.50	0.36		
L.	STS	0	0	0	0	0	0	0	0	0	0	0	0	12	0	2	0	0	0	14	297.81
	exp.	0.61	1.19	0.68	0.04	0.29	0.11	2.02	3.20	0.54	0.47	0.50	0.50	0.43	0.07	1.94	0.54	0.50	0.36		
N.	REP	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	117.99
	exp.	0.09	0.17	0.10	0.01	0.04	0.02	0.29	0.46	0.08	0.07	0.07	0.07	0.06	0.01	0.28	0.08	0.07	0.05		
O.	FEW	1	3	3	0	0	0	1	31	2	2	0	0	0	0	0	5	0	0	48	54.15
	exp.	2.10	4.07	2.34	0.12	0.99	0.37	6.91	10.98	1.85	1.60	1.73	1.73	1.48	0.25	6.66	1.85	1.73	1.23		
P.	FEOM	0	0	0	0	0	0	2	5	0	0	2	0	0	0	2	0	2	1	14	15.33
	exp.	0.61	1.19	0.68	0.04	0.29	0.11	2.02	3.20	0.54	0.47	0.50	0.50	0.43	0.07	1.94	0.54	0.50	0.36		
Q.	FEAM	0	0	2	0	0	0	1	6	0	0	1	0	0	1	3	0	0	1	15	12.14
	exp.	0.66	1.27	0.73	0.04	0.31	0.12	2.16	3.43	0.58	0.50	0.54	0.54	0.46	0.08	2.08	0.58	0.54	0.39		
R.	FESIG	0	1	1	0	0	0	0	7	1	0	0	0	0	0	0	0	0	0	10	25.15
	exp.	0.44	0.85	0.49	0.03	0.21	0.08	1.44	2.29	0.39	0.33	0.36	0.36	0.31	0.05	1.39	0.39	0.36	0.26		
Total:		17	33	19	1	8	3	56	89	15	13	14	14	12	2	54	15	14	10	389	1680.85
Frequency		0.04	0.08	0.05	0.00	0.02	0.01	0.14	0.23	0.04	0.03	0.04	0.04	0.03	0.01	0.14	0.04	0.04	0.03		

Behaviours are identified by their letter codes and acronyms
(for example A = MLSTA). The top number in a row is the observed value and the bottom number the expected value. χ^2 values are given in the far right column.

Table 6.6. Allopatric *T. gigantea* female × *T. saeva* male (TGf × TSm) transition matrix (n = 11).

Allopatric TGf x TSm		Following acts																			
	Preceding acts	A. MLSTA	B. MLCE	C. MLSF	D. MLBW	E. MLRND	F. MLMP	G. FLT	H. MLDVA	I. MLWA	J. MOF	K. FTF	L. STS	M. ACOP	N. REP	O. FEW	P. FEOM	Q. FEAM	R. FESIG	Row Total	Row χ^2
A.	MLSTA	0	11	6	0	1	0	0	9	0	0	0	0	0	0	1	0	0	0	28	37.70
	exp.	1.61	3.03	1.48	1.67	0.32	0.13	3.86	6.63	0.64	0.97	0.84	0.84	0.71	0.00	3.80	0.64	0.32	0.51		
B.	MLCE	9	0	5	19	1	0	17	0	0	0	0	0	0	0	2	0	0	3	56	115.48
	exp.	3.22	6.05	2.96	3.35	0.64	0.26	7.72	13.26	1.29	1.93	1.67	1.67	1.42	0.00	7.60	1.29	0.64	1.03		
C.	MLSF	3	4	0	1	0	2	0	10	0	0	0	0	0	0	1	0	2	1	24	31.01
	exp.	1.38	2.59	1.27	1.43	0.28	0.11	3.31	5.68	0.55	0.83	0.72	0.72	0.61	0.00	3.26	0.55	0.28	0.44		
D.	MLBW	1	18	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	26	85.30
	exp.	1.49	2.81	1.37	1.55	0.30	0.12	3.59	6.16	0.60	0.90	0.78	0.78	0.66	0.00	3.53	0.60	0.30	0.48		
E.	MLRND	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	87.07
	exp.	0.29	0.54	0.26	0.30	0.06	0.02	0.69	1.18	0.11	0.17	0.15	0.15	0.13	0.00	0.68	0.11	0.06	0.09		
F.	MLMP	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	2	78.69
	exp.	0.11	0.22	0.11	0.12	0.02	0.01	0.28	0.47	0.05	0.07	0.06	0.06	0.05	0.00	0.27	0.05	0.02	0.04		
G.	FLT	1	5	0	0	1	0	0	9	0	0	1	0	0	0	40	4	0	1	62	137.96
	exp.	3.56	6.70	3.28	3.71	0.71	0.29	8.55	14.68	1.43	2.14	1.85	1.85	1.57	0.00	8.41	1.43	0.71	1.14		
H.	MLDVA	5	3	8	6	2	0	43	0	7	8	3	1	0	0	10	4	1	2	103	101.62
	exp.	5.92	11.13	5.45	6.16	1.18	0.47	14.21	24.39	2.37	3.55	3.08	3.08	2.60	0.00	13.97	2.37	1.18	1.89		
I.	MLWA	1	0	0	0	0	0	0	8	0	1	0	0	0	0	0	0	0	0	10	21.39
	exp.	0.57	1.08	0.53	0.60	0.11	0.05	1.38	2.37	0.23	0.34	0.30	0.30	0.25	0.00	1.36	0.23	0.11	0.18		
J.	MOF	0	0	2	0	0	0	0	10	0	0	2	0	0	0	0	0	0	0	14	23.02
	exp.	0.80	1.51	0.74	0.84	0.16	0.06	1.93	3.31	0.32	0.48	0.42	0.42	0.35	0.00	1.90	0.32	0.16	0.26		
K.	FTF	0	0	0	0	0	0	0	1	0	0	0	12	0	0	0	0	0	0	13	326.83
	exp.	0.75	1.40	0.69	0.78	0.15	0.06	1.80	3.08	0.30	0.45	0.39	0.39	0.33	0.00	1.76	0.30	0.15	0.24		
L.	STS	0	0	0	0	0	0	0	0	1	0	0	0	11	0	2	0	0	0	14	300.15
	exp.	0.80	1.51	0.74	0.84	0.16	0.06	1.93	3.31	0.32	0.48	0.42	0.42	0.35	0.00	1.90	0.32	0.16	0.26		
N.	REP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	exp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
O.	FEW	0	3	2	0	0	0	0	45	2	2	0	0	0	0	0	1	0	1	56	95.76
	exp.	3.22	6.05	2.96	3.35	0.64	0.26	7.72	13.26	1.29	1.93	1.67	1.67	1.42	0.00	7.60	1.29	0.64	1.03		
P.	FEOM	0	0	0	0	0	0	0	1	0	1	5	0	0	0	2	0	1	0	10	68.98
	exp.	0.57	1.08	0.53	0.60	0.11	0.05	1.38	2.37	0.23	0.34	0.30	0.30	0.25	0.00	1.36	0.23	0.11	0.18		
Q.	FEAM	0	0	0	0	0	0	0	1	0	2	2	0	0	0	0	0	0	0	5	46.15
	exp.	0.29	0.54	0.26	0.30	0.06	0.02	0.69	1.18	0.11	0.17	0.15	0.15	0.13	0.00	0.68	0.11	0.06	0.09		
R.	FESIG	0	3	0	0	0	0	0	1	0	1	0	0	0	0	0	1	1	0	7	19.91
	exp.	0.40	0.76	0.37	0.42	0.08	0.03	0.97	1.66	0.16	0.24	0.21	0.21	0.18	0.00	0.95	0.16	0.08	0.13		
Total:		25	47	23	26	5	2	60	103	10	15	13	13	11	0	59	10	5	8	435	1577.00
Frequency		0.06	0.11	0.05	0.06	0.01	0.00	0.14	0.24	0.02	0.03	0.03	0.03	0.03	0.00	0.14	0.02	0.01	0.02		

Behaviours are identified by their letter codes and acronyms
(for example A = MLSTA). The top number in a row is the observed value and the bottom number the expected value. χ^2 values are given in the far right column.

Table 6.7. Allopatric *T. saeva* female × *T. gigantea* male (TSf × TGm) transition matrix ($n = 9$).

Allopatric TSf x TGm — Following acts

	Preceding acts	A. MLSTA	B. MLCE	C. MLSF	D. MLBW	E. MLRND	F. MLMP	G. FLT	H. MLDVA	I. MLWA	J. MOF	K. FTF	L. STS	M. ACOP	N. REP	O. FEW	P. FEOM	Q. FEAM	R. FESIG	Row Total	Row χ^2
A.	MLSTA	0	7	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	exp.	0.32	1.05	0.41	0.24	0.04	0.00	1.38	1.82	0.12	0.24	0.41	0.36	0.36	0.00	1.14	0.57	0.12	0.41	9	41.32
B.	MLCE	6	0	2	2	1	0	14	2	0	0	0	0	0	0	1	3	0	1		
	exp.	1.15	3.75	1.44	0.86	0.14	0.00	4.90	6.49	0.43	0.86	1.44	1.30	1.30	0.00	4.04	2.02	0.43	1.44	32	41.57
C.	MLSF	0	5	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0		
	exp.	0.36	1.17	0.45	0.27	0.05	0.00	1.53	2.03	0.14	0.27	0.45	0.41	0.41	0.00	1.26	0.63	0.14	0.45	10	20.73
D.	MLBW	0	4	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0		
	exp.	0.22	0.70	0.27	0.16	0.03	0.00	0.92	1.22	0.08	0.16	0.27	0.24	0.24	0.00	0.76	0.38	0.08	0.27	6	27.03
E.	MLRND	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	exp.	0.04	0.12	0.05	0.03	0.00	0.00	0.15	0.20	0.01	0.03	0.05	0.04	0.04	0.00	0.13	0.06	0.01	0.05	1	142.87
F.	MLMP	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0		
	exp.	0.07	0.23	0.09	0.05	0.01	0.00	0.31	0.41	0.03	0.05	0.09	0.08	0.08	0.00	0.25	0.13	0.03	0.09	2	64.93
G.	FLT	0	3	0	0	0	0	0	1	0	0	3	0	0	0	18	5	0	3		
	exp.	1.19	3.86	1.49	0.89	0.15	0.00	5.05	6.69	0.45	0.89	1.49	1.34	1.34	0.00	4.16	2.08	0.45	1.49	33	58.39
H.	MLDVA	0	2	3	3	0	0	18	0	1	3	1	0	0	0	5	6	1	3		
	exp.	1.66	5.39	2.07	1.24	0.21	0.00	7.05	9.32	0.62	1.25	2.07	1.86	1.86	0.00	5.80	2.90	0.62	2.07	46	33.90
I.	MLWA	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0		
	exp.	0.11	0.35	0.14	0.08	0.01	0.00	0.46	0.61	0.04	0.08	0.14	0.12	0.12	0.00	0.38	0.19	0.04	0.14	3	40.91
J.	MOF	0	0	0	0	0	0	0	5	0	0	0	0	0	0	1	0	0	0		
	exp.	0.22	0.70	0.27	0.16	0.03	0.00	0.92	1.22	0.08	0.16	0.27	0.24	0.24	0.00	0.76	0.38	0.08	0.27	6	24.76
K.	FTF	0	0	0	0	0	0	0	0	0	0	0	9	0	0	1	0	0	0		
	exp.	0.36	1.17	0.45	0.27	0.05	0.00	1.53	2.03	0.14	0.27	0.45	0.41	0.41	0.00	1.26	0.63	0.14	0.45	10	170.97
L.	STS	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0		
	exp.	0.32	1.05	0.41	0.24	0.04	0.00	1.38	1.82	0.12	0.24	0.41	0.36	0.36	0.00	1.14	0.57	0.12	0.41	9	191.87
N.	REP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	exp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	
O.	FEW	1	3	1	0	0	0	1	18	1	2	0	0	0	0	0	0	0	1		
	exp.	1.01	3.28	1.26	0.76	0.13	0.00	4.29	5.68	0.38	0.76	1.26	1.14	1.14	0.00	3.53	1.77	0.38	1.26	28	33.47
P.	FEOM	0	1	0	0	0	0	0	3	0	0	6	0	0	0	1	0	2	1		
	exp.	0.50	1.64	0.63	0.38	0.06	0.00	2.14	2.84	0.19	0.38	0.63	0.57	0.57	0.00	1.77	0.88	0.19	0.63	14	51.89
Q.	FEAM	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1		
	exp.	0.11	0.35	0.14	0.08	0.01	0.00	0.46	0.61	0.04	0.08	0.14	0.12	0.12	0.00	0.38	0.19	0.04	0.14	3	40.91
R.	FESIG	0	1	1	1	0	0	0	5	1	0	0	0	0	0	1	0	0	0		
	exp.	0.36	1.17	0.45	0.27	0.05	0.00	1.53	2.03	0.14	0.27	0.45	0.41	0.41	0.00	1.26	0.63	0.14	0.45	10	10.95
Total:		8	26	10	6	1	0	34	45	3	6	10	9	9	0	28	14	3	10	222	996.48
Frequency		0.04	0.12	0.05	0.03	0.00	0.00	0.15	0.20	0.01	0.03	0.05	0.04	0.04	0.00	0.13	0.06	0.01	0.05		

Behaviours are identified by their letter codes and acronyms (for example A = MLSTA). The top number in a row is the observed value and the bottom number the expected value. χ^2 values are given in the far right column.

Table 6.8. Parapatric *T. saeva* female × *T. saeva* male (TSf × TSm) transition matrix ($n = 12$).

Parapatric TSf x TSm		Following acts																			
	Preceding acts	A. MLSTA	B. MLCE	C. MLSF	D. MLBW	E. MLRND	F. MLMP	G. FLT	H. MLDVA	I. MLWA	J. MOF	K. FTF	L. STS	M. ACOP	N. REP	O. FEW	P. FEOM	Q. FEAM	R. FESIG	Row Total	Row χ^2
A.	MLSTA	0	27	3	0	2	0	0	7	0	0	0	0	0	0	1	1	1	0	42	59.67
	exp.	3.56	7.63	1.73	0.51	1.22	0.20	4.88	6.81	0.92	1.02	1.32	1.32	1.22	0.20	4.98	1.73	1.22	1.53		
B.	MLCE	11	0	9	5	4	2	12	1	0	0	0	0	0	1	9	4	6	6	79	78.40
	exp.	6.69	14.35	3.25	0.96	2.30	0.38	9.18	12.82	1.72	1.91	2.49	2.49	2.30	0.38	9.37	3.25	2.30	2.87		
C.	MLSF	3	9	0	0	0	0	0	4	0	0	0	0	0	0	1	0	0	0	17	16.57
	exp.	1.44	3.09	0.70	0.21	0.49	0.08	1.98	2.76	0.37	0.41	0.54	0.54	0.49	0.08	2.02	0.70	0.49	0.62		
D.	MLBW	0	4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	5	37.53
	exp.	0.42	0.91	0.21	0.06	0.15	0.02	0.58	0.81	0.11	0.12	0.16	0.16	0.15	0.02	0.59	0.21	0.15	0.18		
E.	MLRND	9	1	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	13	57.43
	exp.	1.10	2.36	0.54	0.16	0.38	0.06	1.51	2.11	0.28	0.31	0.41	0.41	0.38	0.06	1.54	0.54	0.38	0.47		
F.	MLMP	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	92.28
	exp.	0.17	0.36	0.08	0.02	0.06	0.01	0.23	0.32	0.04	0.05	0.06	0.06	0.06	0.01	0.24	0.08	0.06	0.07		
G.	FLT	2	3	0	0	2	0	0	8	0	0	3	0	0	0	22	6	0	1	47	64.91
	exp.	3.98	8.54	1.93	0.57	1.37	0.23	5.46	7.62	1.02	1.14	1.48	1.48	1.37	0.23	5.58	1.93	1.37	1.71		
H.	MLDVA	3	4	2	0	1	0	20	0	8	6	5	1	0	1	11	4	3	1	70	70.01
	exp.	5.93	12.71	2.88	0.85	2.03	0.34	8.14	11.36	1.53	1.69	2.20	2.20	2.03	0.34	8.31	2.88	2.03	2.54		
I.	MLWA	1	2	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	9	24.83
	exp.	0.76	1.63	0.37	0.11	0.26	0.04	1.05	1.46	0.20	0.22	0.28	0.28	0.26	0.04	1.07	0.37	0.26	0.33		
J.	MOF	0	1	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	10	41.14
	exp.	0.85	1.82	0.41	0.12	0.29	0.05	1.16	1.62	0.22	0.24	0.31	0.31	0.29	0.05	1.19	0.41	0.29	0.36		
K.	FTF	0	0	0	0	0	0	0	1	0	0	0	12	0	0	0	0	0	0	13	311.19
	exp.	1.10	2.36	0.54	0.16	0.38	0.06	1.51	2.11	0.28	0.31	0.41	0.41	0.38	0.06	1.54	0.54	0.38	0.47		
L.	STS	0	0	0	0	0	0	0	0	0	0	0	0	12	0	1	0	0	0	13	338.42
	exp.	1.10	2.36	0.54	0.16	0.38	0.06	1.51	2.11	0.28	0.31	0.41	0.41	0.38	0.06	1.54	0.54	0.38	0.47		
N.	REP	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	92.28
	exp.	0.17	0.36	0.08	0.02	0.06	0.01	0.23	0.32	0.04	0.05	0.06	0.06	0.06	0.01	0.24	0.08	0.06	0.07		
O.	FEW	3	11	2	0	3	0	3	22	0	2	0	0	0	0	0	0	0	2	48	35.74
	exp.	4.07	8.72	1.98	0.58	1.39	0.23	5.58	7.79	1.05	1.16	1.51	1.51	1.39	0.23	5.69	1.98	1.39	1.74		
P.	FEOM	0	5	0	0	0	0	1	2	0	1	3	0	0	0	1	0	0	4	17	27.04
	exp.	1.44	3.09	0.70	0.21	0.49	0.08	1.98	2.76	0.37	0.41	0.54	0.54	0.49	0.08	2.02	0.70	0.49	0.62		
Q.	FEAM	1	4	1	0	0	0	1	2	0	1	1	0	0	0	0	0	0	1	12	9.98
	exp.	1.02	2.18	0.49	0.15	0.35	0.06	1.39	1.95	0.26	0.29	0.38	0.38	0.35	0.06	1.42	0.49	0.35	0.44		
R.	FESIG	1	3	0	0	0	0	1	4	1	0	1	0	0	0	0	2	1	0	14	9.50
	exp.	1.19	2.54	0.58	0.17	0.41	0.07	1.63	2.27	0.31	0.34	0.44	0.44	0.41	0.07	1.66	0.58	0.41	0.51		
Total:		35	75	17	5	12	2	48	67	9	10	13	13	12	2	49	17	12	15	413	1366.90
Frequency		0.08	0.18	0.04	0.01	0.03	0.00	0.12	0.16	0.02	0.02	0.03	0.03	0.03	0.00	0.12	0.04	0.03	0.04		

Behaviours are identified by their letter codes and acronyms
(for example A = MLSTA). The top number in a row is the observed value and the bottom number the expected value. χ^2 values are given in the far right column.

Table 6.9. Parapatric *T. gigantea* female × *T. gigantea* male (TGf × TGm) transition matrix ($n = 8$).

Parapatric TGf x TGm — Following acts

	Preceding acts	A. MLSTA	B. MLCE	C. MLSF	D. MLBW	E. MLRND	F. MLMP	G. FLT	H. MLDVA	I. MLWA	J. MOF	K. FTF	L. STS	M. ACOP	N. REP	O. FEW	P. FEOM	Q. FEAM	R. FESIG	Row Total	Row $\chi 2$
A.	MLSTA	0	6	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	11	37.94
	exp.	0.43	0.70	0.22	0.22	0.00	0.00	1.68	2.38	0.16	0.33	0.65	0.54	0.43	0.16	1.90	0.70	0.22	0.27		
B.	MLCE	3	0	1	1	0	0	6	1	0	0	0	0	0	0	2	3	0	1	18	11.94
	exp.	0.71	1.15	0.35	0.35	0.00	0.00	2.75	3.90	0.27	0.53	1.06	0.89	0.71	0.27	3.10	1.15	0.35	0.44		
C.	MLSF	0	1	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	4	42.51
	exp.	0.16	0.26	0.08	0.08	0.00	0.00	0.61	0.87	0.06	0.12	0.24	0.20	0.16	0.06	0.69	0.26	0.08	0.10		
D.	MLBW	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	4	19.14
	exp.	0.16	0.26	0.08	0.08	0.00	0.00	0.61	0.87	0.06	0.12	0.24	0.20	0.16	0.06	0.69	0.26	0.08	0.10		
E.	MLRND	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	exp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
F.	MLMP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	exp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
G.	FLT	0	1	0	0	0	0	0	2	1	0	3	0	0	0	20	4	0	0	31	47.33
	exp.	1.22	1.99	0.61	0.61	0.00	0.00	4.73	6.72	0.46	0.92	1.83	1.53	1.22	0.46	5.34	1.99	0.61	0.76		
H.	MLDVA	3	0	0	1	0	0	20	0	2	2	1	0	0	2	6	4	1	2	44	40.94
	exp.	1.73	2.82	0.87	0.87	0.00	0.00	6.72	9.54	0.65	1.30	2.60	2.17	1.73	0.65	7.59	2.82	0.87	1.08		
I.	MLWA	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	1	0	0	4	19.44
	exp.	0.16	0.26	0.08	0.08	0.00	0.00	0.61	0.87	0.06	0.12	0.24	0.20	0.16	0.06	0.70	0.26	0.08	0.10		
J.	MOF	0	1	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	6	18.14
	exp.	0.24	0.38	0.12	0.12	0.00	0.00	0.92	1.30	0.09	0.18	0.35	0.30	0.24	0.09	1.03	0.38	0.12	0.15		
K.	FTF	0	0	0	0	0	0	0	0	0	0	0	10	0	0	2	0	0	0	12	139.53
	exp.	0.47	0.77	0.24	0.24	0.00	0.00	1.83	2.60	0.18	0.35	0.71	0.59	0.47	0.18	2.07	0.77	0.24	0.30		
L.	STS	0	0	0	0	0	0	0	0	0	0	0	0	8	0	2	0	0	0	10	133.71
	exp.	0.39	0.64	0.20	0.20	0.00	0.00	1.53	2.17	0.15	0.30	0.59	0.49	0.39	0.15	1.72	0.64	0.20	0.25		
N.	REP	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	3	28.51
	exp.	0.12	0.19	0.06	0.06	0.00	0.00	0.46	0.65	0.04	0.09	0.18	0.15	0.12	0.04	0.52	0.19	0.06	0.07		
O.	FEW	1	2	0	0	0	0	5	22	0	2	0	0	0	0	0	1	0	1	34	35.29
	exp.	1.34	2.18	0.67	0.67	0.00	0.00	5.19	7.37	0.50	1.00	2.01	1.67	1.34	0.50	5.86	2.18	0.67	0.84		
P.	FEOM	0	1	0	0	0	0	0	1	0	0	6	0	0	1	1	0	2	1	13	39.03
	exp.	0.51	0.83	0.26	0.26	0.00	0.00	1.99	2.82	0.19	0.38	0.77	0.64	0.51	0.19	2.24	0.83	0.26	0.32		
Q.	FEAM	0	0	0	0	0	0	0	1	0	0	2	0	0	0	1	0	0	0	4	25.59
	exp.	0.16	0.26	0.08	0.08	0.00	0.00	0.61	0.87	0.06	0.12	0.24	0.20	0.16	0.06	0.69	0.26	0.08	0.10		
R.	FESIG	0	0	0	0	0	0	0	4	0	0	0	0	0	0	1	0	0	0	5	18.98
	exp.	0.20	0.32	0.10	0.10	0.00	0.00	0.76	1.08	0.07	0.15	0.30	0.25	0.20	0.07	0.86	0.32	0.10	0.12		
Total:		8	13	4	4	0	0	31	44	3	6	12	10	8	3	35	13	4	5	203	658.04
Frequency		0.04	0.06	0.02	0.02	0.00	0.00	0.15	0.22	0.01	0.03	0.06	0.05	0.04	0.01	0.17	0.06	0.02	0.02		

Behaviours are identified by their letter codes and acronyms
(for example A = MLSTA). The top number in a row is the observed value and the bottom number the expected value. χ^2 values are given in the far right column.

Table 6.10. Parapatric *T. gigantea* female × *T. saeva* male (TGf × TSm) transition matrix (n = 8).

Parapatric TGf x TSm — Following acts

	Preceding acts	A. MLSTA	B. MLCE	C. MLSF	D. MLBW	E. MLRND	F. MLMP	G. FLT	H. MLDVA	I. MLWA	J. MOF	K. FTF	L. STS	M. ACOP	N. REP	O. FEW	P. FEOM	Q. FEAM	R. FESIG	Row Total	Row χ^2
A.	MLSTA	0	7	2	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	12	30.38
	exp.	0.46	1.57	0.15	0.25	0.20	0.00	1.72	2.28	0.25	0.20	0.41	0.41	0.41	0.35	2.18	0.51	0.25	0.41		
B.	MLCE	2	0	0	3	2	0	12	0	0	0	0	0	0	3	2	7	3	3	37	48.79
	exp.	1.41	4.84	0.47	0.78	0.62	0.00	5.31	7.03	0.78	0.62	1.25	1.25	1.25	1.09	6.71	1.56	0.78	1.25		
C.	MLSF	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	3	33.29
	exp.	0.11	0.39	0.04	0.06	0.05	0.00	0.43	0.57	0.06	0.05	0.10	0.10	0.10	0.09	0.54	0.13	0.06	0.10		
D.	MLBW	0	4	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	5	28.72
	exp.	0.19	0.65	0.06	0.11	0.08	0.00	0.72	0.95	0.11	0.08	0.17	0.17	0.17	0.15	0.91	0.21	0.11	0.17		
E.	MLRND	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	57.97
	exp.	0.15	0.52	0.05	0.08	0.07	0.00	0.57	0.76	0.08	0.07	0.14	0.14	0.14	0.12	0.73	0.17	0.08	0.14		
F.	MLMP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	exp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
G.	FLT	0	0	0	0	0	0	0	3	0	0	1	0	0	0	28	2	0	0	34	84.72
	exp.	1.29	4.45	0.43	0.72	0.57	0.00	4.88	6.46	0.72	0.57	1.15	1.15	1.15	1.00	6.17	1.43	0.72	1.15		
H.	MLDVA	0	0	0	1	1	0	21	0	4	2	4	0	0	2	7	1	1	1	45	56.12
	exp.	1.71	5.89	0.57	0.95	0.76	0.00	6.46	8.54	0.95	0.76	1.52	1.52	1.52	1.33	8.16	1.90	0.95	1.52		
I.	MLWA	1	0	0	0	0	0	0	3	0	1	0	0	0	0	0	0	0	0	5	18.69
	exp.	0.19	0.65	0.06	0.11	0.08	0.00	0.72	0.95	0.11	0.08	0.17	0.17	0.17	0.15	0.91	0.21	0.11	0.17		
J.	MOF	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	4	32.27
	exp.	0.15	0.52	0.05	0.08	0.07	0.00	0.57	0.76	0.08	0.07	0.14	0.14	0.14	0.12	0.73	0.17	0.08	0.14		
K.	FTF	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	8	202.11
	exp.	0.30	1.05	0.10	0.17	0.14	0.00	1.15	1.52	0.17	0.14	0.27	0.27	0.27	0.24	1.45	0.34	0.17	0.27		
L.	STS	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	8	202.11
	exp.	0.30	1.05	0.10	0.17	0.14	0.00	1.15	1.52	0.17	0.14	0.27	0.27	0.27	0.24	1.45	0.34	0.17	0.27		
N.	REP	3	2	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	7	28.53
	exp.	0.27	0.92	0.09	0.15	0.12	0.00	1.00	1.33	0.15	0.12	0.24	0.24	0.24	0.21	1.27	0.30	0.15	0.24		
O.	FEW	0	9	0	1	1	0	0	28	0	1	0	0	0	0	0	0	0	2	42	65.69
	exp.	1.59	5.49	0.53	0.89	0.71	0.00	6.03	7.97	0.89	0.71	1.42	1.42	1.42	1.24	7.62	1.77	0.89	1.42		
P.	FEOM	0	2	1	0	0	0	0	1	0	0	2	0	0	1	1	0	0	2	10	12.72
	exp.	0.38	1.31	0.13	0.21	0.17	0.00	1.43	1.90	0.21	0.17	0.34	0.34	0.34	0.30	1.81	0.42	0.21	0.34		
Q.	FEAM	0	1	0	0	0	0	1	0	0	0	1	0	0	0	2	0	0	0	5	16.58
	exp.	0.19	0.65	0.06	0.11	0.08	0.00	0.72	0.95	0.11	0.08	0.17	0.17	0.17	0.15	0.91	0.21	0.11	0.17		
R.	FESIG	0	4	0	0	0	0	0	2	0	0	0	0	0	1	1	0	0	0	8	12.90
	exp.	0.30	1.05	0.10	0.17	0.14	0.00	0.15	1.52	0.17	0.14	0.27	0.27	0.27	0.24	1.45	0.34	0.17	0.27		
Total:		9	31	3	5	4	0	34	45	5	4	8	8	8	7	43	10	5	8	237	931.59
Frequency		0.04	0.13	0.01	0.02	0.02	0.00	0.14	0.19	0.02	0.02	0.03	0.03	0.03	0.03	0.18	0.04	0.02	0.03		

Behaviours are identified by their letter codes and acronyms
(for example A = MLSTA). The top number in a row is the observed value and the bottom number the expected value. χ^2 values are given in the far right column.

Table 6.11. Parapatric *T. saeva* female × *T. gigantea* male (TSf × TGm) transition matrix ($n = 10$).

Parapatric TSf x TGm		Following acts																			
	Preceding acts	A. MLSTA	B. MLCE	C. MLSF	D. MLBW	E. MLRND	F. MLMP	G. FLT	H. MLDVA	I. MLWA	J. MOF	K. FTF	L. STS	M. ACOP	N. REP	O. FEW	P. FEOM	Q. FEAM	R. FESIG	Row Total	Row χ^2
A.	MLSTA	0	5	2	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	9	33.23
	exp.	0.32	0.92	0.24	0.28	0.36	0.16	1.41	1.77	0.08	0.16	0.56	0.56	0.40	0.04	1.04	0.40	0.08	0.20		
B.	MLCE	0	0	1	3	3	2	13	1	0	0	0	0	0	0	4	3	1	1	32	27.63
	exp.	1.14	3.29	0.86	1.00	1.29	0.57	5.00	6.29	0.29	0.57	2.00	2.00	1.43	0.14	3.71	1.43	0.29	0.71		
C.	MLSF	1	4	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	6	35.91
	exp.	0.21	0.62	0.16	0.19	0.24	0.11	0.94	1.18	0.05	0.11	0.38	0.38	0.27	0.03	0.70	0.27	0.05	0.13		
D.	MLBW	0	3	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	7	25.85
	exp.	0.25	0.72	0.19	0.22	0.28	0.13	1.09	1.38	0.06	0.13	0.44	0.44	0.31	0.03	0.81	0.31	0.06	0.16		
E.	MLRND	5	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	9	67.77
	exp.	0.32	0.92	0.24	0.28	0.36	0.16	1.4	1.77	0.08	0.16	0.56	0.56	0.40	0.04	1.04	0.40	0.08	0.20		
F.	MLMP	0	1	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	4	38.65
	exp.	0.14	0.41	0.11	0.13	0.16	0.07	0.63	0.79	0.04	0.07	0.25	0.25	0.18	0.02	0.46	0.18	0.04	0.09		
G.	FLT	0	2	0	0	3	0	0	2	0	0	8	0	0	1	15	6	0	0	37	58.35
	exp.	1.32	3.80	0.99	1.16	1.49	0.66	5.78	7.27	0.33	0.66	2.31	2.31	1.65	0.17	4.29	1.65	0.33	0.83		
H.	MLDVA	1	1	1	4	2	2	19	0	1	4	1	0	0	0	4	1	0	3	44	49.34
	exp.	1.57	4.52	1.18	1.38	1.77	0.79	6.88	8.64	0.39	0.79	2.75	2.75	1.96	0.20	5.11	1.96	0.39	0.98		
I.	MLWA	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2	91.30
	exp.	0.07	0.21	0.05	0.06	0.08	0.04	0.31	0.39	0.02	0.04	0.13	0.13	0.09	0.01	0.23	0.09	0.02	0.04		
J.	MOF	0	0	0	0	0	0	2	3	0	0	0	0	0	0	0	0	0	0	5	31.35
	exp.	0.18	0.51	0.13	0.16	0.20	0.09	0.78	0.98	0.04	0.09	0.31	0.31	0.22	0.02	0.58	0.22	0.04	0.11		
K.	FTF	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	14	192.89
	exp.	0.50	1.44	0.38	0.44	0.56	0.25	2.19	2.75	0.13	0.25	0.88	0.88	0.63	0.06	1.63	0.63	0.13	0.31		
L.	STS	0	0	0	0	0	0	0	0	0	0	1	0	10	0	3	0	0	0	14	136.58
	exp.	0.50	1.44	0.38	0.44	0.56	0.25	2.19	2.75	0.13	0.25	0.88	0.88	0.63	0.06	1.63	0.63	0.13	0.31		
N.	REP	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	191.42
	exp.	0.04	0.10	0.03	0.03	0.04	0.02	0.16	0.20	0.01	0.02	0.06	0.06	0.04	0.00	0.12	0.04	0.01	0.02		
O.	FEW	0	4	1	0	1	0	0	17	0	0	0	0	0	0	0	0	0	0	23	41.35
	exp.	0.82	2.36	0.62	0.72	0.92	0.41	3.59	4.52	0.21	0.41	1.44	1.44	1.03	0.10	2.67	1.03	0.21	0.51		
P.	FEOM	0	2	0	0	0	0	1	2	0	0	4	0	0	0	0	0	0	1	10	24.29
	exp.	0.38	1.03	0.27	0.31	0.40	0.18	1.56	1.96	0.09	0.18	0.63	0.63	0.45	0.04	1.16	0.45	0.09	0.22		
Q.	FEAM	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	88.21
	exp.	0.07	0.21	0.05	0.06	0.08	0.04	0.31	0.39	0.02	0.04	0.13	0.13	0.09	0.01	0.23	0.09	0.02	0.04		
R.	FESIG	0	0	0	0	0	0	0	3	1	0	0	0	0	0	0	0	1	0	5	30.79
	exp.	0.18	0.51	0.13	0.16	0.20	0.09	0.78	0.98	0.04	0.09	0.31	0.31	0.22	0.02	0.58	0.22	0.04	0.11		
Total:		8	23	6	7	9	4	35	44	2	4	14	14	10	1	26	10	2	5	224	1164.90
Frequency		0.04	0.10	0.03	0.03	0.04	0.02	0.16	0.20	0.01	0.02	0.06	0.06	0.04	0.04	0.12	0.04	0.01	0.02		

Behaviours are identified by their letter codes and acronyms
(for example A = MLSTA). The top number in a row is the observed value and the bottom number the expected value. χ^2 values are given in the far right column.

Graphical representations of the important transitions could be generated. for comparison of the pairing combinations, by reconstructing sequence diagrams from all significant (facilitating) transitions. The transition probabilties - the frequency at which each transition from one behaviour to the next occurred - were also included in these diagrams (Figures 6.3 to 6.6). As an aid to interpretation a summary of the behaviours outlined in section 6.3.1 is provided in Table 6.12. Each figure will be taken in turn:

Table 6.12. Summary of courtship behaviours and their acronyms.

Single letter code	Acronym	Behaviour
A.	MLSTA	MALE STATIONARY
B.	MLCE	MALE CHEMOEXPLORE
C.	MLSF	MALE SHAKE/FLEX WEB
D.	MLBW	MALE BITE WEB
E.	MLRND	MALE RUNS RANDOMLY
F.	MLMP	MALE MOUTHS PALPS/FEET
G.	FLT	FRONT LEG TOUCH
H.	MLDVA	MALE DRUM, VIBRATE ABDOMEN
I.	MLWA	MALE WALK AWAY
J.	MLOF	MALE ORIENTATE TO FEMALE
K.	FTF	FACE-TO-FACE
L.	STS	SIDE-TO-SIDE
M.	ACOP	ATTEMPTED COPULATION
N	REP	MALE REPELLED
O.	FEW	FEMALE WALK
P	FEOM	FEMALE ORIENTATES TO MALE
Q.	FEAM	FEMALE APPROACH MALE
R.	FESIG	FEMALE SIGNAL

FIGURE 6.3. TSF × TSM PAIRINGS:

Phase I courtship behaviours showed major differences between allopatric and parapatric pairings. Allopatric pairings showed frequent transitions between MLSTA and MLSF whereas in parapatry transitions between MLCE and MLSF

A: Allopatric TSf x TSm

REP
MLBW
MLSTA
1.0
MLRND
MLMP
.48
.19
**
.38
MLCE
MLSF
Phase I
Phase II
Female Specific
Male/Female
** $P < 0.01$
All other lines
$P < 0.001$
.28
.08
**
FESIG
.63
**
.21
**
FLT
.64
FEOM
.44
FTF
.28
.81
.41
MLDVA
FEWA
STS
.77
MLWA
MLOF
FEAM
ACOP

B: Parapatric TSf x TSm

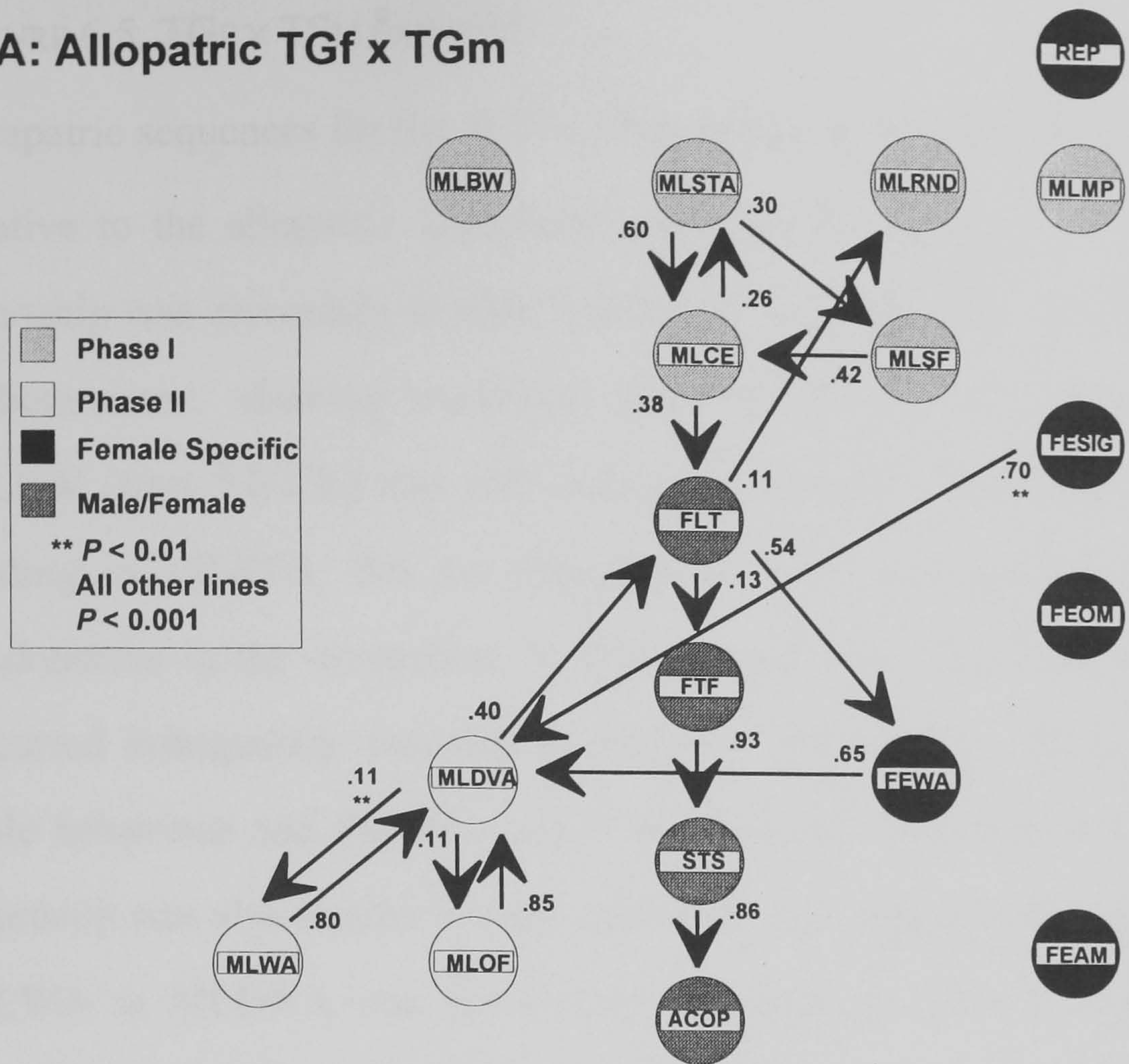
REP
MLBW
MLSTA
.69
MLRND
MLMP
.64
.80
**
.06
.53
**
MLCE
MLSF
Phase I
Phase II
Female Specific
Male/Female
** $P < 0.01$
All other lines
$P < 0.001$
.27
.11
**
FESIG
FLT
.47
.24
FEOM
.18
**
FTF
.29
.92
.46
.11
MLDVA
FEWA
.09
**
STS
.90
.92
.67
MLWA
MLOF
FEAM
ACOP

Figure 6.3. Sequence diagrams of pre-copulatory courtship transitions in allopatric and parapatric TSf × TSm pairings. Numbers indicate the frequency at which each transition from one behaviour to the next occurred.

(i.e. without the male pausing between behaviours) were more significant. Also transitions between MLCE and MLBW were important in parapatry but absent in allopatry. Phase II elements were more complex in parapatry, with transitions between MLDVA and MLWA, and between MLDVA and MLOF being significant. These behaviours, essentially representing the male briefly 'turning his back' to the female during the more directed Phase II courtship, were effectively absent in allopatry. Female specific behaviours were essentially similar between parapatry and allopatry except that FLT was less likely to lead to FEOM in parapatry (suggesting a more cautious female response).

FIGURE 6.4. TGF × TGM PAIRINGS:

The sequence diagram for the parapatric pairings was much simpler than for the allopatric pairings. This suggests an emphasis in a few transitions relative to others. Phase I elements were simplified in parapatry with fewer transitions to and from MLSTA. In allopatry MLSF followed from MLSTA and led directly to MLCE; in parapatry however, MLSF led to MLBW. The only other significant Phase I transition was from MLSTA to MLCE. Phase II sequences were also greatly simplified with transitions between MLDVA and MLWA, and between MLDVA and MLOF much reduced in parapatry; the qualitative opposite of the TSf × TSm case. Female and shared male/female behaviours also showed some small differences. In allopatry (male) FLT occasionally led to FTF (and on to ACOP) i.e. the male approached the female from the front. This was not the case in parapatry where FTF followed FEOM.



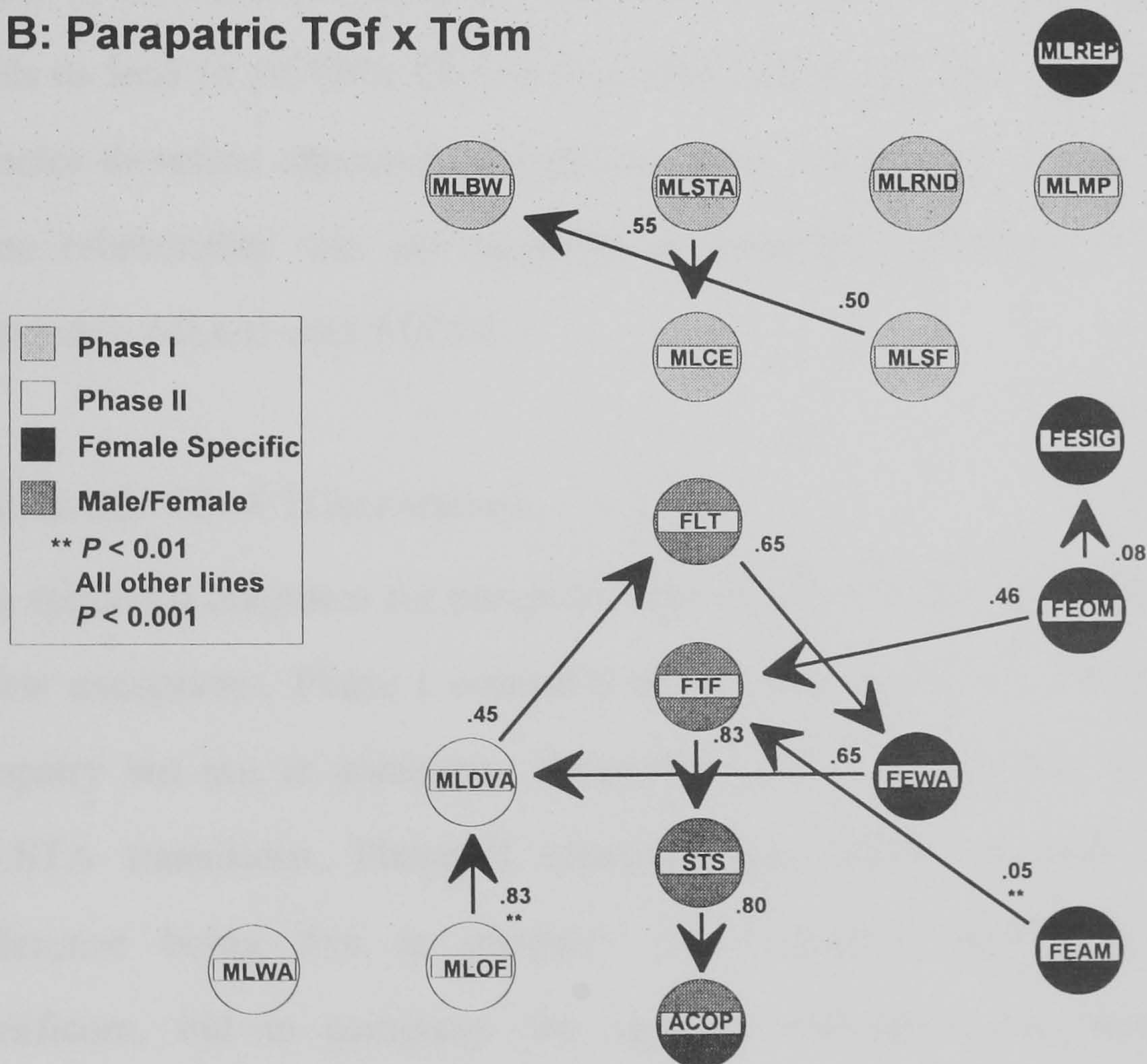


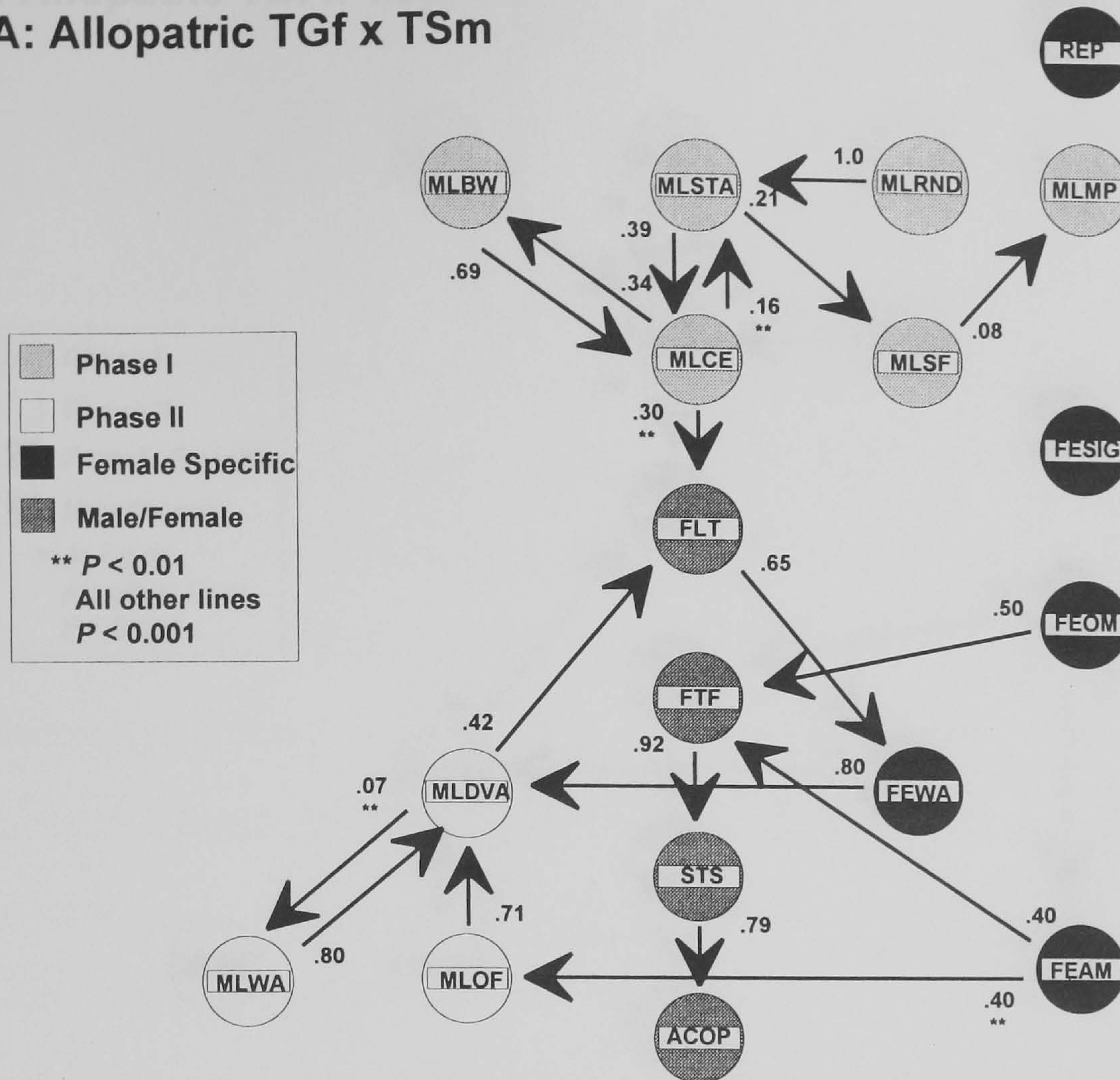
Figure 6.4. Sequence diagrams of pre-copulatory courtship transitions in allopatric and parapatric TGf × TGm pairings. Numbers indicate the frequency at which each transition from one behaviour to the next occurred.

FIGURE 6.5. TGF × TSM PAIRINGS:

Parapatric sequences for the TGf × TSm pairings also showed some simplification relative to the allopatric sequences, though this was not very marked. Phase I courtship was essentially similar in allopatry and parapatry. MLBW was a feature in both cases; showing transitions from MLBW to MLCE but the transition to MLBW (from MLCE) was only defined in allopatry. REP featured in parapatry, leading to MLSTA, but no transition to REP was defined suggesting some randomness in the occurrence of this element. The transition MLSF to MLMP occurred infrequently and only in allopatry. MLMP was effectively a stationary male behaviour and therefore could be regarded as similar to MLSTA. Phase II courtship was also similar in both cases with the exception that the transition from MLWA to MLDVA was not evident in parapatry. Female behaviours differed more than the male behaviours. In allopatry the female was more likely both to orient to the male (FEOM) and approach him (FEAM), resulting in FTF (which tends to lead to ACOP). FEAM also often led to MLOF. Allopatric *T. gigantea* females therefore appeared to make an active acceptance of *T. saeva* males. The same relationship was not apparent in parapatry, although females did often respond to MLCE with FEOM.

FIGURE 6.6. TSF × TGM PAIRINGS:

The sequence diagrams for parapatry and allopatry were qualitatively similar with a few exceptions. Phase I courtship showed the transition MLBW to MLCE in allopatry but not in parapatry. Parapatric males also showed more MLRND to MLSTA transitions. Phase II courtship was simple in both cases, the only difference being that in allopatry the transition MLOF to MLDVA was significant, but in parapatry the opposite transition was significant (but of marginal frequency). In parapatry, FLT was more likely to lead to FTF and to FEOM than in allopatry. But in allopatry FEOM was slightly more likely to lead



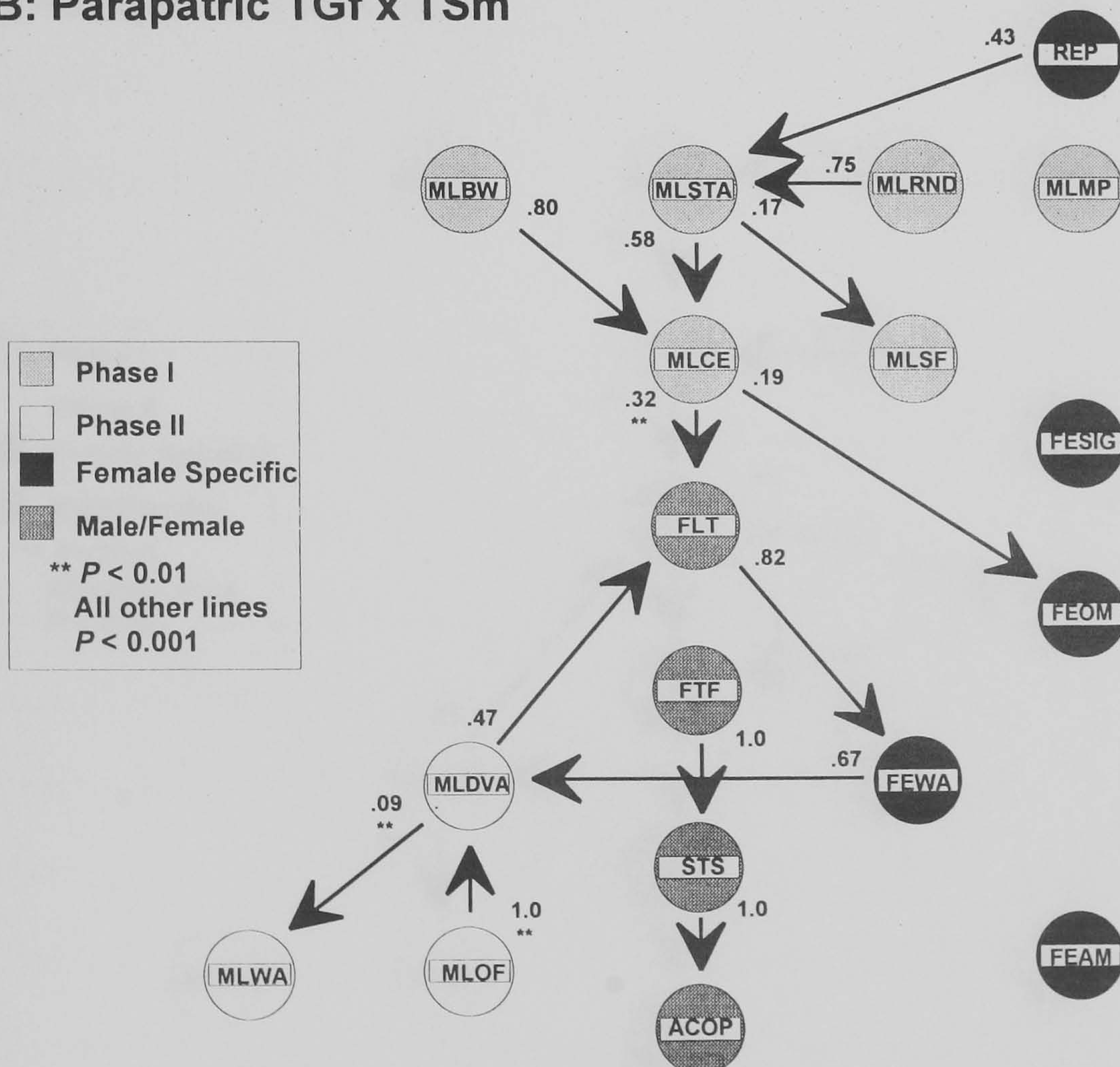
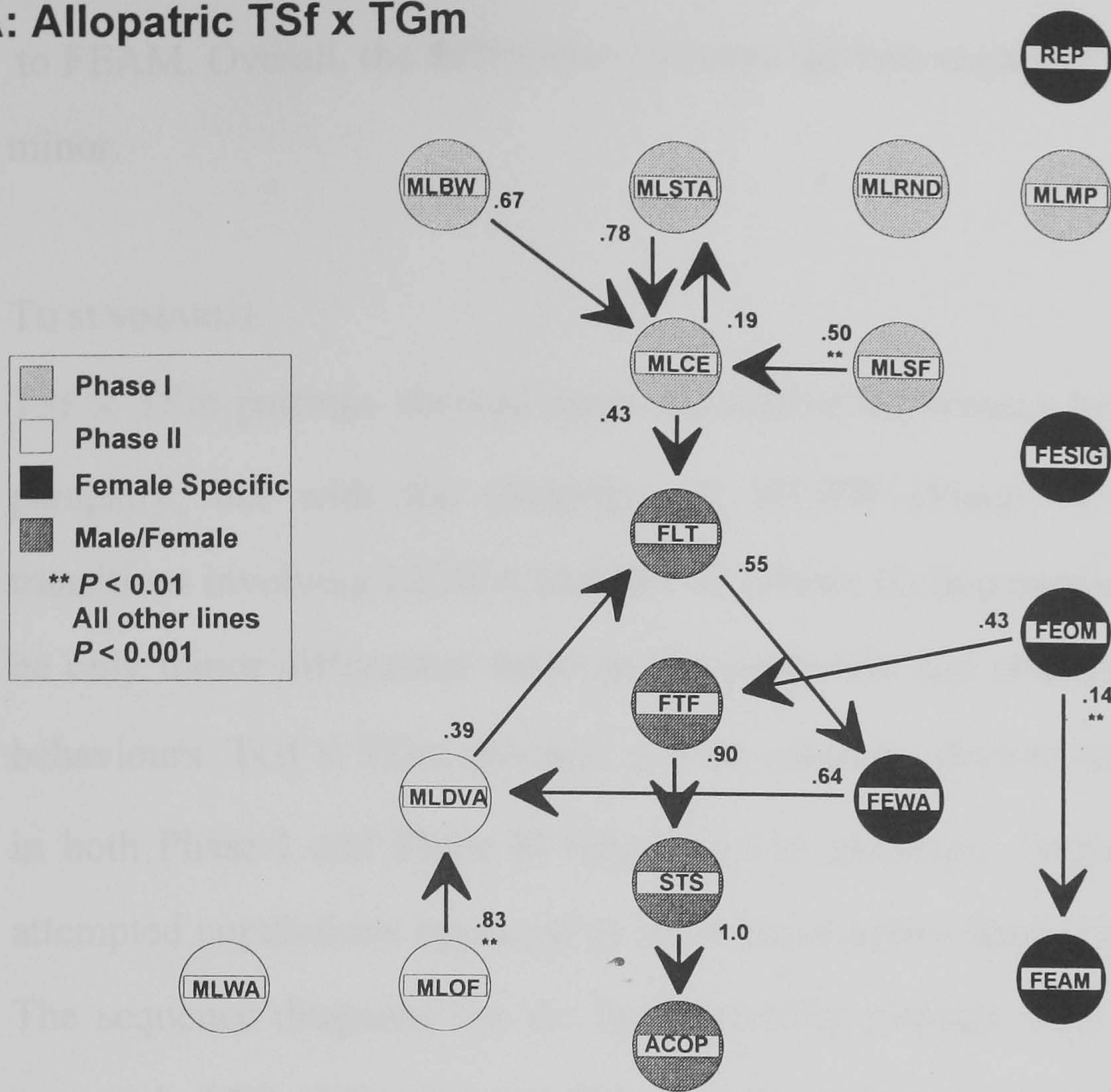


Figure 6.5. Sequence diagrams of pre-copulatory courtship transitions in allopatric and parapatric TGf × TSm pairings. Numbers indicate the frequency at which each transition from one behaviour to the next occurred.



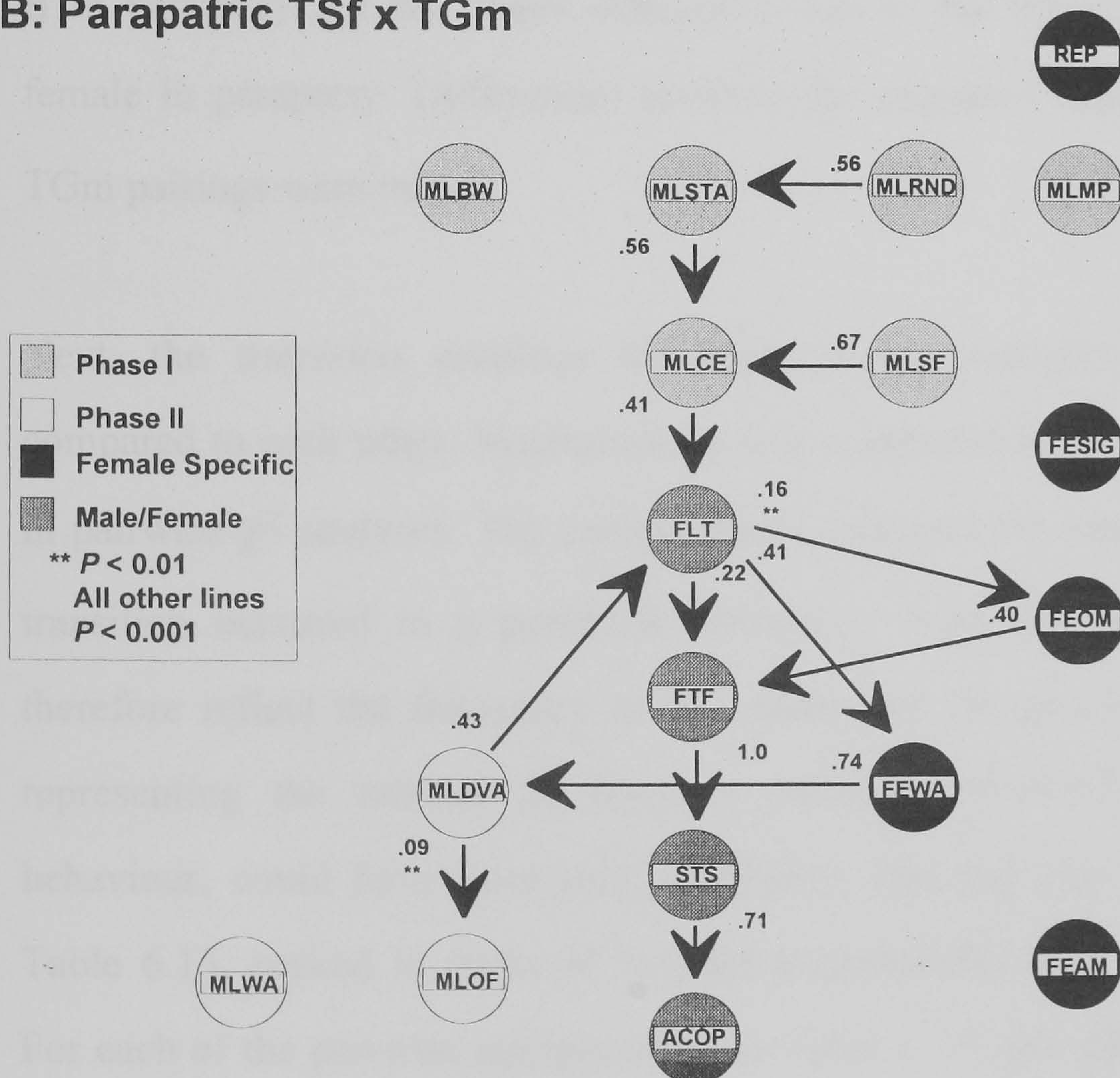


Figure 6.6. Sequence diagrams of pre-copulatory courtship transitions in allopatric and parapatric TSf × TGm pairings. Numbers indicate the frequency at which each transition from one behaviour to the next occurred.

to FEAM. Overall, the differences between the two sequence diagrams appear minor.

TO SUMMARIZE:

TSf × TSm pairings showed some qualitative differences between allopatry and parapatry, but with the exception of MLBW (Phase I) and the significant transitions involving MLWA and MLOF (Phase II) in parapatry, there appeared to be only minor differences between the parapatric and allopatric male and female behaviours. TGf × TGm pairings, on the contrary, showed marked simplification in both Phase I and Phase II behaviours in parapatry. Sequences leading up to attempted copulations appeared to show more active female choices in parapatry. The sequence diagrams for the heterospecific pairings were less revealing. The parapatric TGf × TSm diagram did show some simplification of Phase I and Phase II in parapatry but the major difference was in the reduced receptivity of the female in parapatry. Differences between the sequence diagrams for the TSf × TGm pairings were minor.

Next, the transition matrices for each pairing comparison were explicitly compared to each other. Matrices were first compared by using the column totals in pairwise χ^2 analyses. The column totals represent the total number of times a transition occurred to a particular behaviour from any other behaviour, and therefore reflect the frequency of that behaviour (though of course row totals, representing the number of times a transition occurred *from* a particular behaviour, could have been used similarly). The pairwise results are given in Table 6.13, ranked in order of increasing probability (decreasing significance). For each of the pairwise comparisons the value 1 - P was taken as analogous to a 'distance' measure and a pairwise matrix constructed. A dendrogram was then created using the neighbor-joining algorithm (Saitou and Nei, 1987) to illustrate the relationships between the pairing combinations. The dendrogram was

generated using the program NEIGHBOR from the phylogenetics package PHYLIP (Felsenstein, 1995) running under the Daresbury Laboratory Seqnet Service (Program Manual for the Wisconsin Package, 1994). The dendrogram was drawn using TREEVIEW (Page, 1996) and is shown in Figure 6.7. The comparisons between allopatric TSf × TSm and parapatric TSf × TSm (rank 25, Table 6.13), and between allopatric TGf × TGm and parapatric TGf × TGm (rank 19, Table 6.13), were not significant, but note that the *P* -value for the latter was 4.5 times smaller than for the former. This is reflected in the greater separation of allopatric TGf × TGm and parapatric TGf × TGm on the dendrogram than of allopatric TSf × TSm and parapatric TSf × TSm (Figure 6.7). Also note that both the comparisons between allopatric TSf × TSm and allopatric TGf × TGm (rank 10, Table 6.13), and between parapatric TSf × TSm and parapatric TGf × TGm (rank 2, Table 6.13), were highly significant, with the latter being some 27.4 times more significant than the former. These results are in broad agreement with the interpretation of the sequence diagrams; suggesting that there were differences between *T. gigantea* and *T. saeva* in terms of courtship behaviour and that these differences were greater in parapatry. However, it is not possible to tell whether these greater differences are due to exaggeration of behaviours distinguishing the species in allopatry, or whether they correspond to the exaggeration of different behaviours in parapatry. There were no significant differences between TSf × TGm heterospecific pairings in allopatry and parapatry (rank 24, Table 6.13), however there were very significant differences between TGf × TSm pairings in allopatry and parapatry (rank 3, Table 6.12). Indeed the *P* -value for the latter was 157.2 times smaller than that for the former. These features, reflected in the dendrogram in Figure 6.7, were again in agreement with the sequence diagrams.

The comparisons between the column totals of the transition matrices were taken further by performing a simultaneous *G*-test of heterogeneity (see section 6.3.2) across the transition matrices for each column (using the total frequency of each

Table 6.13. Pairwise comparisons between the transition matrices based on column totals.

Comparison		χ^2 (*df*)	*P*	Rank
Allopatric TGf × TSm:	Parapatric TSf × TSm	45.37 (17)	0.0002	1
Parapatric TSf × TSm:	Parapatric TGf × TGm	41.78 (17)	0.0007	2
Allopatric TGf × TSm:	Parapatric TGf × TSm	37.80 (17)	0.0026	3
Allopatric TGf × TSm:	Parapatric TGf × TGm	36.81 (17)	0.0036	4
Allopatric TGf × TGm:	Allopatric TGf ×TSm	36.07 (17)	0.0045	5
Allopatric TSf × TSm:	Allopatric TGf ×TSm	33.69 (16)	0.0060	6
Parapatric TSf × TSm:	Parapatric TSf × TGm	33.54 (17)	0.0096	7
Allopatric TGf ×TSm:	Parapatric TSf × TGm	33.44 (16)	0.0099	8
Allopatric TGf ×TGm:	Parapatric TSf ×TSm	32.86 (17)	0.0118	9
Allopatric TSf × TSm	Allopatric TGf × TGm	31.14 (17)	0.0192	10
Allopatric TGf × TGm:	Parapatric TSf × TGm	31.09 (17)	0.0195	11
Allopatric TGf × TGm:	Parapatric TGf × TSm	29.14 (17)	0.0332	12
Parapatric TSf ×TSm:	Parapatric TGf × TSm	27.38 (17)	0.0527	13
Allopatric TSf × TSm:	Parapatric TSf × TGm	27.34 (17)	0.0533	14
Allopatric TSf × TSm:	Parapatric TGf × TSm	27.20 (17)	0.0551	15
Allopatric TSf × TSm:	Parapatric TGf × TGm	26.80 (17)	0.0610	16
Allopatric TSf × TGm:	Parapatric TSf × TSm	25.47 (17)	0.0847	17
Allopatric TGf × TGm:	Allopatric TSf × TGm	25.09 (17)	0.0927	18
Allopatric TGf × TGm:	Parapatric TGf × TGm	24.75 (17)	0.1004	19
Parapatric TGf × TSm:	Parapatric TSf × TGm	23.63 (17)	0.1300	20
Allopatric TGf × TSm:	Allopatric TSf × TGm	21.60 (16)	0.1565	21
Parapatric TGf × TGm:	Parapatric TSf × TGm	21.29 (17)	0.2138	22
Allopatric TSf × TGm:	Parapatric TGf × TSm	18.48 (16)	0.2963	23
Allopatric TSf × TGm:	Parapatric TSf × TGm	17.69 (17)	0.4086	24
Allopatric TSf × TSm:	Parapatric TSf × TSm	17.09 (17)	0.4481	25
Allopatric TSf × TSm:	Allopatric TSf × TGm	15.75 (16)	0.4707	26
Parapatric TGf × TGm:	Parapatric TGf × TSm	14.82 (16)	0.5375	27
Allopatric TSf × TGm:	Parapatric TGf × TGm	13.55 (16)	0.6322	28

Given for each comparison are: the χ^2-value with the degrees of freedom in brackets; and the corresponding probability of the χ^2-value. The comparisons are ranked in order of increasing probability (decreasing significance).

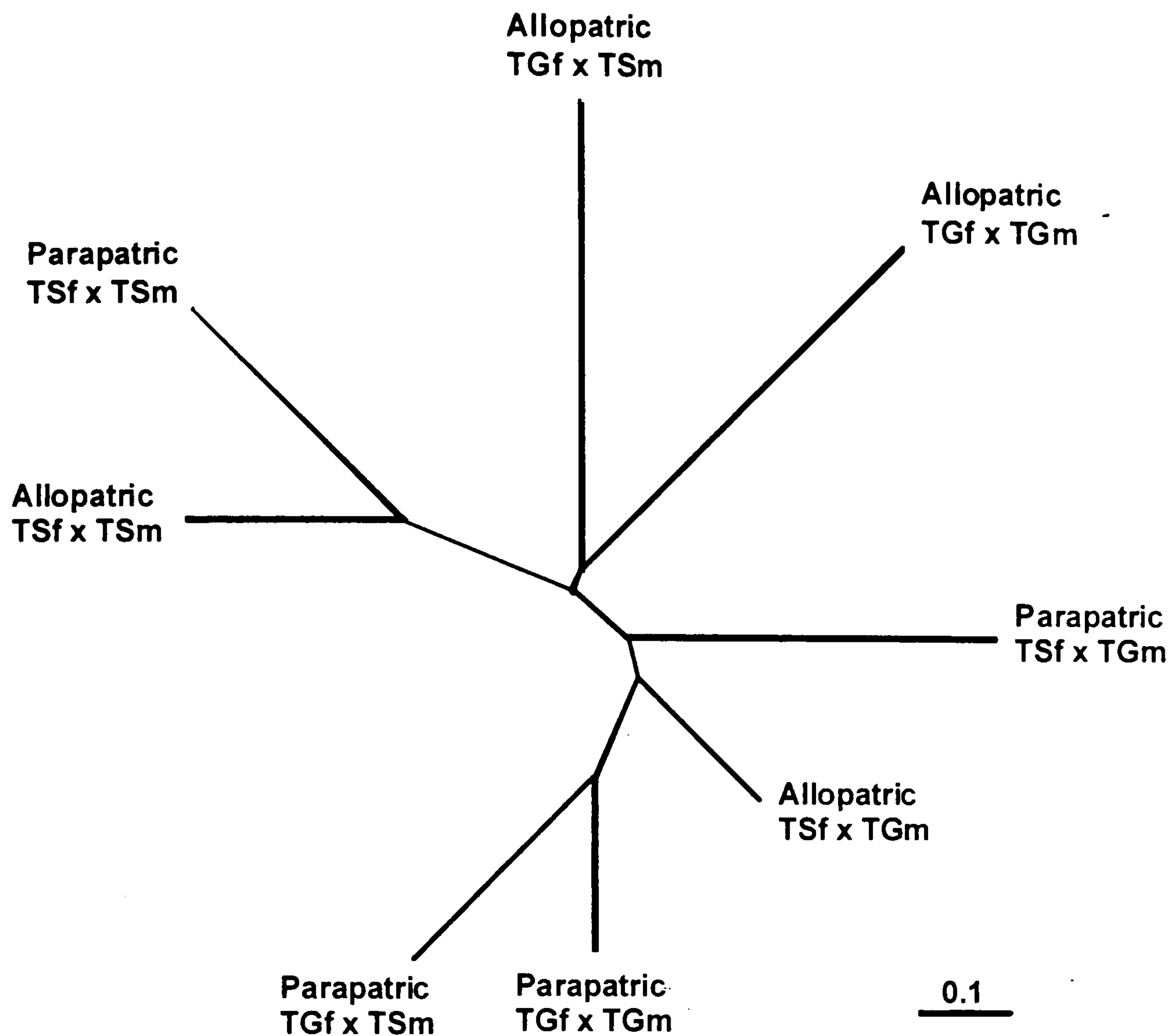


Figure 6.7. A radial dendrogram showing the relationships between courtship behaviours. Constructed using the neighbor-joining algorithm (Saitou and Nei, 1987), from 'distances' generated from the *P*-values (1 - *P*) resulting from pairwise χ^2 comparisons between the column totals of behavioural transition matrices.

behaviour (the column total) and the total frequencies of all other behaviours, in each table as input). All bar three behavioural elements were homogeneous for frequency of transition to the behaviour. The results for the three significantly heterogeneous behaviours are shown in Figure 6.8. MLSTA, MLCE and MLBW are all broadly Phase I elements, the chemoexploratory phase of courtship. Transitions to MLSTA were seen most frequently in TSf × TSm pairings, at quite high frequency in allopatric TGf × TSm pairings and then at lower frequency in all other pairings, with transitions to MLSTA occurring least frequently in TSf × TGm pairings. A higher number of transitions to MLSTA might reflect a

MLSTA: Overall G_H = 15.92; P < 0.05; tabulated χ^2, df = 7.

Parapatric	Allopatric	Allopatric	Parapatric	Parapatric	Parapatric	Allopatric	Parapatric	P_C
TSf × TSm	TSf × TSm	TGf × TSm	TGf × TGm	TGf × TGm	TGf × TSm	TSf × TGm	TSf × TGm	
0.0848	0.0774	0.0575	0.0437	0.0394	0.0380	0.0360	0.0357	
								*
								*

MLCE: Overall G_H = 28.50; P < 0.001; tabulated χ^2, df = 7.

Parapatric	Allopatric	Parapatric	Allopatric	Allopatric	Parapatric	Allopatric	Parapatric	P_C
TSf × TSm	TSf × TSm	TGf × TSm	TSf × TGm	TGf × TSm	TSf × TGm	TGf × TGm	TGf × TGm	
0.1816	0.1452	0.1308	0.1171	0.1081	0.1027	0.0848	0.0640	
								**
								***

MLBW: Overall G_H = 36.59; P < 0.001; tabulated χ^2, df = 7.

Allopatric	Parapatric	Allopatric	Parapatric	Parapatric	Parapatric	Allopatric	Allopatric	P_C
TGf × TSm	TSf × TGm	TSf × TGm	TGf × TSm	TGf × TGm	TSf × TSm	TSf × TSm	TGf × TGm	
0.0598	0.0313	0.0270	0.0211	0.0197	0.0121	0.0097	0.0026	
								*
								***

DECREASING FREQUENCY OF BEHAVIOUR

Figure 6.8. Behaviours significantly heterogeneous in total frequency between transition matrices. The results of a G_H test of heterogeneity are shown for the column totals of the transition matrices of each pairing combination. The three behaviours shown were significantly heterogeneous between matrices; all other behaviours were homogeneous. The frequency of each behaviour (the frequency with which that behaviour follows any other behaviour) is given below each pairing combination in decreasing frequency from left to right. Bars show homogenous groupings with the significance of the critical probability value P_C (the significance obtained when the next set is added; making the set heterogeneous) designated: * 0.05 > P > 0.01; ** 0.01 > P > 0.001; *** = P < 0.001.

generally more complex Phase I courtship for *T. saeva* males (a motionless period often marked a change from one behaviour to another). Transitions to MLCE showed considerable heterogeneity, being most frequent in TSf × TSm pairings and least frequent in TGf × TGm pairings, again perhaps reflecting a more complex Phase I courtship in *T. saeva* males. MLBW was also very heterogeneous, occurring most frequently in heterospecific crosses (most markedly in allopatric TGf × TSm pairings), and much less frequently in allopatric conspecific pairings. The marked increase in MLBW in heterospecific

pairings, and to some degree parapatric conspecific pairings, is interesting as this behaviour may well have a chemosensory role (this trend was also apparent in the sequence diagrams).

The individual dyads were then compared between matrices by using the transition probabilities (P_{ij}) (section 6.2.2) of one matrix to generate the expected values in the next matrix; multiplying the relevant row total by the transition probability from the matrix to be compared for the dyad in question (Baylis, 1976; Clark, 1994). The new row and total χ^2 were calculated using Yate's correction, as for the original matrices, and each dyad χ^2 was again calculated according to Clark's (1994) formula (equation 6.1). All possible comparisons were made (56 matrix comparisons $\times$ 306 dyads = 17,136). Only those dyads significant at $P < 0.001$ were noted. If a dyad showed observed frequencies lower than expected, then the logic was reversed to infer that the transition occurred more frequently in the matrix generating the expected. Hence a table was constructed showing transitions occurring more frequently than expected between matrices (see Table 6.14). This approach allows explicit comparison of the important transitions *between* the matrices and can pin-point differences in male and female behaviour. It is also more robust than the original analysis of significant dyads (above) as the expected values generated tend to be larger or zero (only transitions with observed values can generate expected values in the next matrix). There were no significant differences between any of the matrices in terms of overall χ^2 probabilities (56 comparisons; all $df = 272$, $P > 0.5$). This was to be expected given that many cells were incalculable (zero expected) and therefore contributed nothing to the χ^2. On examining Table 6.14 it can be seen that in general parapatric pairing combinations are distinguished from each other by far fewer transitions than allopatric pairings. This implies a marked and general trend towards courtship simplification in parapatry, or rather, an emphasis on a few transitions relative to others. By way of example, note that allopatric TSf $\times$ TSm pairings and allopatric

Table 6.14. Significant transitions between transition matrices.

	Allopatric TSf × TSm	Allopatric TGf × TGm	Allopatric TGf × TSm	Allopatric TSf × TGm	Parapatric TSf × TSm	Parapatric TGf × TGm	Parapatric TGf × TSm	Parapatric TSf × TGm
Allopatric								
TSf × TSm	-	BO, EA, GP, OB	BO, OB	BO		OB	BO, HP	HP
TGf × TGm	GK, HI	-	DH, GE, GK, HQ, OP, RH	HI	AC, HR	PH	BA, GK	HI,
TGf × TSm	BD, OH	BD, EA, HC, HD	-	BD, GH	BD, OH	BD, DB	AH, BD	BD, HC
TSf × TGm		GP		-			HP	HP
Parapatric								
TSf × TSm		EA, OB	BO, CB, OB	BO, GH, HI	-	OB		BC, HI
TGf × TGm	BP	OG	RH			-		
TGf × TSm	BP	BP, OB	OB		BP	OB	-	GO
TSf × TGm	GK	GP, HD	GK		GK		GK	-

Mean number of significant transitions		
Allopatric / Allopatric	Allopatric / Parapatric	Parapatric / Parapatric
4.17 (25/6)	2.44 (39/16)	1.33 (8/6)

Transition probabilities for individual dyads were used to generate expected frequencies in other matrices. All matrices were compared. The transitions (dyads) occurred more frequently in the pairing combination corresponding to the row in the table, for example: the transitions GK and HI occurred more frequently in allopatric TGf × TGm pairings than in allopatric TSf × TSm pairings. χ^2 analysis for each dyad: $df = 1$; $P < 0.001$. The mean number of significant transitions between the different types of comparison is given at the base of the table.

TGf × TGm pairings are distinguished by six transitions (two occurring more frequently in TGf × TGm pairings and four occurring more frequently in TSf × TSm pairings), whereas in parapatry, these pairing combinations were only separable by one transition. The mean number of differences between the different types of comparison (allopatric / allopatric, allopatric / parapatric, parapatric / parapatric) are tabulated at the base of Table 6.14. In interpreting the single-letter codes for each behavioural element the reader is referred to section 6.3.1 and Table 6.12. As it was possible to distinguish transitions from one behavioural

element to another, it was therefore possible to determine the distinguishing responses (those elements facilitated) and attribute these to the responding sex or courtship phase. Salient comparisons will be taken in turn:

ALLOPATRIC AND PARAPATRIC TSF × TSM PAIRINGS:

Allopatric and parapatric TSf × TSm pairings did not differ: no transitions were significantly more frequent in parapatry or allopatry.

ALLOPATRIC AND PARAPATRIC TGF × TGM PAIRINGS:

Allopatric and parapatric TGf × TGm pairings differed in frequency for two transitions. The transition PH: P. FEOM to H. MLDVA (a transition to a Phase II element), occurred more frequently in allopatric TGf × TGm pairings. The transition OG: O. FEW to G. FLT (to male/female but mediated by the female) occurred more frequently in parapatric TGf × TGm pairings.

ALLOPATRIC AND PARAPATRIC TGF × TSM PAIRINGS:

Allopatric and parapatric TGf × TSm pairings differed in three transitions, with AH: A. MLSTA to H. MLDVA and BD: B. MLCE to MLBW occurring more frequently in allopatric pairings (all transitions to Phase I elements). The transition OB: O. FEW to B. MLCE (a transition to a Phase I element), occurred more frequently in parapatry.

ALLOPATRIC AND PARAPATRIC TSF × TGM PAIRINGS:

The only difference between allopatric and parapatric TSf × TGm pairings was in the significance of the transtion HP: H. MLDVA to P. FEOM (to female) in allopatry.

TSF × TSM AND TGF × TGM PAIRINGS:

In *allopatry*, TSf × TSm and TGf × TGm pairings were distinguished by the frequencies of six transitions. The transitions, BO: B. MLCE to O. FEW (to female), EA: E. MLRND to A. MLSTA (to Phase I), GP: G. FLT to P. FEOM (to female), and OB: O. FEW to B. MLCE (to Phase I), were most frequent in TSf × TSm pairings. The transitions, GK: G. FLT to K. FTF (to male/female), and HI: H. MLDVA to I. MLWA (to Phase II), were most frequent in TGf × TGm pairings. In *parapatry*, TSf × TSm and TGf × TGm pairings were distinguished by the frequencies of only one transition: the transition OB: O. FEW to B. MLCE (to Phase I) occurred more frequently in the TSf × TSm pairings. OB was characteristic of both allopatric and parapatric TSf × TSm pairings.

TO SUMMARIZE:

There was no apparent behavioural difference between allopatric and parapatric TSf × TSm pairings. There were two significant differences between allopatric and parapatric TGf × TGm pairings. These resulted from one significant transition in allopatry, which was attributable to Phase II courtship, and one significant difference in parapatry which was attributable to female behaviour (regarding male/female as dependent on a female response). Allopatric and parapatric TGf × TSm pairings differed in three transitions, two in allopatry and both attributable to Phase I courtship, and one Phase I response in parapatry. TSf × TGm pairings only differed in one transition to a female element in allopatry. In general there were fewer significant transitions in parapatry and these were all atributable to Phase I courtship or female behaviours. Allopatric TSf × TSm and TGf × TGm pairings were distinguished by a range of transitions largely resulting from a diversity of Phase I and female responses in the TSf × TSm pairings. The number of significant transitions in parapatry was much reduced.

6.3.4 Duration of Courtship and Copulatory Parameters

Values for the recorded temporal parameters of courtship and copulation are summarized in Table 6.15. The total duration of courtship ranged from 26 s to 52 min for courtships ending in attempted copulation. Total courtship time "is as much an indication of the receptivity of the female as it is a measure of the persistence of the male" (Stratton, 1997). Given this and the huge variation in this parameter it was unlikely to indicate any differences between the pairing combinations, and indeed there were no significant differences (Kruskal-Wallis Test: $H = 5.71$; $df = 7$; $P > 0.05$, n.s.). Latency to chemoexploratory behaviour (initiation of Phase I courtship) varied from 0 s (immediate exploration) to 23 min 45 s, with both long and short latency being observed in all pairing combinations and those pairings not ending in attempted copulation. Latency to chemoexploratory behaviour was, however, generally short, being less than one minute in 89% of pairings. There were no significant differences in latency to chemoexploratory behaviour between the pairing combinations or those courtships not ending in attempted copulation (Kruskal-Wallis Test: $H = 7.23$; $df = 8$; $P > 0.05$, n.s.). These results agree with those of Stratton (1997) for courtship in *Schizocosa* wolf spiders, in that latency to chemoexploration is not a good indicator of the pattern of courtship to follow or the chances of attempted copulation. Similar results were obtained for the average duration of chemoexploratory bouts (Kruskal-Wallis Test: $H = 5.05$; $df = 8$; $P > 0.05$, n.s.), and also for the total duration of Phase I courtship (Kruskal-Wallis Test: $H = 8.59$; $df = 8$; $P > 0.05$, n.s.). Additionally, there were no differences in the duration of Phase II courtship between pairing combinations (Kruskal-Wallis Test: $H = 1.82$; $df = 7$; $P > 0.05$, n.s.), nor were there any differences in the relative durations of Phase I and Phase II courtships (Kruskal-Wallis Test: $H = 2.06$; $df = 7$; $P > 0.05$, n.s.). Hence, the duration of the broad phases of pre-copulatory behaviour provide no indication of differences between the two species under different pairing regimes. This may be due in part to the fact that these parameters reflect both

Table 6.15. Mean duration of courtship and copulatory parameters for different pairing combinations.

Pairings	Duration court	Latency MLCE	Duration MLCE bouts	Duration Phase I	Duration Phase II	Phase II / Phase I	Insertion duration
	Mean (s) ± S.D.	Mean (s) ± S.D.	Mean (s) ± S.D.	Mean (s) ± S.D.	Mean (s) ± S.D.	Mean ± S.D.	Mean (s) ± S.D.
Allopatric							
TSf × TSm (n = 10)	836.60 ± 418.37	11.90 ± 13.22	175.75 ± 163.40	509.80 ± 472.52	326.80 ± 335.47	2.53 ± 5.14	26.05 ± 14.87
TGf × TGm (n = 12)	856.25 ± 806.66	122.92 ± 410.12	81.53 ± 71.47	311.08 ± 244.86	545.17 ± 717.01	2.48 ± 2.45	45.27 ± 12.07
TGf × TSm (n = 12)	810.67 ± 663.55	27.83 ± 74.70	83.73 ± 48.11*	340.00 ± 453.55	470.67 ± 401.28	3.49 ± 3.86	3.74 ± 7.21
TSf × TGm (n = 9)	539.44 ± 393.72	43.44 ± 114.46	128.33 ± 104.74	147.94 ± 154.27	368.44 ± 361.14	2.75 ± 2.18	11.48 ± 3.33
Parapatric							
TSf × TSm (n = 12)	712.00 ± 505.95	91.17 ± 150.30	101.92 ± 73.30	351.92 ± 293.46	360.08 ± 520.07	6.49 ± 17.01	31.67 ± 16.81
TGf × TGm (n = 8)	579.25 ± 555.66	46.75 ± 104.93	111.22 ± 108.64	224.75 ± 163.18	354.50 ± 448.20	1.98 ± 2.32	33.09 ± 7.96
TGf × TSm (n = 8)	596.13 ± 452.48	38.37 ± 90.05	73.50 ± 63.15	219.50 ± 217.81	376.63 ± 520.95	6.00 ± 9.23	7.12 ± 8.93
TSf × TGm (n = 10)	541.40 ± 528.77	1.90 ± 6.01	83.05 ± 74.16	168.20 ± 123.96	373.20 ± 476.09	2.27 ± 2.44	10.73 ± 7.71
Courtships not ending in copulation	-	26.36 ± 60.06 (n = 14)	88.73 ± 63.68 (n = 14)	422.33 ± 523.87 (n = 12)	-	-	-

The mean duration of each parameter is given for for each pairing combination with the standard deviation. Values for courtships not ending in attempted copulation are given in the bottom row (where calculable) along with sample size. All other sample sizes are as in the first column except * n = 10.

female receptivity and male persistence, as alluded to above, and therefore are of limited value.

The final column in Table 6.15 shows the mean duration of palpal insertions for each pairing combination. For each individual pairing a large number of palpal insertions were timed ($n > 20$) from moment of insertion to moment of removal of the palp from the epigyne (corresponding to one haematodochal inflation). The

average value was taken as the value for that pairing and used to generate the means shown in Table 6.15. The trials revealed that the heterospecific copulations suffered from two forms of mechanical (morphological) difficulty. 1) *Failure to engage* the palp with the epigyne. The palpal tibial apophyses must lock with the posterior facing apophyses on the female epigyne. In all pairings a male may occasionally fail to 'find' the epigyne (if the angle of approach is wrong). However, it was clear in heterospecific pairings, and particularly in TGf × TSm pairings (although this has not been explicitly characterized), that the male repeatedly and commonly failed to locate the epigyne - this could be referred to as 'apophysis mismatch'. 2) *Palpal slippage*: the palpal and epigynal apophyses interdigitate but the conductor fails to enter the copulatory duct or slips away in less that a few seconds (causing the palp to leave the epigyne). It appeared from visual inspection of copulations that heterospecific males had to expend more effort in copulation (adopting more severe angles relative to the female and 'pushing' her upwards with the palp). In recording insertion duration, failures to engage were scored as occupying zero seconds. Slippage occurring in less than one second also scored zero. In each palpal bout (set of attempted copulations using one or other palp and not broken by either of the two sexes moving away) consecutive failures to engage, or slippages, were, for recording consistency, only scored once (it is impossible to define whether the male shaking the palp around the epigyne for a continuous period constitutes one continuous attempt or many repeated attempts to engage). It is therefore likely that these mechanical limitations were 'under-recorded' and that the average insertion duration for heterospecific pairings was actually lower in reality (which would only increase the significance of the results below).

Insertion duration was highly heterogeneous between the pairing combinations (Kruskal-Wallis Test: $H = 51.54$; $df = 7$; $P < 0.001$). There were no significant differences between the conspecific pairings (Kruskal-Wallis Test: $H = 7.44$; $df = 3$; $P > 0.05$, n.s.). Pooling the conspecific values and comparing these to the

pooled heterospecific values confirms, as expected from the mechanical limitation discussed above, that the mean insertion duration was shorter for heterospecific pairings (Mann-Whitney *U*-test: 1-tailed; $P < 0.001$). Closer examination of the heterospecific pairings also reveals heterogeneity (Kruskal-Wallis Test: $H = 9.48$; $df = 3$; $0.05 > P > 0.01$). This heterogeneity was due to shorter insertion durations in TGf × TSm pairings than in TSf × TGm pairings (Mann-Whitney *U*-test: 1-tailed; $0.01 > P > 0.001$), presumably due to the influence from failures to engage the palps. (There were no significant differences between allopatric and parapatric TGf × TSm pairings or between allopatric and parapatric TSf × TGm pairings (in both cases Mann-Whitney *U*-test: 2-tailed; $P > 0.05$), and individually, TGf × TSm pairings and TSf × TGm pairings were very significantly shorter than conspecific pairings (in both cases Mann-Whitney *U*-test: 1-tailed; $P < 0.001$)). The significance values in the above tests were all marked and therefore, despite the multiple testing, it can be stated that heterospecific pairings suffer extreme mechanical limitations on copulation which are even more marked for TGf × TSm pairings than TSf × TGm pairings.

In order to examine further the relationship between insertion duration and morphology, the average insertion duration for each pairing was regressed against a 'genital ratio'. The ratio was the maximum length of the epigyne divided by the maximum combined tegulum and conductor length (average for both palps) (see Chapter 3). These measurements have been shown to be highly correlated with body size (prosoma length) (see Chapter 3), but have been used in preference to body size because genital-genital interactions are of direct relevance here and because there may be differences in allometry between body size and genitalic measurements between allopatric and parapatric populations (see Chapter 3 and discussion in Chapter 7). Allopatric and parapatric trials were pooled to increase sample sizes and regressions were performed for all four pairing combinations. Figure 6.9 shows the results for the conspecific pairings. The regression slope was

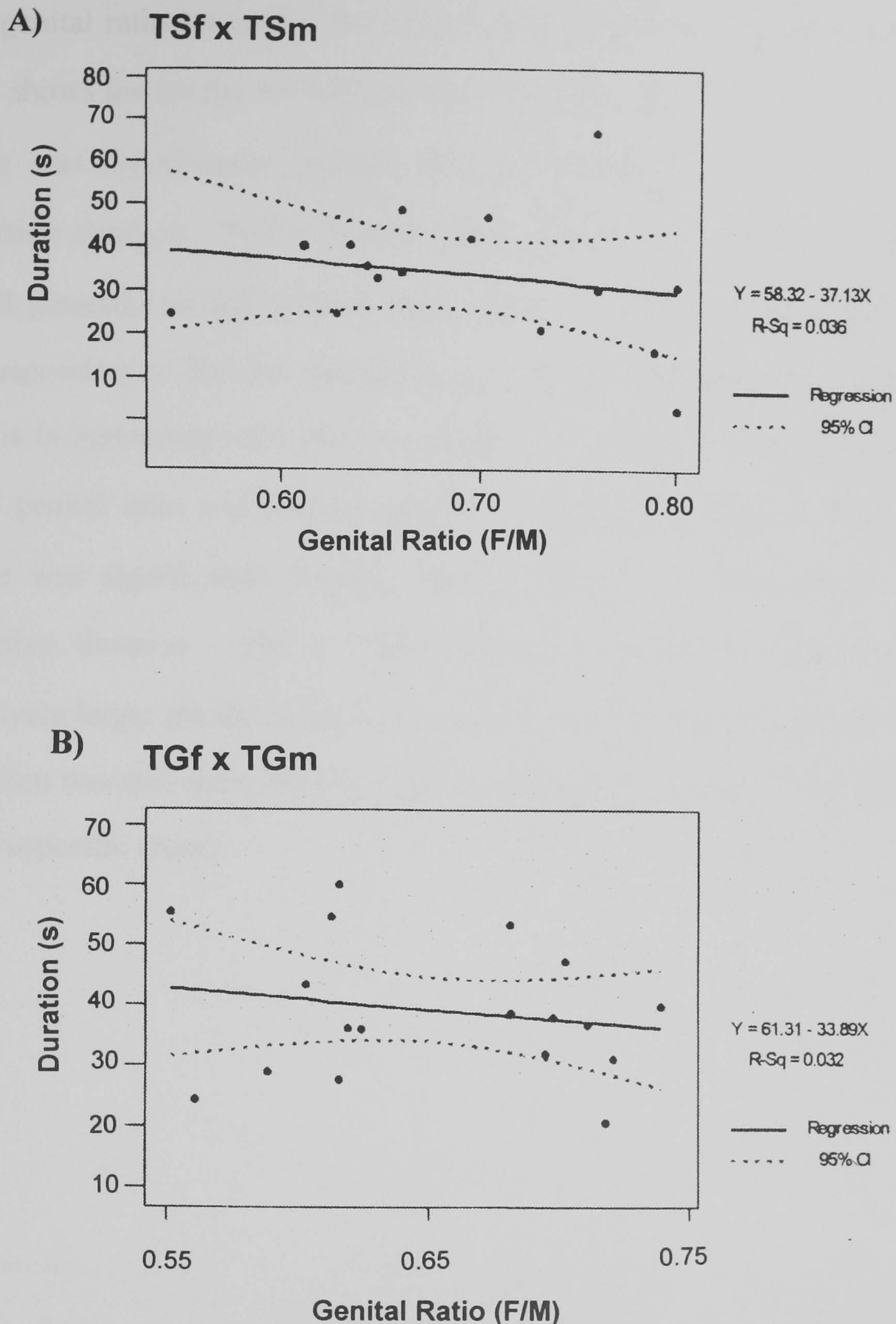


Figure 6.9. Mean palpal insertion duration as a function of female/male genital size: conspecific pairings. The mean duration of palpal insertions for each pairing was regressed against 'genital ratio': the maximum epigyne length of the female divided by the maximum combined conductor + tegulum length of the male. Larger values of 'genital ratio' approximate to smaller males relative to females. **A:** TSf × TSm pairings. **B:** TGf × TGm pairings. Neither slope was significantly different from zero (A: $F = 0.52$, ($df = 1, 14$), $P = 0.48$, n.s.; B: $F = 0.56$, ($df = 1, 17$), $P = 0.47$, n.s.). (R-Sq = R^2). Adjusted $R^2 = 0.00$ in both cases. Genital Ratio accounted for none of the variation in mean insertion duration time in conspecific crosses. (Pairings with mean insertion durations < 1 s were excluded: such low values indicate a persistent failure to engage the palpal apophyses with the epigynal apophyses presumably due to other morphological factors).

not significantly different from zero for either TSf × TSm or TGf × TGm pairings and genital ratio accounted for none of the variation in insertion duration. Figure 6.10 shows the results for the heterospecific pairings. For TGf × TSm pairings the slope was significantly positive and accounted for 30% of the variation in insertion duration - TGf × TSm pairings in which *T. saeva* males had relatively small genitalia tended to have more successful palpal intromissions. The point corresponding to the one successful hybrid cross producing progeny is indicated and is in agreement with this presumption - the point corresponds to a relatively high genital ratio and average insertion duration. For TSf × TGm pairings the slope was significantly negative and accounted for 20% of the variation in insertion duration - TSf × TGm pairings in which *T. gigantea* males were relatively larger tended to have more successful palpal intromissions (although the duration was still short relative to the conspecific pairings, even for the successful heterospecific cross).

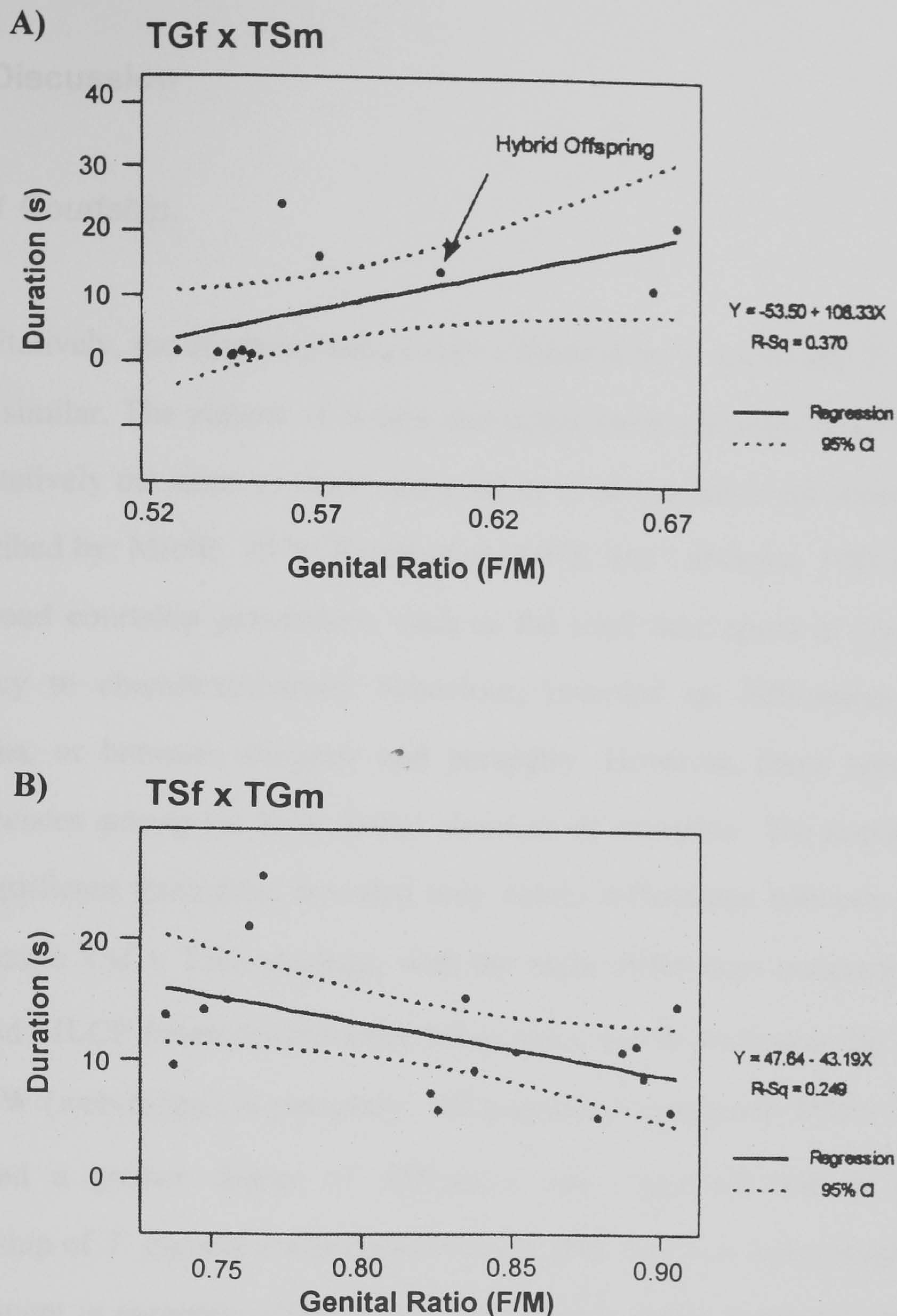


Figure 6.10. Mean palpal insertion duration as a function of female/male genital size: heterospecific pairings. The mean duration of palpal insertions for each pairing was regressed against 'genital ratio': the maximum epigyne length of the female divided by the maximum combined conductor + tegulum length of the male. Larger values of 'genital ratio' approximate to smaller males relative to females. **A:** TGf × TSm pairings. Slope significantly positive (F = 5.29, (df = 1, 9), P = 0.047), adjusted R^2 = 0.30. The one pairing producing offspring is marked. **B:** TSf × TGm pairings. Slope significantly negative (F = 4.96, (df = 1, 15), P = 0.042), adjusted R^2 = 0.20. (R-Sq = R^2). Genital Ratio accounted for 20 - 30% of the variation in mean insertion duration time in heterospecific crosses, with small males more successful in TGf × TSm crosses and large males more successful in TSf × TGm crosses. (Pairings with mean insertion durations < 1 s were excluded: such low values indicate a persistent failure to engage the palpal apophyses with the epigynal apophyses presumably due to other morphological factors).

6.4 Discussion

6.4.1 Courtship.

Qualitatively, the courtship behaviours exhibited by *T. saeva* and *T. gigantea* were very similar. The pattern of events and behavioural elements described were also qualitatively the same as those observed in *T. atrica* (personal observation, and as described by: Mielle, 1978; Krafft *et al.*, 1978; and Leborgne, 1989). The duration of broad courtship parameters, such as the total time spent in courtship, or the latency to chemoexploratory behaviour, revealed no differences between the species, or between allopatry and parapatry. However, there were quantitative differences among the behavioural elements of courtship. The sequence diagrams of significant transitions revealed only subtle differences between allopatric and parapatric TSf × TSm pairings, with the main differences concerning transitions around MLCE (chemoexploratory behaviour), and in particular the importance of MLBW (web-biting) in parapatry. Allopatric and parapatric TGf × TGm pairings showed a greater degree of difference and suggested that in parapatry the courtship of *T. gigantea* was simplified (in other words a few elements were more prominent in parapatry relative to other elements, when compared with allopatry). The diagrams suggested that there were changes in both Phase I, Phase II and female responses. Allopatric and parapatric TSf × TGm sequence diagrams showed few obvious differences, and the equivalent diagrams for TGf × TSm pairings suggested reduced receptivity in the females and a little simplification in male courtship.

Comparisons of the column totals from the transition matrices, by pairwise χ^2 analyses and the resulting dendrogram, were very informative. Examination of the dendrogram (Figure 6.7) clearly illustrates the general conclusions, which

supported and extended the conclusions from the sequence diagrams. The dendrogram reflected the fact that allopatric and parapatric TSf × TSm pairings did not differ very much in behaviour, but that allopatric and parapatric TGf × TGm pairings differed a great deal (this was very marked on the dendrogram, which employed all the data, but was not significant in the individual pairwise comparison). Allopatric and parapatric comparisons of the heterospecific pairings were also in agreement with the sequence diagrams, with TSf × TGm pairings showing no significant differences (and grouping together on the dendrogram) and TGf × TSm pairings showing very significant differences (and a correspondingly large separation on the dendrogram). The fact that TSf × TSm pairings differed little whereas TGf × TGm pairings differed greatly in allopatry and parapatry, together with the fact that TSf × TGm pairings differed little but TGf × TSm pairings differed a great deal, suggests a strong influence of *T. gigantea* female discrimination in parapatry. This analysis also showed that allopatric TGf × TGm and allopatric TSf × TSm behaviours were significantly different, and that the behaviours of the two species were *very* different in parapatry.

The detailed comparisons of individual transitions (dyads) between matrices, using the transition probabilities from one matrix to generate the expected values in the next, powerfully corroborated these results. Furthermore they illustrated that the stronger differences between the two species in parapatry resulted from the relative emphasis of just a *few* behaviours compared with the allopatric differences. Generally, the transitions that distinguished the pairing combinations (particularly in parapatry) were linked to Phase I and female behaviours, suggesting that female responses to male behaviours shown early in courtship were the most important. This fact was also supported from the G_H -tests on the individual column totals, which revealed three Phase I elements to be heterogeneous. Perhaps the most interesting of these was MLBW. Web-biting did not appear to correspond to the male cutting the web in order to gain easier access,

or to reduce the web size. Cutting-away the female web is a behaviour employed by males of some spider families and is exemplified by the linyphiids, *Lepthyphantes leprosus* and *Linyphia triangularis*, in which the male may reduce the female web by more than 90%, presumably to confine her location (Platnick, 1971). Web-biting behaviour was observed most frequently in heterospecific pairings, and next most frequently in parapatric conspecific pairings. There could be a number of explanations for this observation. The male could be displaying an extreme signal to the female in response to receiving inappropriate responses from her (the release of the web often resulted in a very audible 'snap'), or perhaps this constitutes a displacement behaviour. More likely, web-biting serves a chemosensory role. As previously mentioned, the male spider possesses numerous chemosensory sensilla mostly located on the tarsi of the legs and palps (in addition to the tarsal organ which has been shown to have a chemosensory function in detecting air-borne female odours in *Cupiennius salei* (Tietjen and Rovner, 1982)). Males possess an unusually large number of these sensilla on the palpal tibia as a secondary sexual characteristic (Foelix, 1996; Tietjen and Rovner, 1982). These hairs no doubt play an important role in the male location and chemoexploration of the female web. Web-biting could reflect an additional way to sample the web-pheromone through the chemosensory sensilla of the maxillae or even through cells in the pharynx which have been implicated in the sensation of taste (Foelix, 1996). But despite all this, more or less all males go on to court females of either species, so why should they take in this extra information? Possibly there are differences in silk pheremones between the species but not enough to stop the male courting. However, the male may need to taste the slightly different pheremone more (to exceed some excitatory threshold) before continuing with courtship. Alternatively, or perhaps in combination, the male spider could be daubing a male pheromone on the web from the gnathocoxal glands, which may aim to suppress female aggression (Section 6.1.2 and, Legendre and Lopez, 1974; Krafft, 1982). The cat-stretch-like motion frequently employed in this behaviour could facilitate the silk passing through

the maxillae and over these glands. In connection with this, the observation that the male spread his spinnerets widely and often laid thick bands of silk during chemoexploration, corresponds to the observation, and male pheromone implication, that male lycosids secrete greater amounts of silk than usual during chemoexploration and trail following (Tietjen and Rovner, 1982).

The transition analysis had limitations. Only changes from one behavioural element to a different behavioural element were recorded. Therefore nothing can be said about the effect of the repetition of individual elements or their overall duration. Also, with the exception of the analysis employing transition frequencies in one matrix to generate expected values in the next matrix, the expected values were calculated from the column sums (frequencies) of the observed data; therefore any large facilitatory or inhibitory effect within any given row would tend to 'push' other cells in the same row toward a significant deviation in the opposite direction (Baylis, 1976). However, given that it was probably changes in the relative frequency of transitions that was most important, this seems to be a minor detraction. A more serious limitation was that only pairings resulting in attempted copulation could be employed. This would not be a problem if there were no significant differences in copulation frequency between the different pairing combinations (Table 6.1; page 273). In fact, *T. saeva* females responded significantly negatively towards *T. gigantea* males compared to conspecific males. Also there was a trend towards *T. saeva* females being more discriminatory in parapatry (although this was not significant, Table 6.1). In addition, the only overt and sustained displays of aggression recorded were from *T. saeva* females towards *T. gigantea* males (resulting in the male being removed in three cases). Aggression also showed a non-significant trend towards higher frequency in parapatry. These pairings were excluded from the transition analyses which were therefore biased towards the more tolerant females. Of course, this bias only applies to *T. saeva* females, and not to *T. saeva* males, therefore the lack of divergence between allopatric and parapatric TSf × TSm pairings may reflect a real lack of

divergence in male behaviour, while not taking account of a real divergence in female behaviour.

It was also observed that males of both species, almost without fail, will initiate courtship when placed upon the web of a virgin female of either species (see above). This suggests, as in the case of the *Schizocosa* species discussed in Section 6.1.2 (Stratton and Uetz, 1981), that the female web presents a chemosensory trigger, but lack of specificity. Male courtship could also have been triggered by an unrecorded female stimulus. This seems unlikely, but should be tested by placing males on conspecific and heterospecific female webs without females. It has also been suggested that male eagerness to court in similar experiments could be a laboratory artefact. Prolonged stimulus deprivation of the test animals could alter response thresholds such that courtship can be triggered by subnormal stimuli (Tietjen and Rovner, 1982). This can only be circumvented by using freshly caught, wild males, for each trial; something that would require very careful planning. It would also be valuable to try placing males on the webs of virgin female *T. atrica*, *T. parietina*, *T. domestica* etc. to see how general the effect of the female pheremone is in intitiating courtship.

6.4.2 Mechanical Isolation

The mating trials showed that *T. saeva* and *T. gigantea* experience a large degree of mechanical isolation, as exemplified by the measure of palpal insertion duration. Two routes to mechanical isolation have been identified: failure to engage the palp with the epigyne and slippage of the conductor (and consequently the palp) from the copulatory duct. The former effect appeared to make mechanical isolation more severe in TGf × TSm pairings. Observation also suggested that parapatric TSf × TGm pairings may also experience a degree of 'failure to engage' not seen in allopatry, but the data on insertion

duration did not support this. It also seemed likely that this mechanical barrier was sufficient in most cases to prevent fertilization (in other words that the embolus fails to reach far enough into the copulatory duct to deposit sperm in the spermatotheca). Support for this, although from only one sample, comes from the fact that in the single reproductively successful heterspecific pairing, there was no difference in mean egg viability compared with conspecific crosses. Nor did the eggs that had failed to develop, in any cross, appear to have aborted; they were simply unfertilized (although one cannot rule out other barriers to fertilization or effects of the female tract).

The results of comparing palpal insertion duration with 'genital ratio' fit with expectation. Both maximum epigyne length and combined conductor and tegulum length are larger (relative to body size) in *T. saeva* than in *T. gigantea* (see Chapter 3). Consequently, if the male palp is to 'span the gap' between the epigynal apophysis and the copulatory duct, without over- or under-reaching, then one might expect larger *T. gigantea* males to 'do better' with smaller *T. saeva* females and smaller *T. saeva* males to 'do better' with larger *T. gigantea* females. This seems to be the case. Additionally *T. saeva* males with relatively smaller conductors might be better suited to negotiating the more oblique copulatory opening of *T. gigantea* females. On a mechanical basis one might also expect the average genital ratio between conspecific partners of either species to be approximately the same (optimal), with similar insertion durations in both species; variation being due largely to other factors (morphological, behavioural, or physiological). The data presented above support this view. Genital ratio is just an index of relative genitalic size and is a gross simplification of the complex morphological variation present. However it provides a rare opportunity to predict a (probably) simple ontogenetically controlled route on which selection might act to increases mechanical isolation - make the genitalia of *T. saeva* larger and/or *T. gigantea* smaller. This hypothesis will be discussed, in Chapter 7 in terms of the morphological data (Chapter 3).

6.4.3 Character Displacement/Reinforcement?

The degree of mechanical isolation between *T. saeva* and *T. gigantea* provides a clear disadvantage to courting/accepting and attempting to mate with the wrong species. Whether the mechanical isolation evolved in allopatry or in an historical secondary contact under selection due to genetic incompatibilities, hybrid sterility, hybrid breakdown etc., cannot be known. Without further breeding and behavioural experiments to generate more F_1 hybrids, and F_2 and backcross hybrid generations, it is difficult to say very much about hybrid disadvantage. The evidence presented in the other chapters of this thesis will be employed in Chapter 7 to discuss whether the apparent parapatric divergences observed here reflect character displacement *or* reinforcement.

The transition analyses have revealed that, although the behaviours of *T. saeva* and *T. gigantea* were qualitatively similar, they differed in detail and differed to a much greater degree in the area of overlap. Further, these differences were largely due to changes in the courtship of male *T. gigantea* and the receptivity of female *T. gigantea*. There was only weak evidence for changes in the behaviour of *T. saeva* from the transition analyses. However, analyses of copulation frequency showed that *T. saeva* females discriminated significantly against *T. gigantea* males compared to conspecific males. There was a trend towards greater discrimination of *T. saeva* females in parapatry with a concurrent increase in aggression towards heterospecifics. The pattern of divergence seen in the data appears to fulfil the criteria set out at the end of Section 6.1.4 and is *suggestive* of character displacement or reinforcement; it is certainly the qualitative opposite of the pattern to be expected if these processes were not occurring (whereby one might expect the behaviours of the two species to be more similar in parapatry as a result of hybridization and introgression and more divergent in allopatry). Importantly, the evidence suggests that there is not just divergence in male traits, but that females are more discriminating in parapatry. The data also fulfilled the prediction that

traits modified in parapatry to ensure species recognition should not necessarily be those that distinguish the species in allopatry.

The analyses presented here were very cursory. As outlined above, there is a great need to perform more such experiments, not only to produce more hybrids but to add to the data so far collected and to examine the detail of the important discriminatory traits. For instance, Phase I elements were important (as expected - species recognition should occur early in courtship) and there is a clear need to examine the vibrations produced by the abdominal taps observed during chemoexploration. In similar species the frequency of the abdominal vibration has been shown to be species-specific (Boulanger *et al.*, 1986; Krafft, 1978; 1982; Krafft *et al.*, 1978; Leborgne, 1984; 1989, Leborgne and Krafft, 1979; Leborgne *et al.* 1980; Mielle, 1978). This should be investigated in parapatric and allopatric pairings together with the structure of the other vibrations such as those produced by palpal drumming. Characterization of the female sex pheromone, and investigation of any that may be produced by the male, would be invaluable.

7 General Discussion

7.1 The *Tegenaria atrica* group

The mtDNA phylogeny of *Tegenaria*, although based upon few species, supported the traditional systematic placement of *T. atrica*, *T. saeva* and *T. gigantea* in a group of closely-related sister species: the *T. atrica* group. The genus *Tegenaria* appears to contain marked phylogenetic divisions, with maximum CO1 nucleotide divergences of 20-25% for the species studied, representing a divergence time of around 10 Ma. Within the *T. atrica* group, the estimated divergence time for *T. saeva* and *T. gigantea* was ca. 1.4 Ma, and the estimated time for the splitting of the ancestor of these two species from *T. atrica* was ca. 2.5 Ma. Although these estimates were based on relatively few individuals, and despite the cautionary points associated with molecular 'clocks' outlined in Chapter 5, these estimates should not be dismissed. Once the errors associated with molecular 'clocks' have been acknowledged the information from rate estimates may still be informative (Lunt *et al.*, 1998). Indeed these estimates of divergence time corresponded remarkably well with the Northern European Quaternary climatic record, placing the divergence of *T. saeva* and *T. gigantea* around the period of the second major glaciation (Praetiglian), and the divergence of the ancestor of these two species from *T. atrica* during the first major glaciation (Eburonian) (Chapter 5). This, and the current European distribution pattern of members of the *T. atrica* group, suggest that the speciational and distributional history of these species has been moulded by the glaciation events of the Quaternary period, with repeated retreats into southern refugia. This would not be surprising considering the array of organisms in which patterns of postglacial range expansion from southern European or eastern European/Asian refugia have been described. Examples of these include: the brown bear *Ursus arctos* (Taberlet and Bouvet, 1994); the woodmouse *Apodemus sylvaticus* (Michaux *et al.*, 1996); the bank vole *Clethrionomys glareolus*, pigmy shrew *Sorex minutus* and common shrew

Sorex araneus (Bilton *et al.*, 1998); and the grasshopper *Chorthippus parallelus* (Hewitt 1990, 1996), among others. (For reviews on glaciations and refugia see for example: Hewitt, 1996; and also Haffer, 1982; Turner, 1982). It is conceivable that *T. saeva* and/or *T. gigantea* colonized Britain (but not Ireland) without human intervention, before the landbridges flooded, following expansion from southern European (Iberian) refugia (see Chapter 2). However, two factors suggest that the current distributions of members of the *T. atrica* group probably result largely from transport by people. First, these species are known to be readily transported by people (for example the recent colonizations of Iceland, Ireland and America - see Chapter 2). Second, recent studies suggest that the post-glacial colonization of Europe by many species was from eastern or northern Balkan refugia, as opposed to southern Mediterranean refugia (Bilton *et al.*, 1998). It would be valuable to know the European distributions of the members of the *T. atrica* group in more detail, and to attempt to explore their phylogeographic history through mtDNA sequencing. There is also a clear need to investigate the phylogeny of the genus *Tegenaria* in more detail, by including more species in the analysis, and in particular by including the putative fourth member of the *T. atrica* group, *T. aliquoi*, for which very little information exists. If this currently isolated Sicilian endemic *really* belongs in the *T. atrica* group then its phylogenetic placement and estimated divergence times, relative to the other species, would inform the understanding of the phylogeographic history of this group. Any future analyses of mtDNA will of course have to take account of the presence of introgressed mtDNA molecules (Chapter 5 and below). Further, the identification of a putative pseudogene of CO1 in *T. parietina* emphasizes the ever present need for caution in interpreting sequence data.

7.2 Hybridization and Introgression in the T. atrica *group*

The evidence that the members of the *T. atrica* group can, and do, hybridize in the wild is overwhelming. The observations of Merrett (1980), Oxford and

Smith (1987) and Oxford and Plowman (1991) on the morphology of the male palps of wild-caught *T. gigantea* and *T. saeva*, together with the (albeit preliminary) report on laboratory crosses from Kennett and Dalingwater (1986), left little doubt that *T. gigantea* and *T. saeva* form hybrids. There have, however, been no reports in the literature of morphological intermediates between *T. gigantea* and *T. atrica*. This may be because these species do not hybridize in the wild, or simply because intermediates have not been found, or have been over-looked. It should perhaps also be remembered that much of the interest in the *T. atrica* group has come from British arachnologists: *T. atrica* rarely occurs in Britain. However, it does seem likely that *T. saeva* and *T. atrica* can and do hybridize. Although the interactions between these two species have not been specifically examined here, observations by Locket (1975), and by Barrientos and Ribera (1988), on Spanish material, examination of Icelandic specimens (where both species are established in ports) (Agnarsson, 1996; and G. S. Oxford, pers. comm.), and personal observations on material from Dublin, Eire (coll. J. O'Connor), and on material from Nancy, France (coll. R. Leborgne), suggest strongly that hybridization occurs.

The detailed morphometrical analyses presented here have confirmed the presence of individuals that are intermediate, to varying degrees, between *T. saeva* and *T. gigantea*. These patterns were suggestive of not only interspecific hybridization, but of extensive introgression. The ability to form F_1 hybrids was confirmed by the successful crossing of a female *T. gigantea* with a male *T. saeva* and the successful rearing of a small number of the progeny (although only one successful crossing was performed, and this was using parents from the area of parapatry in southern England - who were therefore of unknown ancestry). The males from this cross were intermediate in morphology as recorded by the discriminant analysis and thus supported the morphometric conclusions that intermediacy did reflect hybrid ancestry. The production of F_1 females was important because all previous morphological surveys failed to record intermediate females. However, the hybrid females were not readily distinguished by the discriminant analysis. (This is not meant to imply that

hybrid females resemble one parental species more than the other - they may indeed be intermediate but the characters used in the discriminant analysis are not sensitive enough to show a clear pattern. An examination of internal characters of the epigyne might lead to a more robust separation of the species and consequent placement of the hybrids). This highlights the general difficulties associated with the identification of females and suggests that hybrid females, although present in nature, are generally overlooked. Hybridization in the wild was confirmed by the pattern of allozyme allele frequencies (for instance, the 'over-spill' between *T. saeva* and *T. gigantea* in the area of parapatry - zones 3 and 4 - in southern England). Further confirmation, and proof that female hybrids *must* exist (at least from crosses of *T. gigantea* females and *T. saeva* males), was provided by the evidence for substantial introgression of *T. gigantea* mtDNA into *T. saeva* populations.

The patterns of hybridization and introgression will now be discussed, starting with southern England. In southern England the two species exhibited quite distinct ranges, with *T. saeva* occupying the west of the country and *T. gigantea* the east; with the two species forming a parapatric boundary in eastern Dorset. Most specimens from southern England were easily identified as *T. saeva* or *T. gigantea* from the visual assessment of genital morphology, suggesting little hybridization. Only two males, and four females, were not readily assigned to species. The discriminant function on males was more informative than that on females and revealed that a number of individuals originally identified as *T. saeva*, in parapatry, were in fact intermediate. The distribution of the parapatric *T. saeva* was skewed towards *T. gigantea* with a tail that was indicative of introgression. No such evidence was apparent for *T. gigantea*, indicating that hybridization and introgression were asymmetrical in southern England. The mtDNA data also showed that introgression was asymmetrical in that *T. gigantea* haplotypes were frequently revealed in individuals that were identified as *T. saeva* on morphological grounds. Such introgression was detected far away from the contact area, with the most distant example being a specimen of *T. saeva* from Trewen, Cornwall -

approximately 175 km west of the contact boundary. There was no evidence for introgression of *T. saeva* mtDNA molecules into *T. gigantea* populations. Because mtDNA is maternally inherited, this implies that female F_1 hybrids between *T. gigantea* females and *T. saeva* males are able to backcross successfully with *T. saeva* males, whereas female F_1 hybrids between *T. saeva* females and *T. gigantea* males (if they are ever produced) are unable to backcross successfully with *T. gigantea* males. However, this of course provides no indication of how successful backcross matings may be between female F_1 hybrids of *T. gigantea* female and *T. saeva* male parentage and *T. gigantea* males, nor for backcross matings between female F_1 hybrids of *T. saeva* female and *T. gigantea* male parentage and *T. saeva* males. Neither does it provide any indication about the ability of male F_1 hybrids to engage successfully in backcrosses. Such information can only be gained from careful laboratory breeding studies; however the morphometrical evidence suggests that backcrossing to *T. gigantea* may be infrequent. The allozyme data also revealed evidence of hybridization at the contact zone, showing an 'over-spill' of allele frequencies between the parapatric populations ('zones 3 and 4', Chapter 4). The apparent exchange of nuclear markers between species appears however to be largely limited to this area of parapatry - there was little evidence of long-range introgression away from the hybrid zone. In other words, although the clines for mtDNA and the nuclear allozyme markers appear to be coincident, they are not concordant. (It is important to state the need for more detailed transects of populations to either side of, and across the hybrid zone, in order to investigate possible clinal patterns in morphology, nuclear and mtDNA markers more thoroughly; the investigations presented here have all been based on rather large arbitrary geographic divisions. Furthermore, it is also important to investigate allozymes, morphology and mtDNA all within the same individual, so that one marker can be related to others). In fact, the allozyme analysis indicated that the parapatric *T. gigantea* 'population' actually contains relatively more alleles of *T. saeva* origin than the parapatric *T. saeva* 'population' contains alleles of *T. gigantea* origin. This observation is counter-intuitive in light of the above discussions of mtDNA

introgression and morphology, but perhaps it could be explained in the following, speculative, way. A large percentage of the individuals employed as samples in the allozyme electrophoresis were females (53% for the parapatric *T. saeva* (zone 3), and 66% for the parapatric *T. gigantea* (zone 4)), and as stated previously, hybrid females will generally be over-looked. *If* only crosses between *T. gigantea* females and *T. saeva* males result in fertile F_1 offspring (or at least they are far more frequently successful than the reciprocal cross), and *if* only backcrosses between F_1 females and *T. saeva* males are successful, then the presence of a number of unidentified F_1 hybrid females in the parapatric *T. gigantea* 'population' could explain this observation (because F_1 hybrids inherit half their genes from each parental species whereas backcrossed individuals in the *T. saeva* 'population' would contain more *T. saeva* genes). Of course, if F_1 hybrid females were also likely to be 'mis-classified' as *T. saeva* then this speculative argument would be undermined. Again, the need for more detailed transects of the hybrid zone, and knowledge of the viability and fertility of F_1, backcross, and preferably higher generation hybrids is clear. An example of higher generation hybrid breakdown is seen in the grasshopper *Caledia captiva.* In this species, F_1 hybrids between different chromosomal races are completely fertile but F_2 hybrids are totally inviable and backcrosses suffer approximately 50% inviability, both as a result of embyronic breakdown (Marchant *et al.*, 1988; Shaw *et al.*, 1988).

There are a number of possible explanations for the apparently long-distance introgression of *T. gigantea* mtDNA into *T. saeva* populations, compared with the narrow clines associated with the nuclear markers and morphology. This may truly represent introgression of mtDNA, originating at the hybrid zone. 'Cytoplasmic capture', where the mtDNA (or often cpDNA in plants) of one species is found to occur against a nuclear background that predominantly belongs to the 'capturing' species, has been reported frequently (Avise, 1994). One possible explanation for the often apparent ease of mtDNA transfer compared to nuclear markers is that the genes contributing to reproductive isolation will usually be predominantly housed in the nucleus; thus allozyme

markers, for instance, although often regarded as selectively neutral may themselves be selected against, or more frequently linked genes on the chromosomal segments that they mark may be selected against, as a result of disruption of coadapted gene complexes in the heterologous background of hybrids and backcrosses (Avise, 1994; Barton and Jones, 1983; Harrison *et al.*, 1987). Such effects may of course frequently lead to different clinal patterns in different markers (the studies presented here were not of sufficient geographic resolution to reveal non-concordance in allozyme or morphological markers). For similar reasons, mtDNA could introgress relatively easily if Haldane's rule is operating. In spiders, males are the heterogametic sex; if F_1 hybrid males (at least from *T. gigantea* female/*T. saeva* male crosses) are sterile or show reduced fertility then mtDNA introgression will occur more readily because 'the fertile females leave open an avenue for interspecific cytoplasmic exchange' (Avise, 1994), whereas the loci for male sterility map to sex chromosomes and autosomes and are subject to strong selection. This sexual asymmetry in hybrid fertility may explain the apparent ease with which mtDNA crosses some species boundaries, for instance in some *Drosophila* species (Avise, 1994) and in the voles, *Clethrionomys rutilus* and *C. glareolus* (Tegelström *et al.*, 1988). (See Coyne and Orr (1998) for a recent review of the possible genetic mechanisms behind Haldane's rule.)

There are also other possible explanations for the apparent long-range introgression of *T. gigantea* mtDNA into *T. saeva* populations. The observed pattern could simply reflect the translocation of individual spiders (by people). This would seem to be very likely given the ease with which these species are transported by people. Similarly, the individuals with captured mtDNA may reflect past hybridization events in other areas which simply remain as polymorphisms in the current population distributions. The pattern could also suggest that *T. saeva* has some competitive advantage over *T. gigantea* where they meet (perhaps ecological, but probably due to the genetics of hybridization) and that *T. saeva* has invaded from the west, gradually displacing *T. gigantea*, with a hybrid zone moving eastwards and leaving a trail

of introgressed mtDNA in its wake (as proposed for the hybrid zone between the chromosomal races of *Caledia captiva* in S. E. Queensland (Marchant *et al.*, 1988; Shaw *et al.*, 1988). The allozyme and morphometrical data do not support this - although selection could feasibly act to remove this evidence by acting against *T. gigantea* nuclear DNA as described above. Advancing hybrid zones have not been commonly reported (Hewitt, 1988) and this would seem to be an unlikely scenario. In the few cases that have been recognized this may have been a result of human activities (Hewitt, 1988), and the distributions of large house spiders are certainly affected to some extent by the movement of people (Chapter 2). Once again more detailed studies of this hybrid zone and laboratory studies of hybrids are required. Repeat surveys of Dorset and Hampshire in a few years time might reveal if the hybrid zone is moving. Finally, the asymmetric patterns of introgression observed in the mtDNA and morphometrical data, may not simply be a result of direct genetic effects but could also result from strong asymmetries in mate choice or mechanical barriers to interspecific mating. These possibilities are discussed below (section 7.4).

The comparisons made by Oxford and Smith (1987) between their plots of conductor + tegulum length against prosoma length and those of Merrett (1980), together with the morphometrical analyses of Oxford and Plowman (1991), suggested that there was a greater degree of hybridization in the York area compared to southern England (from where most of Merrett's (1980) specimens originated). Given the recent colonization of the York area by *T. saeva* and *T. gigantea*, and the possible existence of a longer-standing parapatric boundary between the two species in southern England, this suggested that there could be evidence of reinforcement in southern England and was the original inspiration for the work presented here. Although this may be true, it will be argued presently that *recent* species contact and frequent hybridization in the York area, compared to the apparently more stable southern situation, is in itself not a fair indication of the likelihood of reinforcement. This is because the contact between the two species in these two regions differs fundamentally in structure, in a way that is likely to

influence the likelihood of observing reinforcement for reasons other than the recency of contact. Nonetheless, both the morphometrical and the allozyme data presented here did reveal a high degree of hybridity in the specimens from the York area. Both the plots of tegulum + conductor length against prosoma length and the distributions of allele and genotype frequencies showed a great degree of overlap between *T. saeva* and *T. gigantea* in the York area compared to the quite discreet patterns for these characters in the two species from southern England. Furthermore, although both the morphometrical discriminant analysis and the allozyme frequencies suggested that few individuals of either species were likely to be 'pure' in the York area, the allozyme analyses and the regression and ANCOVA analyses of tegulum + conductor length against prosoma length indicated that *T. saeva* was experiencing a far greater degree of introgression in York than was *T. gigantea*. In each of these characters *T. saeva* showed a significant shift towards *T. gigantea*. A similar shift for *T. gigantea* towards *T. saeva* was not so apparent; indeed the York *T. gigantea* was very similar to the *T. gigantea* from southern England in terms of these characters (although the morphometrical discriminant analysis suggested that a shift was present). This also supports the observation, from southern England, that hybridization and introgression is largely asymmetric and that *T. saeva* has been affected more severely.

Comparison with the results of Oxford and Smith (1987) reveals an interesting discrepancy between the observations presented here and theirs. On the basis of their regressions of tegulum + conductor length against prosoma length they concluded that *T. saeva* from the York area were very similar to *T. saeva* from southern England (Merrett's (1980) data), whereas *T. gigantea* from the city of York and the immediate surrounding area were shifted towards *T. saeva*. Furthermore, they argued that this shift was probably not an artifact of different people performing the measurements in York and in southern England because one would have then expected a shift in *T. saeva* as well. They suggested that introgression was having more of a morphological impact on *T. gigantea* than *T. saeva* in the York area, and that this could result from the

numerical superiority of *T. saeva* in and around the city. The opposite pattern would then have been expected in 'the rest of Yorkshire' where *T. gigantea* is more numerous than *T. saeva*, however their data indicated that this was not this case: *T. gigantea* was also morphologically shifted towards *T. saeva* in the rest of Yorkshire (however the two species are patchily distributed across Yorkshire and assuming an homogenous mixture of the two species is not really valid (G. S. Oxford, pers. comm.). This intriguing discrepancy between the current data and that of Oxford and Smith (1987) may suggest that the observable pattern of introgression in the York area really has changed during the last 10 years, and that these recently arrived species have not yet found an equilibrium with regard to their interspecific interactions. As mentioned above, Oxford and Smith (1987) attributed their observed shift in *T. gigantea* towards *T. saeva* to the numerical inferiority of the former in York; which would therefore be more likely to engage in hybridization and backcrossing. This could certainly be a factor. It has already been suggested that backcrossing of F1 hybrids to *T. gigantea* may be very infrequent, however both Oxford and Smith's (1987) observations, as described above, and the present discriminant analyses of morphology (in particular) indicate that it has occurred. The numerical inferiority of *T. gigantea* in York leading to an increased likelihood of interspecific and backcross matings would multiply the chances of successful backcrossing being observed, but another possibly important factor could be the degree of introgression experienced by *T. saeva.* York *T. saeva*, being shifted towards *T. gigantea* presumably through introgression, by definition would contain a significant proportion of *T. gigantea* genes and therefore 'F_1' hybrids would contain, on average, more *T. gigantea* genes than *T. saeva* genes and perhaps be more likely occasionally to produce more viable/fertile backcross progeny with *T. gigantea* than those from southern England.

7.3 Postzygotic Isolation

Two major assumptions of this thesis have been that 1) the members of the *T. atrica* group are truly distinct species and do not simply reflect some geographical variation in the morphology of a single species, and 2) that some degree of postzygotic isolation exists between them (hybrid zygote inviability, hybrid sterility, hybrid breakdown). These two assumptions are interlinked; adherence to the biological species concept (BSC) implies, by definition, that the species are reproductively isolated. The presence of gene flow between *T. saeva* and *T. gigantea* clearly negates their species status under a strict interpretation of the BSC and they should therefore more correctly be regarded as subspecies or incipient species. However, the BSC - like all strict 'species concepts' - has practical difficulties (see Chapter 1) and, as Mallet (1995) forcefully points out: 'the interesting part of speciation is divergence into genetic clusters that can co-exist, not the final demise of gene transfer' (although it *is* the demise of gene transfer that is of central interest in this thesis). In a line of reasoning that encompasses the views of Darwin and Wallace and incorporates modern genetics, Mallet (1995) highlights the real yet dynamic nature of species as distinct morphological (or better, genotypic) clusters: gene flow can exist between two parapatric or sympatric species but they will still exhibit a bimodal distribution, and will not fuse, if the force of disruptive selection exceeds the ability of gene flow to fuse the clusters. If the disruptive selection encouraging divergence is maintained then the two species may become *completely* isolated. Gene flow is therefore not only a fundamental part of the speciation process but is a force that continues to influence real species (unless completely allopatric). Accepting that gene flow can and does occur between many species, without negating their specific status, there are a number of reasons for believing that *T. atrica*, *T. saeva* and *T. gigantea* are valid species. Firstly, the species are separated by considerable differences in mitochondrial CO1 DNA sequence, placing their divergences in the early part of the Pleistocene. If they simply represent morphological variants then they have must have survived a number of substantial range changes over a

considerable period of time. Secondly, *T. atrica*, *T. saeva* and *T. gigantea* are consistently recognized throughout their distributions, even where introduced, despite the frequent presence of intermediate individuals. Thirdly, in southern England where *T. saeva* and *T. gigantea* meet they show a distinct boundary with little overlap, distinct gene frequencies, and asymmetric hybridization - they therefore do not freely interbreed.

Accepting that the members of the *T. atrica* group are distinct species then it seems likely that there must be a degree of postzygotic isolation between the species. As has already been emphasized, there is a great need for detailed laboratory work to quantify the effects of hybrid zygote inviability, hybrid sterility, and hybrid breakdown in crosses between these species. Without the empirical observations of breeding trials, and microscopical examinations of gametogenesis in the gonads of hybrids, the patterns of postzygotic isolation between *T. saeva* and *T. gigantea* can only be inferred from the patterns of hybridization and introgression observed in the wild. These patterns of hybridization and introgression are themselves interpreted partly through assumptions about postzygotic barriers; hence without empirical data on the magnitude of postzygotic isolation there is some unavoidable circularity in the arguments. Nonetheless, the distinctness of the species both in terms of genetic divergence, morphology, and range, the apparent asymmetry in hybridization, and the lack of any obvious differences in ecology or habitat preferences strongly suggests that postzygotic isolation does have a strong impact on these species. The hybrid zone in southern England therefore seems likely to be a tension zone, maintained by the balance between dispersal and selection against hybrids. The small amount of evidence available from the breeding studies (one *T. gigantea* female x *T. saeva* male produced progeny and 13 offspring were reared with a mortality rate of around 54% compared to about 4% in non-hybrid progeny) suggested that F_1 hybrids may suffer from some developmental instability.

Clearly the genetic basis of any postzygotic isolation that exists between *T. saeva* and *T. gigantea* (and between these two species and *T. atrica*) is unknown. However, it seems most likely that genetic incompatibilities would result from the inability of different alleles to function properly when brought together in a heterologous genetic background (Coyne and Orr, 1998) - the disruption of coadapted epistatic interactions between loci (Avise, 1994; Palmer and Strobeck, 1986). The Dobzhansky-Muller Model (Coyne and Orr, 1998) suggests a possible mechanism by which postzygotic isolation can evolve. If changes in a single gene are involved it is hard to conceive how isolation can evolve: if one species has genotype *AA* and the other species *aa*, and *Aa* hybrids are completely sterile, then the alternative allele can never be fixed because the first mutant would have genotype *Aa* and therefore be sterile. However, if changes at more than one locus are involved then hybrid inviability and sterility can evolve unimpeded by selection: if the ancestral species has genotype *aabb*, then a new mutation (allele *A*) could become fixed by selection or drift in an isolated population (*Aabb* and *AAbb* genotypes are perfectly fit). Similarly a new mutation (allele *B*) could become fixed in another population. When the *AAbb* and *aaBB* populations come into contact the resulting *AaBb* hybrids could be sterile or inviable because the *A* and *B* alleles have never been tested together within a genome and may not function properly when united in hybrids. It seems unlikely that major chromosomal rearrangements (for example, fusions) play a role in the postzygotic isolation of these *Tegenaria* species. Although the karyotypes of the *T. atrica* group warrant further investigation (especially using banding techniques), it is probable that all three species exhibit the same chromosome number: 42 acrocentric chromosomes in males and 44 acrocentric chromosomes in females. The karyotype of *T. atrica* was published by Král (1995), and also by Revell (1947) - although this latter paper almost certainly examined *T. gigantea* and not *T. atrica* as reported (G. S. Oxford, pers. comm.). A limited examination of karyotypes from embryos of *T. saeva* and *T. gigantea* also suggested a chromosome number of 42/44 (pers. obs.). The presence and effects of minor deletions, translocations, inversions, and so on cannot be ruled out.

7.4 Prezygotic Isolation, Reinforcement, and Character Displacement

Two aspects of prezygotic isolation between *T. saeva* and *T. gigantea* were investigated in this thesis: 1) mechanical isolation and, 2) sexual/behavioural isolation (courtship and mating behaviour). Other forms of possible prezygotic isolation were not explicitly investigated but, as discussed in Chapter 1, there appear to be no obvious differences in ecology (habitat preferences) or phenology between the species. Observations of the copulatory behaviour of *T. saeva* and *T. gigantea* demonstrated clearly that there were potentially strong mechanical barriers to heterospecific copulation between these two species. These mechanical barriers stemmed directly from differences in the morphology of the palps and epigynes of the two species and manifested themselves through visibly more awkward palpal insertions that were very significantly shorter in duration for heterospecific pairings than for conspecific pairings. Males in heterospecific pairings experienced both slippage of the palp (conductor) from the female copulatory duct, and failure to engage the palp with the epigyne (failure to locate the palpal tibial apophyses with the apophysis on the epigyne - 'apophysis mismatch'). Palpal insertion duration was significantly shorter for pairings between *T. gigantea* females and *T. saeva* males than for the reciprocal cross. These mechanical limitations on heterospecific pairing could well have been sufficient to limit the chances of fertilization and hence interspecific hybridization. Whether or not the males simply fail to reach the female spermatheca with the embolus, and therefore fail to place their sperm in the correct place, should be examined by microscopy post-mating. Some support for this came from the observation that most undeveloped eggs in the mating trials appeared unfertilized (yolky), rather than having experienced arrested development (however the possibility of successful fertilization followed by very early developmental arrest can not be ruled out). Although mechanical isolation may have a very strong effect on the likelihood of a reproductively successful heterospecific copulation this would seem unlikely to be able to explain the asymmetry of introgression observed. F_1 hybrids, that are intermediate in morphology (the morphometrical analyses

suggest that this is true for males but the intermediacy of female hybrids was not clear), would be expected to experience less mechanical isolation when backcrossing with parental types.

The regressions of mean palpal insertion duration on the female/male genital size ratio for the heterospecific crosses in Chapter 6 indicated that males with small palps relative to the female epigyne enjoyed longer palpal insertions in *T. gigantea* female x *T. saeva* male pairings, whereas relatively larger males might be more successful in *T. saeva* female x *T. gigantea* male pairings. This led to the suggestion that a small increase in the genital size of *T. saeva* (males and females), or a small decrease in the genital size of *T. gigantea* might substantially increase the mechanical isolation between the species. A small change in overall size of these genitalia characters relative to body size would probably involve a relatively simple ontogenetic change, especially since in *Tegenaria* maturity can be reached in several instars and therefore the potential for early or late maturity already exists (Merrett, 1980). However, despite evidence suggesting that increasing mechanical isolation would be relatively straightforward, the morphometrical analyses of Chapter 3 provided no evidence of reinforcement in the parapatric populations from southern England compared to the allopatric populations (no divergence in mean character values and no decrease in character variance in parapatry). This result was important for two reasons. First, it indicated that reinforcement of mechanical isolation has not been occurring in *T. saeva* and *T. gigantea*. Secondly, it refutes the 'lock-and-key' hypothesis despite having demonstrated a simple route by which a small change in genital morphology could increase mechanical isolation in these species. Previous studies that have failed to find evidence in favour of the 'lock-and-key hypothesis' in spiders, for example that of Ware and Opell (1989), had the weakness of simply looking for 'character displacement' in the sclerites of the genitalia in areas of sympatry or parapatry - without any precise knowledge of how the sclerites, or *which* sclerites, might interact to facilitate mechanical isolation. The results presented here argue against the differences in genital morphology, seen so commonly in closely related spiders

(and other invertebrates), having evolved to facilitate mechanical isolation. They therefore lend support to alternative hypotheses such as cryptic female choice (Eberhard, 1985).

Perhaps the lack of evidence in support of reinforcement and the 'lock-and-key hypothesis' of mechanical isolation is not surprising. It seems obvious that potential partners should endeavour to assess each other's quality (i.e. specific status) as soon as possible - early in courtship - and certainly before copulation occurs (Ware and Opell, 1989) (although females of many species may assess mate quality during copulation and manipulate sperm or modulate their future mating strategies accordingly (Andersson, 1994; Eberhard, 1985)). There were no dramatic differences in the form of distinct behavioural elements between the courtship behaviours of *T. saeva* and *T. gigantea*. However the courtship of these species did differ in the overall frequency of different behavioural elements. These differences were significant in comparisons between the allopatric populations of the two species (P = 0.0192) but were *much greater* in comparisons between the parapatric populations (P = 0.0007, Table 6.12). Further, comparisons between individual transitions (dyads) across transition matrices revealed that these greater differences were a result of an emphasis on a few behavioural transitions - there were fewer significant transitions in parapatry than in allopatry. This could be interpreted as a 'fine-tuning' of behaviour, or a reduction in behavioural variance. A greater divergence between the species and less variance in behaviour both fit with selection for species-specific mate recognition. The courtship data also suggested that a greater proportion of the difference observed between the species was attributable to fine-tuning of the courtship behaviour in *T. gigantea*; with *T. saeva* showing few differences between allopatry and parapatry. It was also suggested that the differences between allopatry and parapatry resulted to a large degree from more 'cautious' responses of *T. gigantea* females to courting males in parapatry. In agreement with the idea that discrimination should occur as early as possible in interactions between the two sexes, most significant differences between the species, both in

allopatry and parapatry, resulted from either early male behaviours ('Phase 1' or chemoexploratory behaviours) or from female responses. In general males of both species, from allopatry and parapatry, initiated courtship with females regardless of species. Although the transition analysis had limitations (see Chapter 6), the results strongly suggested a divergence and refinement of courtship behaviour of the two species in parapatry, with this being much more evident in *T. gigantea*. There was also some evidence of aggressive discrimination by *T. saeva* females against heterospecific males - but there was no evidence for a significant increase in discrimination in parapatry. Overall, the behavioural analyses appear to support the reinforcement hypothesis - is this true and what alternative explanations might there be?

As stated previously, the apparently recent contact of *T. saeva* and *T. gigantea* in the York area compared to the longer-standing contact area in southern England, plus the apparently greater degree of hybridity in the York area, suggested that evidence for reinforcement might be observable in southern England. Any study claiming to demonstrate reinforcement must fulfil four requirements (Butlin, 1989), as outlined in Chapter 1 (section 1.2.4). To restate, these requirements are:

"1) that gene flow occurs between the taxa, or did occur when they originally met,
2) that components of the mate recognition system have diverged in the area of contact and in the time since contact was established,
3) that this divergence is sufficient to alter the pattern of mating in a way that decreases the frequency of production of unfit hybrid genotypes, and, ideally,
4) that divergence is not a result of other selection pressures on the mate recognition system".

The first of these requirements - that gene flow occurs between the taxa - has been demonstrated quite conclusively and discussed in detail above. The behavioural data indicate that male courtship and female responses have

diverged in the area of contact in southern England (requirement 2), and that this divergence has been in such a way as to make the behaviour of *T. saeva* and *T. gigantea* more distinct and less variable - and is therefore likely to reduce the frequency at which hybrid genotypes are produced (requirement 3). However more extensive studies would be valuable and should include not only more mating trials but choice experiments where females are presented with males of both species. Sequential mating experiments could also be performed in which females are presented with a males of one species, allowed to mate, and then presented with a male of the other species. Courtship and copulatory parameters are timed, and sperm priority patterns assessed. Sequential mating experiments may be more natural than simultaneous choice experiments. A test of assortative mating is required to ensure that apparent divergences in behaviour have not resulted from other unforeseen pressures on the mating system (requirement 4) - although it is hard to imagine what these might be. The behavioural analyses were very preliminary; examining the frequencies of transitions between behaviours. Therefore future work should also aim to explore species-specific differences within behavioural elements, especially the vibrational signals generated by the males - the abdomen kicks performed during chemoexploratory behaviour ('Phase I' courtship), and the palpal drumming and abdomen vibrations performed in Phase II courtship. Previous studies (reviewed in Chapter 6) suggest that these are generally-species specific (for example Boulanger *et al.*, 1986; Krafft, 1978; 1982; Krafft *et al.*, 1978; Leborgne, 1984; 1986; 1989; Leborgne and Krafft, 1979; Leborgne *et al.*, 1980; Mielle, 1978). The possibility that the web pheromones of females differ between species also warrants investigation.

The observation that *T. gigantea* showed more evidence of possible reinforcement fits with the observation that there has been little introgression of *T. saeva* genes into *T. gigantea* populations in southern England. Low levels of introgression (backcrossing) imply low levels of recombination between *T. saeva* and *T. gigantea* genes in *T. gigantea* populations. The selection-recombination antagonism constitutes the major theoretical objection to models

of reinforcement (Butlin, 1989; Hostert, 1997; Howard, 1993). Any reduction in recombination would make reinforcement more likely (Hostert, 1997; Trickett and Butlin, 1994). The low levels of introgression experienced by *T. gigantea* populations (presumably because backcrosses are inviable) compared to *T. saeva* population implies that reinforcement should be more evident in *T. gigantea*; this appears to be the case (of course it could be argued that low levels of introgression were observed *because* of reinforcement). Female *T. gigantea* that mate with male *T. saeva* would suffer a loss of fitness for three reasons. First, these pairings would tend only rarely to produce offspring; secondly, the offspring may suffer some developmental problems or sterility; thirdly, in the absence of backcrossing, the genes of the *T. gigantea* females are lost to the *T. gigantea* gene pool.

The discussion of prezygotic isolation and reinforcement has focused on the populations of *T. saeva* and *T. gigantea* from southern England. Prezygotic isolation was not examined in samples from the York area where there is a much greater degree of hybridization compared to southern England. Increased hybridization may partly reflect the recency of contact in the York area compared to southern England, where processes such as reinforcement have had more time to act, but may also be a function of species distribution patterns. Distribution patterns can greatly affect the frequency of hybridization between taxa and the possibility of reinforcement. Patchiness in the distribution of species populations within a hybrid zone (a 'mosaic'), often as a result of environmental heterogeneity, could increase the frequency of contact between individuals. Patch size could have a critical effect with typical tension zones existing at the boundaries of large patches and small patches equating to a uniform environment. Intermediate patches could result in a broader zone of contact with some advantageous parental alleles being maintained in patches where they confer greater fitness - and increasing the likelihood of reinforcement (Harrison and Rand, 1989; Hewitt, 1989). Howard (1993) emphasizes that no regular geographical pattern of divergence in an isolating trait would be observed in a mosaic hybrid zone, and also that

pressure exerted by disadvantageous hybrid matings on a taxon would be greatest where that taxon is least abundant. The likelihood of reinforcement would therefore vary from patch to patch in a mosaic and with location within a 'linear' hybrid zone. The general conclusion from these studies is that the pooling of results from different localities within the contact zone should be avoided. This represents a valid, but inescapable, criticism of the data presented here and would reduce the chances of detecting reinforcement. A recent paper by Servedio and Kirkpatrick (1997) modelled the effects of gene flow on reinforcement. In their paper they investigated the likelihood of reinforcement occurring under patterns of symmetrical and asymmetrical migration. They concluded that reinforcement was possible under a range of conditions when there was symmetrical migration, but that it was unlikely when migration was asymmetrical. In their system asymmetrical migration was modelled as one-way migration from a large population to a smaller island population; reinforcement was unlikely because of the dilution and elimination of novel mate-preference genes by the influx of parental genes from the mainland. The distributional survey presented here, and the survey of Oxford and Smith (1987), both indicate that *T. saeva* is most abundant in and around the city of York, with *T. gigantea* prevailing in the surrounding rural areas. It is proposed therefore, that the distribution of the two species in the York area in some ways resembles an island model - with *T. saeva* occupying an island with continual immigration pressure from the surrounding *T. gigantea* populations. Under such a scenario reinforcement may be less likely to occur than in the parapatric situation in southern England. The likelihood of reinforcement developing in the two study areas is therefore a function of two processes - the recency of contact and species distribution patterns - which cannot at present be disentangled. It would be valuable to examine areas of contact between *T. saeva* and *T. gigantea* in other regions.

7.5 Conclusions

The evolutionary interactions between the members of the *T. atrica* group are complex, and are likely to vary in detail from area to area dependent upon population structure and recency of contact. It seems likely that crosses between *T. saeva* females and *T. gigantea* males are very unlikely to be successful, whereas crosses between *T. gigantea* females and *T. saeva* males may occasionally produce progeny. The viability and fertility of these F_1 progeny remains unclear, but it seems likely that they do experience some developmental difficulties. Asymmetrical patterns of introgression indicated by all markers suggest that backcrossing between hybrid individuals and *T. gigantea* parental types will be rare (and that this stems from a postzygotic, i.e. genetic barrier), whereas the reciprocal backcross must occur. It also seems possible that F_1 hybrid males are sterile, although this has not been demonstrated. If these inferred patterns of asymmetric postzygotic isolation are correct then they are very similar to those that have been recognized in the crickets *Gryllus pennsylvanicus* and *G. firmus* (Harrison *et al.*, 1987). There was no evidence of reinforcement of mechanical isolation (arguing against the 'lock-and-key hypothesis' for the evolution of species-specific genitalia) in parapatry in southern England. However, there were indications of reinforcement in species-specific courtship behaviours. This occurred most notably in *T. gigantea* which suffered less introgression, and therefore less selection-recombination antagonism which would hinder the reinforcement process. A greater degree of hybridization and introgression was observed in the York area. This may partly result from recency of contact but also from the geographical population structure of the two species in this region.

Some 318 pages ago this thesis opened with a quote from *Charlotte's Web* by E. B. White. 'Charlotte A. Cavatica' was clearly a species of *Araneus*. If she had not been, and if she had been addressing an evolutionary biologist or an arachnologist and not a young pig named Wilbur, then she would undoubtedly have been a member of the *Tegenaria atrica* group. In terms of representing a

novel system - a remarkable and tangled web of species' interactions - with which to explore extant problems of evolutionary theory, she and her sisters, although perhaps not particularly flashy, certainly will 'do'. It is hoped that the work presented here has gone some way to unraveling the interactions between these species and will provide an, albeit rudimentary, base from which more detailed studies can develop. There is much to do. Had W. S. Bristowe been able to overcome his abhorrence of the long-legged *Tegenaria* (although apparently the disgust was not sufficient to overcome a contrary apparent predilection for ingesting them on several occasions!), and had he been aware that his '*T. atrica*' would actually turn out to be not *T. atrica* at all but *two* related species, *T. saeva* and *T. gigantea*, then I am sure he *would* have embarked on a long and detailed study of house spiders.

Appendices

The following appendices provide additional information relating to Chapters 4 (A.1), 5 (A.2), and 6 (A.3).

A.1 Allozyme Reagents, Procedures, and Genetic Distances (Chapter 4)

The following series of tables describe:

1) the main reagents (Table A.1.1) and procedures (Table A.1.2) used for the four polymorphic enzyme systems analyzed in Chapter 4;

2) the various matrices of genetic distances used to generate the 'population' trees in Chapter 4 (Tables A.1.3, A.1.4, A.1.5).

Table A.1.1. The chemicals used in the electrophoretic analysis.

Chemical	Abbreviation*	Supplier¤	Storage conditions Dilution§	Temperature (°C)
cis-Acontic acid		Sigma	70 mM	-20
Citric acid (monhydrate)	Citrate	BDH		20
D-fructose-6-phosphate	F6P	Sigma	0.16 M	-20
Glucose-1-phosphate	G1P	Sigma		-20
Glucose-6-phosphate dehydrogenase (E.C. 1.1.1.49)	G6PD	Sigma	1 i.u. μl^{-1}	4
Glycine (ammonia free)		Sigma		20
Hydrochloric acid	HCl	BDH	1.0 M, 5.0 M	20
Iodine		BDH		20
Isocitric dehydrogenase (E.C. 1.1.1.42)	IDH	Sigma	0.5 i.u. μl^{-1}	4
Potassium Iodide	KI	Fisons		20
Magnesium acetate	Mg acetate	Sigma	0.25 M	4
Mercaptoethanol		Sigma	#	20
MTT		Sigma	24 mM	4
NAD		Sigma	1.5 mM	-20
NADP		Sigma	1.3 mM	-20
PMS		Sigma	8.2 mM	4
Sodium *di*hydrogen orthophosphate	$NaH_2PO_4.2H_2O$	BDH		20
*Di*sodium hydrogen orthophosphate	Na_2HPO_4	BDH		20
Sodium hydroxide	NaOH	BDH	2.0 M	20
Tris (Trizma base)		Sigma		20
Triton X-100	TX100	Sigma		20

* As used in Table A.1.2.

¤ Sigma Chemical Co. Ltd., Poole, Dorset and BDH Chemicals Ltd., Poole, Dorset.

§ All dilutions aqueous.

1 µl of a 10% solution added to sample wells for electrophoresis of *Gpi-1*, to remove sub-bands.

Table A.1.2. The electrophoretic method for each of the four polymorphic enzymes.

Enzyme	Locus	Tank/gel buffer	Run time (min)	Volt-age	Position of appli-cation* (mm)	No. of appli-cations	Stain develop-ment time (min)	Approx. volume of stain solution (ml)¤	Constituents of stain solution on gel (approximate final concentration)¤
Aconitase (aconitate hydratase) E.C. 4.2.1.3	*Aco-1*	40 mM Tris 10 mM Citrate pH 7.6	40	200	36/6	1	20-60	2.6	77 mM Tris HCl pH 8.0 19 mM Mg acetate 0.5 mM NADP 2.9 i.u. ml^{-1} IDH 26.5 mM cis-Aconitic acid 0.31 mM PMS 0.92 mM MTT
Amylase E.C. 3.2.1.1	*Amy-1*	30 mM $NaH_2PO_4.2H_2O$ 10 mM Na_2HPO_4 pH 6.3	120	150	24	1	30§	- §	1% I_2 / 0.5 % KI solution§
Glucose-6-phosphate isomerase E.C. 5.3.1.9	*Gpi-1*	25 mM Tris 0.19 M Glycine pH 8.5	40	200	42/6	1	1-5	2.4	250 mM Tris HCl pH 8.0 21 mM Mg acetate 0.54 mM NADP 14 mM F6P 2 i.u. ml^{-1} G6PD 0.68 mM PMS 2 mM MTT
Phospho-glucomutase E.C. 5.4.2.2	*Pgm-1*	25 mM Tris 0.19 M Glycine pH 8.5	35	200	42/6	1	15-30	3.3	182 mM Tris HCl pH 8.0 76 mM Mg acetate 59 mM NADP 0.33 M G1P 3 i.u. ml^{-1} G6PD 0.25 mM PMS 0.73 mM MTT

* Position is measured from the cathodal edge of the plate. ¤ Volumes/concentrations excluding the 4 ml agar overlay. § Staining procedure for amylase was different: after electrophoresis the cellulose acetate plate was placed face down on a 1% agar / 0.25% starch plate, at room temperature in the dark for 30 minutes. 1% I2 / 0.5% KI solution was then applied to both agar and cellulose acetate plates to resolve the banding pattern.

Table A.1.3. Pairwise genetic distances from allozyme data: *Dg* and Roger's distances

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Zone 1&2	0.000	0.341	0.861	0.952	0.702	0.785	0.831	0.898	0.911
(2) Zone 3	0.058	0.000	0.776	0.912	0.524	0.701	0.726	0.848	0.848
(3) Zone 4	0.238	0.195	0.000	0.485	0.469	0.370	0.276	0.406	0.378
(4) Zone 5&6	0.336	0.289	0.120	0.000	0.747	0.669	0.555	0.348	0.378
(5) York *T. saeva*	0.165	0.123	0.080	0.191	0.000	0.404	0.384	0.602	0.626
(6) York *T. saeva?*	0.214	0.178	0.074	0.181	0.069	0.000	0.255	0.504	0.612
(7) York *T. saeva/gigantea?*	0.225	0.182	0.046	0.125	0.080	0.101	0.000	0.542	0.380
(8) York *T. gigantea?*	0.283	0.237	0.071	0.122	0.122	0.086	0.104	0.000	0.546
(9) York *T. gigantea*	0.277	0.230	0.087	0.069	0.143	0.156	0.077	0.122	0.000

Above the diagonal: *Dg*, an estimate of genetic distance calculated as 1 - *Sg*, where *Sg* is the pairwise estimate of 'genotypic overlap. Below the diagonal: Roger's (1972) genetic distance with Wright's (1978) modification as calculated by the BIOSYS-1 program (Swofford and Selander, 1981).

Table A.1.4. Pairwise genetic distances from allozyme data: Nei's and Cavalli-Sforza distances

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Zone 1&2	0.000	0.016	0.144	0.306	0.063	0.108	0.122	0.256	0.178
(2) Zone 3	0.004	0.000	0.095	0.211	0.027	0.078	0.066	0.171	0.103
(3) Zone 4	0.076	0.051	0.000	0.052	0.034	0.015	0.015	0.054	0.026
(4) Zone 5&6	0.142	0.102	0.011	0.000	0.110	0.097	0.061	0.019	0.040
(5) York *T. saeva*	0.036	0.020	0.008	0.037	0.000	0.027	0.021	0.082	0.041
(6) York *T. saeva?*	0.061	0.042	0.007	0.029	0.006	0.000	0.036	0.075	0.064
(7) York *T. saeva/gigantea?*	0.069	0.045	0.002	0.016	0.009	0.013	0.000	0.056	0.019
(8) York *T. gigantea?*	0.107	0.074	0.006	0.009	0.019	0.009	0.013	0.000	0.052
(9) York *T. gigantea*	0.108	0.074	0.010	0.007	0.028	0.031	0.008	0.019	0.000

Above the diagonal: The Cavalli-Sforza distance (Cavalli-Sforza and Edwards, 1967). Below the diagonal: Nei's (1972) genetic distance. Both calculated by the GENDIST program from the PHYLIP sofware package (Felsenstein, 1995).

Table A.1.5. Pairwise genetic distances from allozyme data: Reynolds distance.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Zone 1&2	0.000								
(2) Zone 3	0.014	0.000							
(3) Zone 4	0.201	0.139	0.000						
(4) Zone 5&6	0.304	0.233	0.037	0.000					
(5) York *T. saeva*	0.104	0.058	0.027	0.106	0.000				
(6) York *T. saeva?*	0.174	0.122	0.025	0.092	0.021	0.000			
(7) York *T. saeva/gigantea?*	0.170	0.114	0.009	0.046	0.025	0.042	0.000		
(8) York *T. gigantea?*	0.274	0.202	0.024	0.034	0.065	0.036	0.046	0.000	
(9) York *T. gigantea*	0.237	0.171	0.030	0.020	0.075	0.093	0.022	0.062	0.000

Reynolds, Weir, and Cockerham's (1983) genetic distance is shown, as calculated by the GENDIST program from the PHYLIP sofware package (Felsenstein, 1995).

A.2 Molecular Reagents (Chapter 5)

Details for the main reagents used in DNA extraction and gel electrophoresis follow. All other reagents are specified in the text or follow the appropriate manufacturer's instructions and supplies.

A.2.1 DNA Extraction

Reagent:	*Storage conditions:*
1. *Lysis Buffer:* 50 mM Tris HCl, pH 8.0 1% Sodium dodecyl sulfate (SDS) 100 mM NaCl 1% beta-mercaptoethanol	-20°C in 10 ml aliquots
2. 180 i.u. ml-1 Proteinase K	-20°C in 1 ml aliquots
3. 1000 i.u. ml-l RNAse A	-20°C in 0.5 ml aliquots
4. 5 M NaCl	4°C
5. Absolute ethanol	-20°C in 20 ml aliquots
6. 70% Ethanol	-20°C in 20 ml aliquots
7. *1X TE Buffer*: 10 mM TRIS pH 7.4 1 mM EDTA pH 8.0	4°C

A.2.2 Gel Electrophoresis

Reagent:	*Storage conditions:*
1. *1X TAE Electrophoresis Buffer:*	4°C
40 mM Tris acetate pH 8.0	
1 mM EDTA pH 8.0	
2. Orange G Loading Dye:	20°C
50% w/v Glycerol	
10 mM Tris HCl pH 8.0	
25 mM EDTA	
Orange G Powder (Sigma) as suffucient.	

A.3 Behavioural Transitions (Chapter 6)

The following two tables (Table A.3.1 and Table A.3.2) display the preceding behavioural acts that significantly ($P < 0.01$) facilitated or inhibited a following act in the preliminary transition matrix analyses for each pairing combination (type of cross) (see section 6.3.2).

Table A.3.1. ALLOPATRIC PAIRINGS: Preceding acts that significantly facilitate or inhibit following acts ($P < 0.01$).

	TSf x TSm		TGf x TGm		TGf x TSm		TSf x TGm	
Behaviour	**Facilitates**	**Inhibits**	**Facilitates**	**Inhibits**	**Facilitates**	**Inhibits**	**Facilitates**	**Inhibits**
A. MLSTA	B, C		B, C		B, C		B	
B. MLCE	G, Q		A, G	H	A, D, G	H	A, G	
C. MLSF	******		B,		******		******	
D. MLBW		F, I		E, F, M, N, R	B		******	
E. MLRND	A	F, I	******		A	F		D, I, J, Q
F. MLMP	B	D, E, I, Q		D, N		E, Q		E, I, Q
G. FLT	O, P		E, K, O		O		O	
H. MLDVA	G		G, I, J	B	G, I		G	
I. MLWA		D, E, F, Q	******		******			E
J. MLOF	******		******		******		******	
K. FTF	L		L		L		L	
L. STS	M		M		M		M	
N. REP	******			D, F	******		******	
O. FEW	H		H		H	G	H	
P. FEOM	K		******		K		K, Q	
Q. FEAM		F, I	******		J, K	F		E
R. FESIG	******		******		******		******	

Asterixes indicate transition matrix rows with a row chi^2 less than 33.41 ($P < 0.01$; 17 d.f.) which were excluded from the analysis.
TSf = *T. saeva* female, TGm = *T. gigantea* male, and so on. The body of the table uses the single-letter codes for each behaviour (see section 6.3.1 or Table 6.12).

Table A.3.2. PARAPATRIC PAIRINGS: Preceding acts that significantly facilitate or inhibit following acts ($P < 0.01$).

	TS x TS		TG x TG		TGf x TSm		TSf x TGm	
Behaviour	**Facilitates**	**Inhibits**	**Facilitates**	**Inhibits**	**Facilitates**	**Inhibits**	**Facilitates**	**Inhibits**
A. MLSTA	B		B		B, C		******	
B. MLCE	C, D, G	H	******		G, P		******	
C. MLSF	******		D		******		B	N
D. MLBW	B	F, N	******		B		******	
E. MLRND	A		******		A		A	
F. MLMP		D, N	******		******			N
G. FLT	O		O		O		K, O, P	
H. MLDVA	G, I, J		G		G, I		G, J	
I. MLWA	******		******		******			N, Q
J. MLOF	H		******		******		******	
K. FTF	L		L		L		L	
L. STS	M		M		M		M	
N. REP		D, F	******		******			C, D, F, I, J, Q, R
O. FEW	H		H		H		H	
P. FEOM	******		K		******		******	
Q. FEAM	******		******		******			I, N, Q
R. FESIG	******		******		******		******	

Asterixes indicate transition matrix rows with a row chi^2 less than 33.41 ($P < 0.01$; 17 d.f.) which were excluded from the analysis.
TSf = *T. saeva* female, TGm = *T. gigantea* male, and so on. The body of the table uses the single-letter codes for each behaviour (see section 6.3.1 or Table 6.12).

Glossary

A *selected* glossary of terms follows. This glossary aims to clarify mainly those relevant arachnological terms (and some phylogenetic terms) which were not defined elsewhere in the text. Definitions follow those in the glossary of the British Arachnological Society Members' Handbook (BAS, 1989) except entries marked * which follow Ridley (1993).

Autapomorphy	A derived character unique to a given species or other monophyletic group.
Abdomen	See Opithosoma.
Allopatric*	Geographically separated populations; as in *allopatric speciation*.
Apomorphy	A relative term referring to a character hat is derived from and differs from an ancestral or generalized condition. (*cf.* Plesiomorphy).
Apophysis	A sclerotized process or projection.
Araneomorphae (=labidognatha)	A sub-order of the Araneae containing the 'modern' spiders with pincer-like (diaxial) articulation of the chelicerae. (*cf.* Mygalomorphae).
atrica	(L: *atrium*) a hall: of the house.
Carapace	The shield of exoskeleton covering the dorsal surface of the prosoma.
Cephalothorax	See Prosoma.
Chelicerae	The paired jaws, consisting of a large basal portion (paturon) and a fang.
Clypeus	The area between the anterior row of eyes and the anterior edge of the carapace.
Conductor	A semi-membranous structure in the male palp which supports and guides the embolus during copulation.

Coxa — The segment of leg nearest the body. Modified in the palp to form the maxilla.

Cymbium — The broadened, hollowed out, tarsus of the male palp to which the palpal bulb is attached. The hollowed out region is known as the alveolus or operculum.

Dorsal view — Viewed from above (the 'back' of the spider).

Ectal view — Viewed from the outside (usually referring to a paired asymmetrical structure).

Embolus — The structure containing the terminal portion of the ejaculatory duct and its opening in the male palp; whip-like in *Tegenaria.*

Entelegyne — Spiders in which the females have external genitalia in the form of an epigyne having two symmetrical halves. (*cf.* haplogyne).

Epigastric fold/furrow — A fold and groove separating the region of the book lungs and epigyne from the more posterior portion of the ventral surface of the abdomen.

Epigyne (or epigynum) — A more or less sclerotized and modified external structure associated with the reproductive openings of the adult females of most spider species.

gigantea — (L: *giganteus*) of the giants.

Haplogyne — Spiders in which the females have little or no external geintalic structure or epigyne. (*cf.* Entelegyne).

Haematodocha — A balloon of connective tissue between groups of sclerites in the male palp. There may be up to three haematodochae - proximal, middle, and distal - separating the three groups of sclerites in the palp. The haematodochae distend under haemolymph pressure during copulation thus causing the palpal sclerites to separate and rotate.

Homologous — Similar structures (for example mitochondrial genes) having a *common* origin.

Homoplasy	Structural resemblance resulting from parallelism, reversal, or convergent evolution rather than common ancestry (for example an identical base change at a position in a DNA sequence in two taxa for which the common ancestor possesses a different base at this position).
Labium	The lip, ventral to the mouth opening, between the maxillae and attached to the anterior border of the sternum.
Lateral view	Viewed from the side (referring to a bilaterally symmetrical structure).
Maxillae	The mouthparts ventral to the mouth opening and lateral to the labium. Formed from the modified coxae of the palps. Also known as 'endites'.
Mesal view	Viewed from the inside (usually referring to a paired asymmetrical structure).
Metatarsus	The sixth segment of the leg, counting from the proximal end; absent in the palps.
Monophyletic group	1. A group of taxa descended from a single ancestral species. 2. The ancestral species and all descendent species.
Mygalomorphae (= orthognatha)	A sub-order of the Araneae containing families with non-opposing (paraxial) articulation of the chelicerae: Theraphosidae ('tarantulas'), Ctenizidae, and Atypidae.
Opithosoma	The posterior of the two major bodily divisions of a spider; used interchangeably with abdomen.
Palp (or palpus)	The second appendage of the prosoma, originating behind the chelicerae but in front of the legs; its coxa also forms the maxilla;it lacks a metatarsal segment. In adult males it is modified, often greatly, for sperm transfer. More correctly (but rarely by arachnologists) referred to as the pedipalp.

Palpal bulb (or genital bulb) — The structures of the male palpal organ that arise from the alveolus of the cymbium. In the most complete form there are three groups of sclerites separated by three haematodochae in the following order, beginning at the attachment to the cymbium: proximal haematodocha; subtegulum; middle haematodocha; tegulum and median apophysis; distal haematodocha; apical division (embolus and conductor). Many of these structures are frequently fused, reduced, or absent.

Parapatric* — Geographically contiguous populations; as in *parapatric speciation.*

Paraphyletic group — A group of taxa derived from a single ancestral taxon but one which does not contain all the descendents of the most recent ancestor; a category based upon the common possession of plesiomorphic characters.

Pheromone — A chemical, secreted by an animal in minute amounts, which effects a behavioural response in another animal, frequently the opposite sex of the same species, or other members of a social community.

Plesiomorphy — A relative term implying an ancestral or primitive character state. (*cf.* Apomorphy).

Polyphyletic group — A group determined by non-homologous similarities or characters which therefore does not share a single ancestor.

Prosoma — The anterior of the two major bodily divisions of a spider; used interchangeably with cephalothorax.

Receptacle (or copulatory duct/pore) — The external openings and ducts in the female genitalia through which the male inseminates the female.

saeva — (L: *saevus*) fierce or cruel.

Sclerite — A discrete sclerotized (hardened chitin) structure. of particular shape. Several such structures may be interconnected by flexible membranes for example in the male palps, or female epigyne.

Spermatheca — An internal-sac in female spiders used for the reception and storage of spermatozoa.

Spinnerets (or spinners)	The paired appendages at the posterior end of the abdomen, from which silk is extruded through the spigots.
Sternum	The heart-shaped shield of exoskeleton covering the ventral surface of the prosoma, lying posterior to the labium and between the leg coxae.
Sympatric*	Geographically overlapping populations; as in *sympatric speciation.*
Symplesiomorphy	A relatively ancestral character shared among taxa.
Synanthropic	Living close to human habitation.
Synapomorphy	A relatively derived homologous character shared among taxa.
Syntopic	Two or more species occurring in the same locality or habitat.
Tarsus	The most distal segment of a leg or palp.
Taxon	Any taxonomic unit (for example family or species) whether named or not.
Tibia	The fifth segment of the leg or palp counting from the proximal end.
Tegenaria	(L: *teges*) a mat; (L: *arium*) a place.
Tegulum	A scerite forming, along with the median apophysis, the middle of the three divisions of the male palpal bulb; often a broad ring-like structure. In the *Tegenaria atrica* group the tegulum is fused with the conductor.
Ventral view	Viewed from below.

References

AGNARSSON, I. 1996. *Íslenskar Könulœr.* In: Ólafsson, E. (ed) *Fjölrit Náttúrufrœðistofnunar*, **31**, pp. 48-49. Náttúrufræðistofnun Íslands. Reykjavík.

AIROLDI, J. -P. AND FLURY, B. K. 1988. An application of common principle component analysis to cranial morphometry of *Mictrous californicus* and *M. ochrogaster* (Mammalia, Rodentia). *J. Zool., Lond.*, **216**, 21-36.

ANDERSSON, M. B. 1994. *Sexual Selection.* Princeton University Press, Princeton.

ARNOLD, M. L. 1997. *Natural Hybridization and Evolution.* Oxford University Press, Oxford.

ASSMANN, T. AND WEBER, F. 1997. On the allozyme differentiation of *Carabus punctatoauratus* Germar (Coleoptera, Carabidae). *J. Zool. Syst. Evol. Res.*, **35**, 33-43.

AVISE, J. C. 1991. Matriarchal liberation. *Nature*, **352**, 192.

AVISE, J. C. 1994. *Molecular Markers, Natural History and Evolution.* Chapman and Hall, New York.

AVISE, J. C., ARNOLD, J., BALL, R. M., BERMINGHAM, E., LAMB, T., NEIGEL, J. E., REEB, C. A. AND SAUNDERS, N. C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* **18**, 489-522.

AVISE, J. C., NEIGEL, J. E. AND ARNOLD, J. 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Mol. Evol.*, **20**, 99-105.

AYALA, F. J. 1983. Enzymes as taxonomic characters. In: Oxford, G. S. and Rollinson, D. (eds) *Protein Polymorphism: Adaptive and Taxonomic Significance. Systematics Association Special Volume No. 24*, pp. 3-26. Academic Press, London.

BARRIENTOS, J. A. AND RIBERA, C. 1988. Algunas reflexiones sobre las especies del grupo "*atrica*", en la Península Ibérica (Araneae, Agelenidae, *Tegenaria*). *Rev. Arachnol.*, **7** (4), 141-162.

BARTON, N. H. AND HEWITT, G. M. 1985. Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.*, **16**, 113-148.

BARTON, N. H. AND JONES, J. S. 1983. Mitochondrial DNA: new clues about evolution. *Nature*, **306**, 317-318.

BAS. 1989. Glossary. In: *The British Arachnological Society Members Handbook.* (BAS, occasional private publication).

BAYLIS, J. R. 1976. A quantitative study of long-term courtship: II. A comparative study of the dynamics of courtship in two New World cichlid fishes. *Behaviour* LIX, **141**, 117-161.

BILTON, D., MIROL, P., MASCHERETTI, S., FREDGA, K. AND SEARLE, J. B. 1998. The mediterranean regions of Europe as an area of endemism for small mammals rather than a major source in the post-glacial colonisation of central and northern Europe. *Proc. R. Soc. Lond. B*, (in press).

BOULANGER, P., LEBORGNE, R. AND KRAFFT, B. 1986. Barrière éthologique et complémentarité des signaux chez deux espèces sympatriques de Tegenaires (Araneae, Agelenidae). *Biol. Behav.*, **11**, 145-156.

BOYCE, T. M., ZWICK, M. E. AND AQUADRO, C. F. 1989. Mitochondrial DNA in the bark weevils: size, structure and heteroplasmy. *Genetics*, **123**, 825-836.

BRIGNOLI, P. M. 1978. Quelques notes sur les Agenidae, Hahniidae, Oxyopidae et Pisauridae de France et d'Espagne. *Rev. Suisse Zool.*, **85**, 265-294.

BRISTOWE, W. S. 1958. *The World of Spiders*. Collins, London.

BROWER, A. V. Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl. Acad. Sci. USA*, **91**, 6491-6495.

BROWN, W. M. 1985. The mitochondrial genome of animals. In: MacIntyre, R. J. (ed) *Evolutionary Genetics*, pp. 95-130. Plenum Press, New York.

BROWN, W. M., GEORGE, M., JR. AND WILSON, A. C. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA*, **76**, 1967-1971.

BUI, E. T. N., BRADLEY, P. J. AND JOHNSON, P. J. 1996. A common evolutionary origin for mitochondria and hydrogenosomes. *Proc. Natl. Acad. Sci. USA*, **93**, 9651-9656.

BUTLIN, R. K. 1987. Speciation by reinforcement. *Trends Ecol. Evol.*, **2**, 8-13.

BUTLIN, R. K. 1989. Reinforcement of premating isolation. In: Otte, D. and Endler, J. A. (eds) *Speciation and its Consequences*, pp. 158-179. Sinauer Associates Inc., Sunderland, MA.

CAIN, A. J. 1983. Concluding remarks. In: Oxford, G. S. and Rollinson, D. (eds) *Protein Polymorphism: Adaptive and Taxonomic Significance. Systematics Association Special Volume No. 24*, pp. 391-397. Academic Press, London.

CAVALLI-SFORZA, L. L. AND EDWARDS, A. W. F. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution*, **21**, 550-570.

CLARIDGE, M. F., DAWAH, H. A. AND WILSON, M. R. 1997. Practical approaches to species concepts for living organisms. In: Claridge, M. F., Dawah, H. A. and Wilson, M. R. (eds) *The Units of Biodiversity*, pp 1-15. Chapman and Hall, London.

CLARK, D. L. 1994. Sequence analysis of courtship behaviour in the dimorphic jumping spider *Maevia inclemens* (Araneae, Salticidae). *J. Arachnol.*, **22**, 94-107.

CLARY, D. O. AND WOLSTENHOLME, D. R. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *J. Mol. Evol.*, **22**, 252-271.

COYLE, F. A. 1985. Two-year life cycle and low palpal character variance in a Great Smoky Mountain population of the lampshade spider (Araneae, Hypochilidae, *Hypochilus*). *J. Arachnol.*, **13**, 211-218.

COYNE, J. A. AND ORR, H. A. 1997. Patterns of speciation in *Drosophila* revisited. *Evolution*, **51**, 295-303.

COYNE, J. A. AND ORR, H. A. 1998. The evolutionary genetics of speciation. *Phil. Trans. R. Soc. Lond. B*, **353**, 287-305.

CRACRAFT, J. 1989. Speciation and its ontology: the empiracal consequences of alternative species concepts for understanding patterns and processes of differentiation. In: Otte, D. and Endler, J. A. (eds) *Speciation and its Consequences*, pp. 3-27. Sinauer Associates Inc., Sunderland, MA.

CRAWFORD, R. AND LOCKET, G. H. 1976. The occurrence of *Tegenaria gigantea* Chamberlain and Ivie (1935) in North America and Western Europe. *Bull. Br. arachnol. Soc.*, **3**, 199.

CROZIER, R. H. 1993. Molecular methods for insect phylogenetics. In: Oakeshott, J. and Whitten, M. J. (eds.) *Molecular Approaches to Fundamental and Applied Entomology*, pp. 164-221. Springer Series in Experimental Entomology. Springer-Verlag, New York.

DEN HOLLANDER, J., DIJKSTRA, H., ALLEMAN, H. AND VLIJM, L. 1973. Courtship behaviour as a species barrier in the *Pardosa pullata* group (Araneae, Lycosidae). *Tijd. Entomol.*, **116** (1), 1-22.

DEWALT, T. S., SUDMAN, P. D., HAFNER, M. S. AND DAVIS, S. K. 1993. Phylogenetic relationships of pocket gophers (*Cratogeomys* and *Pappogeomys*) based on mitochondrial DNA cytochrome b sequences. *Mol. Phylogenet. Evol.*, **2**(3), 193-204.

DINGLE, H. 1969. A statistical and information analysis of aggressive communication in the mantis shrimp *Gonodactylus bredini* Manning. *Anim. Behav.*, **17**, 561-575.

DOBZHANSKY, T. 1935. A critique of the species concept in biology. *Phil. Sci.*, **2**, 344-345.

DOBZHANSKY, T. 1970. *Genetics of the Evolutionary Process.* Columbia University Press, New York.

EBERHARD, W. G. 1985. *Sexual Selection and the Evolution of Animal Genitalia.* Harvard University Press, Cambridge, MA.

EFRON, B. 1979. Bootstrap methods: another look at the jackknife. *Ann. Statist.*, **7**, 1-26.

ELGAR, M. A. 1995. The duration of copulation in spiders: comparative patterns. *Rec. Western Australian Mus.*, **52**, 1-11.

ELGAR, M. A. 1997. Sperm competition and sexual selection in spiders and other arachnids. In: Birkhead, T. R. and Møller, A. P. (eds) *Sperm Competition and Sexual Selection*. Academic Press, London.

ELLIOTT, N. C., WOLFF, R. J., JOHNSON, W. E., AND ENGEMANN, J. G. 1982. Genetic differentiation among three species of *Pardosa* (Arachnida: Lycosidae), *The Great Lakes Entomologist*, **15** (1), 55-58.

ENDLER, J. A. 1982. Pleistocene forest refuges: fact or fancy? In: Prance, G. T. (ed) *Biological Diversification in the Tropics. Proc. 5th int. Symp. Ass. Trop. Biol.*, pp. 641-657. Columbia University Press, New York.

FAHEY, B. F. AND ELGAR, M. A. 1997. Sexual cohabitation as mate-guarding in the leaf-curling spider *Phonognatha graeffei* Keyserling (Araneoidea, Araneae). *Behav. Ecol. Sociobiol.*, **40**, 127-133.

FARRIS, J. S. 1989. The retention index and the rescaled consistency index. *Cladistics*, **5**, 417-419.

FELSENSTEIN, J. 1981a. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution*, **35**, 124-138.

FELSENSTEIN, J. 1981b. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.*, **17**, 368-376.

FELSENSTEIN, J. 1983. Parsimony in systematics: biological and statistical issues. *Annu. Rev. Ecol. Syst.*, **14**, 313-333.

FELSENSTEIN, J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annu. Rev. Genet.*, **22**, 521-565.

FELSENSTEIN, J. 1995. *PHYLIP (Phylogeny Inference Package)*, Version 3.57c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

FITCH, W. M. 1971. Toward defining the course of evolution: minimum change for a specified tree topology. *Syst. Zool.*, **20**, 406-416.

FITCH, W. M. AND MARGOLIASH, E. 1967. Construction of phylogenetic trees. *Science*, **155**, 279-284.

FOELIX, R. F. 1996. *Biology of Spiders*. Second Edition. Harvard University Press, Massachusetts and London.

FORSTER, L. 1982. Visual communication in jumping spiders (Salticidae). In: Witt, P. N. and Rovner, J. S. (eds) *Spider Communication: Mechanisms and Ecological Significance*, pp.161-212. Princeton University Press, New Jersey.

FOWLER, K. AND WHITLOCK, M. C. 1994. Fluctuating asymmetry does not increase with moderate inbreeding in *Drosophila melanogaster*. *Heredity*, **73**, 373-376.

FRANCESCOLI, G. AND COSTA, F. G. 1992. Postemergence development in *Lycosa carbonelli* Costa and Capocasale, *L. thorelli* (Keyserling), and their hybrid progeny (Araneae, Lycosidae): a comparative laboratory study. *Can. J. Zool.*, **70**, 380-384.

GILLESPIE, R. G., CROOM, H. B. AND PALUMBI, S. R. 1994. Multiple origins of a spider radiation in Hawaii. *Proc. Natl. Acad. Sci. USA*, **91**, 2290-2294.

GILLESPIE, R. G. AND OXFORD, G. S. 1998. Selection on the colour polymorphism in Hawaiian happy-face spiders: evidence from genetic structure and temporal fluctuations. *Evolution* (in press).

GOLDBERG, T. L. AND RUVOLO, M. 1997. Molecular phylogenies and historical biogeography of east African chimpanzees. *Biol. J. Linn. Soc.*, **61**, 301-324.

GRAY, M. W. 1989a. Origin and evolution of mitochondrial DNA. *Annu. Rev. Cell. Biol.*, **5**, 25-50.

GRAY, M. W. 1989b. The evolutionary origins of organelles. *Trends Genet.*, **5**, 294-299.

GUNNARSSON, B. AND ANDERSSON, A. 1992. Chromosome variation in embryos of a solitary spider, *Pityohyphantes phrygianus*, with skewed sex ratio. *Hereditas*, **117**, 85-91.

GYLLENSTEN, U.B., WHARTON, D., JOSEFSSON, A. AND WILSON, A. C. 1991. Paternal inheritance of mitochondrial DNA in mice. *Nature*, **352**, 255-257.

GYLLENSTEN, U. AND WILSON, A. C. 1987. Interspecific mitochondrial DNA transfer and the colonization of Scandinavia by mice. *Genet. Res. Camb.*, **49**, 25-29.

HAFFER, J. 1982. General aspects of the refuge theory. In: Prance, G. T. (ed) *Biological diversification in the tropics. Proc. 5th int. Symp. Ass. Trop. Biol.*, pp. 6-24. Columbia University Press, New York.

HALE, L. R. AND SINGH, R. S. 1987. Mitochondrial DNA variation and genetic structure in populations of *Drosophila melanogaster. Mol. Biol. Evol.*, **4**, 622-637.

HARRIS, H. 1966. Enzyme polymorphisms in man. *Proc. R. Soc. Lond. B*, **164**, 298-310.

HARRISON, R. G. 1986. Pattern and process in a narrow hybrid zone. *Heredity*, **56**, 337-350.

HARRISON, R. G. 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends Ecol. Evol.* **4**, 6-11.

HARRISON, R. G. 1990. Hybrid zones: windows on the evolutionary process. *Oxford Surv. Evol. Biol.*, **7**, 69-128.

HARRISON, R. G. 1993. Hybrids and hybrid zones: historical perspective. In: Harrison, R. G. (ed) *Hybrid Zones and the Evolutionary Process*, pp. 3-12. Oxford University Press, Oxford.

HARRISON, R. G. AND RAND, D. M. 1989. Mosaic hybrid zones and the nature of species boundaries. In: Otte, D. and Endler, J. A. (eds) *Speciation and its Consequences*, pp. 111-133. Sinauer Associates Inc., Sunderland, MA.

HARRISON, R. G., RAND, D. M. AND WHEELER, W. C. 1987. Mitochondrial DNA variation in field crickets across a narrow hybrid zone. *Mol. Biol. Evol.*, **4**, 144-158.

HEBERT, P. D. N. & BEATON, M. J. 1993. *Methodologies for Allozyme Analysis Using Cellulose Acetate Electrophoresis.* (First Revision). Helena Laboratories, Beaumont, Texas.

HEDIN, M. C. 1997. Molecular phylogenetics at the population/species interface in cave spiders of the Southern Appalachians (Araneae: Nesticidae: *Nesticus*). *Mol. Biol. Evol.*, **14**(3), 309-324.

HEWITT, G. M. 1988. Hybrid zones - natural laboratories for evolutionary studies. *Trends Ecol. Evol.*, **3**, 158-167.

HEWITT, G. M. 1989. The subdivision of species by hybrid zones. In: Otte, D. and Endler, J. A. (eds) *Speciation and its Consequences*, pp. 85-110. Sinauer Associates Inc., Sunderland, MA.

HEWITT, G. M. 1990. Divergence and speciation as viewed from an insect hybrid zone. *Can. J. Zool.*, **68**, 1701-1715.

HEWITT, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.*, **58**, 247-276.

HILLIS, D. M. 1991. Discriminating between phylogenetic signal and random noise in DNA sequences. In: Miyamoto, M. M. and Cracraft, J. (eds) *Phylogenetic Analysis of DNA Sequences*, pp. 278-294. Oxford University Press, New York.

HILLIS, D. M., ALLARD, M. W. AND MIYAMOTO, M. M. 1993. Analysis of DNA sequence data: phylogenetic inference. *Methods Enzymol.*, **224**, 456-487.

HILLIS, D. M. AND HUELSENBECK, J. P. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. *J. Hered.*, **83**(3), 189-195.

HILLIS, D. M. AND MORITZ, C. 1990. *Molecular Systematics.* Sinauer, Sunderland, Massachussetts.

HOSTERT, E. E. 1997. Reinforcement: a new perspective on an old controversy. *Evolution*, **51** (3), 697-702.

HOWARD, D. J. 1993. Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. In: Harrison, R. G. (ed) *Hybrid Zones and the Evolutionary Process*, pp. 46-69. Oxford University Press, Oxford.

HUBER, K. C., HAIDER, T. S., MÜLLER, M. W., HUBER, B. A., SCHWEYEN, R. J. AND BARTH, F. G. 1993. DNA sequence data indicates the polyphyly of the family Ctenidae (Araneae). *J. Arachnol.*, **21**, 194-201.

HUBBY, J. AND LEWONTIN, R. C. 1966. A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics* **54**, 577-594.

HUFF, R. P. AND COYLE, F. A. 1992. Systematics of *Hypochilus sheari* and *Hypochilus coylei*, two southern Appalachian lampshade spiders (Araneae, Hypochilidae). *J. Arachnol.*, **20**, 40-46.

HULL, D. L. 1978. A matter of individuality. *Phil. Sci.*, **45**, 335-360.

HULL, D. L. 1997. The ideal species concept - and why we can't get it. In: Claridge, M. F., Dawah, H. A. and Wilson, M. R. (eds) *The Units of Biodiversity*, pp 357-380. Chapman and Hall, London.

HUNT, W. G. AND SELANDER, R. K. 1973. Biochemical genetics of hybridisation in European house mice. *Heredity*, **31**, 11-33.

INNIS, M. A., GELFAND, D. H., SNINSKY, J. J. AND WHITE, T. J. 1990. *PCR Protocols: A Guide to Methods and Applications*. Academic Press, New York.

JACKSON, R. R. 1979. Comparative studies of *Dictyna* and *Mallos* (Araneae, Dictynidae) II. The relationship between courtship, mating, agression and cannibalism in species with differing types of social organization. *Rev. Arachnol.*, **2** (3), 103-132.

JACKSON, R. R. 1982. The behaviour of communicating in jumping spiders (Salticidae). In: Witt, P. N. and Rovner, J. S. (eds) *Spider Communication: Mechanisms and Ecological Significance*, pp. 15-66. Princeton University Press, New Jersey.

JAMES, F. C. AND MCCULLOCH, C. E. 1990. Multivariate analysis in ecology and systematics: panacea or pandora's box? *Annu. Rev. Ecol. Syst.* **21**, 129-166.

JOHANNESEN, J., BAUMANN, T., SEITZ, A. AND VEITH, M. 1998. The significance of relatedness and gene flow on population genetic structure in the subsocial spider *Eresus cinnaberinus* (Araneae: Eresidae). *Biol. J. Linn. Soc.* (in press).

JONES, R.D. 1983. *The Country Life Guide to Spiders of Britain and Northern Europe*. Hamlyn, London.

JUKES, T. H. AND CANTOR, C. R. 1969. Evolution of protein molecules. In: Munro, H. N. (ed) *Mammalian Protein Metabolism*, pp. 21-132. Academic Press, New York.

KELLY, J. K. AND NOOR, M. A. F. 1996. Speciation by reinforcement: a model derived from studies of *Drosophila*. *Genetics*, **143**, 1485-1497.

KENNETT, J. A. B. AND DALINGWATER, J. E. 1986. A progress report on using aspects of reproduction as criteria for distinguishing two closely related *Tegenaria* species. In: W.G. Eberhard, Y. D. Lubin and B. C. Robinson (eds.), *Proc. 9th int. Congr. Arachnol., Panama 1983*, pp. 139-142. Smithsonian Institution Press, Washington.

KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, **16**, 111-120.

KLUGE, A. G. AND FARRIS, J. S. 1969. Quantitative phyletics and the evolution of the anurans. *Syst. Zool.* **18**, 1-32.

KRAFFT, B. 1978. The recording of vibratory signals performed by spiders during courtship. *Symp. zool. Soc. Lond.* **42**, 59-67.

KRAFFT, B. 1982. The significance and complexity of communication in spiders. In: Witt, P. N. and Rovner, J. S. (eds) *Spider Communication: Mechanisms and Ecological Significance*, pp. 15-66. Princeton University Press, New Jersey.

KRAFFT, B., ROLAND, C., MIELLE, D., LEBORGNE, R. AND HOREL, A. 1978. Contribution a l'étude des signaux intervenant lors du comportement sexuel d'araignées du genre *Amaurobius*, *Coelotes* et *Tegenaria*. *103e Congrès national des sociétés savantes, Nancy, sciences, fasc. III.*, pp. 19-29.

KRÁL, J. 1995. The karyotype studies in central Europe's species of the superfamilies Amaurobidae and Dictynoidea (Araneida). *Proc. 15th European Coll. Arachnol., České Budějovice*, pp. 224 (abstract).

LAKE, J. A. 1987. A rate-dependent technique for analysis of nucleic acid sequences: evolutionary parsimony. *Mol. Biol. Evol.*, **4**, 167-191.

LEBORGNE, R. 1984. Spécificité d'organisation des comportments de cour des mâles de trois espèces d'*Amaurobius* (Araneae, Dictynidae). *Rev. Arachnol.*, **5** (4), 239-245.

LEBORGNE, R. 1986. Le rapprochement des sexes chez *Coelotes terrestris* (Wider) (Araneae, Agelenidae): etude des phénomènes vibratoires transmis par la toile lors de la rencontre des partenaires. *Biol. Behav.*, **11**, 205-216.

LEBORGNE, R. 1989. La communication chez les Araignées: étude des phénomènes vibratoires transmis par les toiles lors du rapprochement des sexes chez quelques *Dictynidae* et *Agelenidae*. Unpublished Ph.D. Thesis, University of Nancy, France.

LEBORGNE, R. AND KRAFFT, B. 1979. Technique d'enregistrement et d'analyse des signaux vibratoires intervenant dans les comportements des araignées sédentaires. *Rev. Arachnol.*, **2** (4), 173-182.

LEBORGNE, R., ROLAND, C., AND HOREL, A. 1980. Quelques aspects de la communication chimique et vibratoire chez certaines Agelenidae. *Proc. 8th int. Arachnol. Congr., Vienna*, pp. 215-220.

LEGENDRE, R. M. M. AND LOPEZ, A. 1974. Variations morphologiques sexuelles des glandes gnathocoxales chez les Araneidae (Araneae). *C. R. Acad. Sc. Paris* **279**, 1769-1771.

LI, W. -H., AND GRAUR, D. 1991. *Fundamentals of molecular evolution.* Sinauer, Sunderland, MA.

LIOU, L. W. AND PRICE, T. D. 1994. Speciation by reinforcement of premating isolation. *Evolution*, **48**, 1451-1459.

LIU, Y-B. AND HAYNES, K. F. 1994. Evolution of behavioral responses to sex pheromone in mutant laboratory colonies of *Trichoplusia ni*. *J. Chem. Ecol.*, **20** (2), 231-238.

LOCKET, G. H. 1975. The identity of Blackwall's *Tegenaria saeva* (Araneae, Agelenidae). *Bull. Br. arachnol. Soc.* **3**, 85-90.

LOCKET, G. H. AND MILLIDGE, A. F. 1951. *British spiders* **1**. Ray Society, London.

LOCKET, G. H. AND MILLIDGE, A. F. 1953. *British spiders* **2**. Ray Society, London.

LOCKET, G. H., MILLIDGE, A. F. AND MERRETT, P. 1974. *British spiders* **3**. Ray Society, London.

LOWE, J. J. AND WALKER, M. J. C. 1984. *Reconstructing quaternary environments*. John Wiley and Sons, New York.

LUBIN, Y. D. AND CROZIER, R. H. 1985. Electrophoretic evidence for population differentiation in a social spider *Achaearanea wau* (Theridiidae). *Ins. Soc.*, **32** (3), 297-304.

LUNT, D. H., IBRAHIM, K. M. AND HEWITT, G. M. 1998. mtDNA phylogeography and postglacial patterns of subdivision in the meadow grasshopper *Chorthippus parallelus*. *Heredity*, **80**, 633-641.

LUNT, D. H., ZHANG, D.-X., SZYMURA, J. M. AND HEWITT, G. M. 1996. The Insect Cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Mol. Biol.*, **5**(3), 153-165.

LYNCH, J. M. AND HAYDEN, T. J. 1995. Genetic influences on cranial form: variation among ranch and feral American mink *Mustela vison* (Mammalia: Mustelidae). *Biol. J. Linn. Soc.*, **55**, 293-307.

MADDISON, W. P. AND MADDISON, D. R. 1992. *MacClade: Analysis of Phylogeny and Character Evolution.* Version 3.0. Sinauer Associates, Sunderland, Massachusetts.

MALLET, J. 1995. A species definition for the modern synthesis. *Trends Ecol. Evol.*, **10**, 294-299.

MANLY, B. F. J. 1994. *Multivariate Statistical Methods. A Primer.* Second edition. Chapman and Hall, London.

MARCHANT, A. D., ARNOLD, M. L. AND WILKINSON, P. 1988. Gene flow across a chromosomal tension zone. I. Relicts of an ancient hybridization. *Heredity*, **61**, 321-328.

MARGULIS, L. 1981. *Symbiosis in Cell Evolution.* W. H. Freeman, San Francisco.

MAURER, R. 1992. *Checkliste der Europäischen Agelenidae nach der Roewerschen Systematik 1954 - unter Berücksichtigung Angrenzender Östlicher Gebiete 1: Listen.* Privately printed.

MAYDEN, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. In: Claridge, M. F., Dawah, H. A. and Wilson, M. R. (eds) *The Units of Biodiversity*, pp 381-424. Chapman and Hall, London.

MAYR, E. 1942. *Sytematics and the Origin of Species.* Columbia University Press, New York.

MAYR, E. 1969. *Principles of Systematic Zoology.* McGraw-Hill: New York.

MAYR, E. 1970. *Populations, Species, and Evolution. An Abridgement of 'Animal Species and Evolution'.* Harvard University Press, Cambridge, MA.

MCMILLAN, W. O., JIGGINS, C. D. AND MALLET, J. 1997. What initiates speciation in passion-vine butterflies? *Proc. Natl. Acad. Sci. USA*, **94**, 8628-8633.

MEDRANO, J. F., AASEN, E. AND SHARROW, L. 1990. DNA extracted from nucleated red blood cells. *Biotechniques*, **8**(1), 43.

MERRETT, P. 1980. Notes on the variation, identification and distribution of British species of the *Tegenaria atrica* group (Araneae, Agelenidae). *Bull. Br. arachnol. Soc.*, **5**, 1-8.

MERRETT, P. 1982. New county records of British spiders. *Bull. Br. arachnol. Soc.*, **5**, 332-336.

MERRETT, P. 1989. Twelve hundred new county records of British spiders. *Bull. Br. arachnol. Soc.*, **8**, 1-4.

MERRETT, P., LOCKET, G. H. AND MILLIDGE, A. F. 1985. A checklist of British spiders. *Bull. Br. arachnol. Soc.*, **6**, 381-403.

MICHAUX, J. R., FILIPPUCCI, M. G., LIBOIS, R. M., FONS, R. AND MATAGNE, R. F. 1996. Biogeography and taxonomy of *Apodemus sylvaticus* (the woodmouse) in the Tyrrhenian region; enzymatic variation and mitochondrial DNA restriction pattern analysis. *Heredity*, **76**, 267-277.

MIELLE, D. 1978. *Contribution a l'Etude du Comportement Predatuer et Des Mécanismes de Tolerance dans le Genre* Tegenaria *(Araneae, Agelenidae).* Unpublished Ph.D. Thesis, University of Nancy, France.

MINDELL, D. P. AND HONEYCUTT, R. L. 1990. Ribosomal RNA in vertebrates: evolution and phylogenetic applications. *Annu. Rev. Ecol. Syst.*, **21**, 541-566.

MINELLI, A. AND FODDAI, D. 1997. The species in terrestrial non-insect invertebrates (earthworms, arachnids, myriapods, woodlice and snails). In: Claridge, M. F., Dawah, H. A. and Wilson, M. R. (eds) *The Units of Biodiversity*, pp 309-324. Chapman and Hall, London.

MIROL, P. M. 1996. *Molecular Phylogenies and Karyotype Evolution in Small Mammals. The Examples of* Sorex araneus *and in Eurasia and* Ctenomys *in South America.* Unpublished D. Phil. thesis, University of York.

MISHLER, B. D. AND DONOGHUE, M. J. 1982. Species concepts: a case for pluralism. *Syst. Zool.*, **31**, 491-503.

MITHEN, R., RAYBOULD, A. F. AND GIAMOUSTARIS, A. 1995. Divergent selection for secondary metabolites between wild populations of *Brassica oleracea* and its implication for plant-herbivore interaction. *Heredity*, **75**, 472-484.

MIYASHITA, T. AND HAYASHI, H. 1996 Volatile chemical cue elicits mating behaviour of cohabiting males of *Nephila clavata* (Araneae, Tetragnathidae). *J. Arachnol.*, **24**, 9-15.

MORITZ, C., DOWLING, T. E. AND BROWN, W. M. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annu. Rev. Ecol. Syst.* **18**, 269-292.

MULLIS, K. AND FALOONA, F. 1987. Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. *Methods Enzymol.*, **155**, 335-350.

MULLIS, K., FALOONA, F., SCHARF, S., SAIKI, R., HORN, G. AND ERLICH, H. 1986. Specific enzymatic amplification of DNA *in vitro:* the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.*, **51**, 263-273.

MURPHY, R. W., SITES, J. W.-JR., BUTH, D. AND HAUFLER, C. H. 1990. Proteins I: isozyme electrophoresis. In: Hillis, D. M. and Moritz, C. (eds) *Molecular Systematics*, pp. 45-126. Sinauer, Sunderland, Massachusetts.

MØLLER, A. P. 1997. Developmental stability and fitness: a review. *Am. Nat.*, **149** (5), 916-932.

NAPPI, A. J. 1965. Notes on the courtship and mating habits of the wolf spider *Lycosa helluo* Walckenaer. *Am. Midl. Nat.*, **74** (2), 368-373.

NEI, M. 1972. Genetic distance between populations. *Am. Nat.*, **106**, 283-292.

NEI, M. 1996. Phylogenetic analysis in molecular evolutionary genetics. *Annu. Rev. Genet.*, **30**, 371-403.

NEI, M., TAJIMA, F. AND TATENO, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. *J. Mol. Evol.*, **19**, 153-170.

NELSON, G. 1989. Species and taxa: systematics and evolution. In: Otte, D. and Endler, J. A. (eds) *Speciation and its Consequences*, pp. 3-27. Sinauer Associates Inc., Sunderland, MA.

NIKOLIC, F. AND POLENEC, A. 1981. *Catalogus Faunae Jugoslaviae*. Ljubljana. [NOT SEEN]

NOOR, M. A. 1995. Speciation driven by natural selection in *Drosophila*. *Nature*, **375**, 674-675.

NORUSIS, M. J./SPSS INC. 1994. *SPSS Professional Statistics Release 6.1*.

ORTA-OCAÑA, J. M., MOYA-LARAÑO, AND BARRIENTOS, J. A. 1996. Precopulatory male ethograms of three species of *Lycosa* Latrielle 1804 (Araneae: Lycosidae) of the Iberian peninsula. *Proc. 13th int. Congr. Arachnol., Geneva, 1995. Rev. Suisse Zool.*, vol. hors série, 515-522.

OXFORD, G. S. 1991. Visible morph-frequency variation in allopatric and sympatric populations of two species of *Enoplognatha* (Araneae: Theridiidae). *Heredity*, **67**, 317-324.

OXFORD, G. S. 1993. Patterns of sperm usage in large house spiders (*Tegenaria* spp.): genetics of esterase markers. *Heredity*, **70**, 413-419.

OXFORD, G. S. AND CHESNEY, H. C. G. 1994. Large house spiders *Tegenaria* Spp. in Northern Ireland: previously overlooked species or recent introductions? *Ir. Nat. J.*, **24** (9), 354-357.

OXFORD, G. S. & CROUCHER, P. J. P. 1997. Gone fishing.....for Tegenaria. *Newsl. Br. arachnol. Soc.*, **78** (March).

OXFORD, G. S. AND PLOWMAN, A. 1991. Do large house spiders *Tegenaria gigantea* and *T. saeva* (Araneae, Agelenidae) hybridise in the wild? - A multivariate approach. *Bull. Br. arachnol. Soc.*, **8** (9), 293-296.

OXFORD, G. S. AND SMITH, C. J. 1987. The distribution of *Tegenaria gigantea* Chamberlain and Ivie, 1935 and *T. saeva* Blackwall, 1844 (Areaneae, Agelenidae) in Yorkshire. *Bull. Br. arachnol. Soc.*, **7** (4), 123-127.

PAGE, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences.* **12**, 357-358.

PALMER, A. R. AND STROBECK, C. 1986. Fluctuating asymmetry: measurement, analysis, patterns. *Annu. Rev. Ecol. Syst.*, **17**, 391-421.

PALMER, A. R. AND STROBECK, C. 1992. Fluctuating asymmetry as a measure of developmental stability: implications of non-normal distributions and power of statistical tests. *Acta Zool. Fennica*, **191**, 57-72.

PARKER, J. R. 1984. Synanthropic spiders and other things - Part 2. *Newsl. Br. arachnol. Soc.*, **39**, 1-2.

PESARINI, C. 1994. Arachnida Araneae. In: Minelli, A., Ruffo, S. and La Posta, S. (eds.) *Checklist delle specie della fauna italiana*, **23**. Calderini, Bologna.

PLATNICK, N. 1971. The evolution of courtship behaviour in spiders. *Bull. Br. arachnol. Soc.*, **2** (3), 40-47.

PORTER, A. H. AND JAKOB, E. M. 1990. Allozyme variation in the introduced spider, *Holocnemus pluchei* (Araneae, Pholcidae) in California. *J. Arachnol.*, **18**, 313-319.

POST, R. J., FLOOK, P. K. AND MILLES, A. L. 1993. Methods for the preservation of insects for DNA studies. *Biochem. Syst. Ecol.*, **21**(1), 85-92.

POWELL, J. R., CACCONE, A., AMATO, G. D. AND YOON, C. 1986. Rates of nucleotide substitutions in *Drosophila* mitochondrial DNA and nuclear DNA are similar. *Proc. Natl. Acad. Sci. USA*, **83**, 9090-9093.

PRAGER, E. M., SAGE, R. D., GYLLENSTEN, U., THOMAS, W. K., HÜBNER, R., JONES, C. S., NOBLE, L. SEARLE, J. B. AND WILSON, A. C. 1993. Mitochondrial DNA sequence diversity and the colonization of Scandinavia by house mice from East Holstein. *Biol. J. Linn. Soc.*, **50**, 85-122.

PRENTER, J., MONTGOMERY, W. I. AND ELWOOD, R. W. 1995. Multivariate morphometrics and sexual dimorphism in the orb-web spider *Metellina segmentata* (Clerck, 1757) (Araneae, Metidae). *Biol. J. Linn. Soc.*, **55**, 345-354.

PROGRAM MANUAL FOR THE WISCONSIN PACKAGE, Version 8, September 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin.

PROMEGA PROTOCOLS AND APPLICATIONS GUIDE, 3rd ed. 1996. Promega Corporation, Madison, WI.

RAMIREZ, M. G. AND BECKWITT, R. D. 1995. Phylogeny and historical biogeography of the spider genus *Lutica* (Araneae, Zodariidae). *J. Arachnol.*, **23**, 177-193.

RAMIREZ, M. G. AND FANDINO, L. B. 1996. Genetic variability and gene flow in *Metepeira ventura* (Araneae, Araneidae). *J. Arachnol.*, **24**, 1-8.

RAMIREZ, M. G. AND FROEHLIG, J. L. 1997. Minimal genetic variation in a coastal dune arthropod: the trapdoor spider *Aptostichus simus* (Cyrtaucheniidae). *Cons. Biol.*, **11**, 256-259.

RAND, D. M. AND HARRISON, R. G. 1986. Mitochondrial DNA transmission genetics in crickets. *Genetics*, **114**, 955-970.

REISKIND, J. AND CUSHING, P. E. 1996. Study of a narrow hybrid zone between two wolf spiders, *Lycosa ammophila* and *Lycosa ericeticola*, in north Florida (Araneae, Lycosidae). *Proc. 13th int. Congr. Arachnol., Geneva, 1995. Rev. Suisse Zool.*, vol. hors série, 543-554.

RETTIG, J. E., FULLER, R. C., CORBETT, A. L. AND GETTY, T. 1997. Fluctuating asymmetry indicates levels of competition in an even-aged poplar clone. *Oikos*, **80**, 123-127.

REVELL, S. H. 1947. Controlled X-segregation at meiosis in *Tegenaria*. *Heredity*, **1**, 337-347.

REYNOLDS, J. WEIR, B. S. AND COCKERHAM, C. C. 1983. Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics*, **105**, 767-779.

RICHARDSON, B. J., BAVERSTOCK, P. R., & ADAMS, M. 1986. *Allozyme Electrophoresis: A Handbook for Animal Systematics and Population Studies.* Academic Press: London.

RIDLEY, M. 1993. *Evolution.* Blackwell Scientific Publications, London.

ROBERTS, M. J. 1993. *The Spiders of Great Britain and Ireland* **1** (compact edition). Harley Books, Colchester.

ROBERTS, M. J. 1995. *Spiders of Britain and Northern Europe*. Harper-Collins, London.

ROGERS, J. S. 1972. Measures of genetic similarity and genetic distance. *Studies in Genetics VII, University of Texas Publication 7213*, pp. 145-153.

ROVNER, J. S. 1982. Eplilogue. In: Witt, P. N. and Rovner, J. S. (eds) *Spider Communication: Mechanisms and Ecological Significance*, pp. 387-391. Princeton University Press, New Jersey.

SAETRE, G. -P., MOUM, T., BURES, S., KRÁL, M., ADAMJAN, M. AND MORENO, J. 1997. A sexually selected character displacement in flycatchers reinforces premating isolation. *Nature*, **387**, 589-592.

SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. AND ERLICH, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-491.

SAITOU, A. G. AND NEI, M. 1986. The number of nucleotides required to determine the branching order of three species, with special reference to the human-chimpanzee-gorilla divergence. *J. Mol. Evol.*, **24**, 189-204.

SAITOU, A. G. AND NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406-425.

SAMBROOK, J., FRITSCH, E. F. AND MANIATIS, T. 1989. *Molecular Cloning: A Laboratory Manual, 2nd ed.* Cold Spiring Harbour Laboratory Press, Cold Spiring Harbour, NY.

SANGER, F., NICKLEN, S. AND COULSON, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

SCHIKORA, H. -B. 1995. Intraspecific variation in taxonomic characters, and notes on distribution and habitats of *Meioneta mossica* Schikora and *M. saxatilis* (Blackwall), two closely related spiders from northern and central Europe (Araneae: Linyphiidae). *Bull. Br. arachnol. Soc.*, **10**, 65-74.

SEARLE, J. B. 1983. *Robertsonian chromosomal variation in the common shrew* Sorex araneus *L.* Unpublished PhD thesis, University of Aberdeen.

SERVEDIO, M. R. AND KIRKPATRIC, M. 1997. The effects of gene flow on reinforcement. *Evolution*, **51** (6), 1764-1772.

SEUTIN, G., BRAWN, J., RICKLEFS, R. AND BERMINGHAM, E. 1993. Genetic divergence among populations of a tropical passerine, the streaked saltator (*Saltator albicollis*). *Auk*, **110**, 117-126.

SHAW, D. D., MARCHANT, A. D., ARNOLD, M. L. AND CONTRERAS, N. 1988. Chromosomal rearrangments, ribosomal genes and mitochondrial DNA: introgression across a narrow hybrid zone. In: Brandham, P. E. and Bennett, M. D. (eds) *Kew Chromosome Conference III*, pp.121-129. Allen and Unwin, London.

SIMON, C., FRATI, F., BECKENBACH, A., CRESPI, B., LIU, H. AND FLOOK, P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.*, **87**, 651-701.

SLATKIN, M. 1985. Rare alleles as indicators of gene flow. *Evolution*, **39**, 53-65.

SLATKIN, M. AND MADDISON, W. P. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*, **123**, 603-613.

SMITH, C. J. 1985. The distribution of the larger *Tegenaria* spider species (Agelenidae) in the York area. *Newsl. Br. arachnol. Soc.*, **42**, 5-6.

SMITH, D. R. R. 1986. Population genetics of *Anelosimus eximius* (Araneae, Theridiidae). *J. Arachnol.*, **14**, 210-217.

SMITH, D. R. AND ENGEL, M. S. 1994. Population structure in an Indian cooperative spider, *Stegodyphus sarasinorum* Karsch (Eresidae). *J. Arachnol.*, **22**, 108-113.

SOKAL, R. R. AND ROHLF, F. J. 1995. *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd Edition. W. H. Freeman and Company: New York.

STANNERS, D. AND BOURDEAU, P. (eds). 1995. *Europe's Environment: The Dobříš Assessment.* European Environment Agency, Copenhagen.

STEINER, W. W. M. AND GREENSTONE, M. H. 1992. Variation in *Schizocosa* (Araneae: Lycosidae), *Metaphidippus* and *Phidippus* (Araneae: Salticidae). *J. Arachnol.*, **20**, 35-39.

STRATTON, G. E. 1997. Investigation of species divergence and reproductive isolation of *Schizocosa stridulans* (Araneae: Lycosidae) from Illinois. *Bull. Br. arachnol. Soc.*, **10** (9), 313-321.

STRATTON, G. E. AND UETZ, G. W. 1981. Acoustic communication and reproductive isolation in two species of wolf spiders. *Science*, **214**, 575-577.

STRATTON, G. E. AND UETZ, G. W. 1983. Communication via substratum-coupled stridulation and reproductive isolation in wolf spiders (Araneae: Lycosidae). *Anim. Behav.*, **31** (1), 164-172.

STRATTON, G. E. AND UETZ, G. W. 1986. The inheritance of courtship behavior and its role as a reproductive isolating mechanism in two species of *Schizocosa* wolf spiders (Araneae, Lycosidae). *Evolution*, **40** (1), 129-141.

STRATTON, G. E. AND UETZ, G. W. 1987. The inheritance of courtship behavior in *Schizocosa* wolf spiders (Araneae, Lycosidae). In: Huettel, M. D. (ed) *Evolutionary Genetics of Invertebrate Behaviour.* Plenum Publishing, New York.

SULLIVAN, M. S., ROBERTSON, P. A. AND AEBISCHER, N. A. 1993. Fluctuating asymmetry measurement. *Nature*, **361**, 409-410.

SWADDLE, J. P., WITTER, M. S. AND CUTHILL, I. C. 1994. The analysis of fluctuating asymmetry. *Anim. behav.*, **48**, 986-989.

SWOFFORD, D. L. 1993. *PAUP: Phylogenetic Analysis Using Parsimony*, version 3.1.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

SWOFFORD, D. L. AND OLSEN, G. J. 1990. Phylogeny reconstruction. In: Hillis, D. M. and Moritz, C. (eds) *Molecular Systematics*, pp. 411-501. Sinauer, Sunderland, MA.

SWOFFORD, D. L. AND SELANDER, R. B. 1981. *BIOSYS-1. A computer program for the analysis of allelic variation in genetics.* Release 1. University of Illinois, Urbana, IL.

TABERLET, P. AND BOUVET, J. 1994. Mitochondrial DNA polymorphism, phylogeography and conservation genetics of the brown bear *Ursus arctos* in Europe. *Proc. R. Soc. Lond. B*, **255**, 195-200.

TEGELSTRÖM, H. 1987. Transfer of mitochondrial DNA from the northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*). *J. Mol. Evol.*, **24**, 218-227.

TEGESTRÖM, H. AND GELTER, H. P. 1990. Haldane's rule and sex biased gene flow between two hybridizing flycatcher species (*Ficedula albicollis* and *F. hypoleuca*, Aves: Muscicapidae). *Evolution*, **44**, 2012-2021.

TEGELSTRÖM, H., WYONI, P. -I, GELTER, H. AND JAAROLA, M. 1988. Concordant divergence in proteins and mitochondrial DNA between two vole species in the genus *Clethrionomys. Biochem. Genet.*, **26**, 223-237.

TEMPLETON, A. R. 1989. The meaning of species and speciation: a genetic perspective. In: Otte, D. and Endler, J. A. (eds) *Speciation and its Consequences*, pp. 3-27. Sinauer Associates Inc., Sunderland, MA.

TERRANOVA, A. C. AND ROACH, S. H. 1987a. Electrophoretic key for distinguishing South Carolina species of the genus *Phidippus* (Araneae: Salticidae) as spiderlings and adults. *Ann. Entomol. Soc. Am.*, **80** (3), 346-352.

TERRANOVA, A. C. AND ROACH, S. H. 1987b. Genetic differentiation in the genus *Phidippus* (Araneae, Salticidae). *J. Arachnol.*, **14**, 385-391.

THOMPSON, J. D., HIGGINS, D. G. AND GIBSON, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl. Acids Res.*, **22**, 4673-4680.

THORPE, R. S. 1988. Multiple group principle component analysis and population differentiation. *J. Zool., Lond.* **216**, 37-40.

TIETJEN, W. J. AND ROVNER, J. S. 1982. Chemical communication in lycosids and other spiders. In: Witt, P. N. and Rovner, J. S. (eds) *Spider Communication: Mechanisms and Ecological Significance*. pp.249-279. Princeton University Press, New Jersey.

TRICKETT, A. J. AND BUTLIN, R. K. 1994. Recombination suppressors and the evolution of new species. *Heredity*, **73**, 339-345.

TURNER, J. R. G. 1982. How do refuges produce biological diversity? Allopatry and parapatry, extinction and gene flow in mimetic butterflies. In: Prance, G. T. (ed) *Biological Diversification in the Tropics. Proc. 5th int. Symp. Ass. Trop. Biol.*, pp. 309-335. Columbia University Press, New York.

UETZ, G. W. AND STRATTON, G. E. 1982. Acoustic communication and reproductive isolation in spiders. In: Witt, P. N. and Rovner, J. S. (eds) *Spider Communication: Mechanisms and Ecological Significance*, pp. 123-159. Princeton University Press, New Jersey.

VANLERBERGHE, F., BOURSOT, P., NIELSEN, J. T., BONHOMME, F. 1988. A steep cline for mitochondrial DNA in Danish mice. *Genet. Res. Camb.*, **52**, 185-193.

VENANZETTI, F., CESARONI, D., MARIOTTINI, P. AND SBORDONI, V. 1993. Molecular phylogenies in *Dolichopoda* cave crickets and mtDNA rate calibration. *Mol. Phylogenet. Evol.*, **2**, 275-280.

VLIJM, L. AND DIJKSTRA, H. 1966. Comparative research of the courtship behaviour in the genus *Pardosa* (Arach., Araneae). *Senck. Biol.* **47** (1), 51-55.

WAKELEY, J. 1996. The excess of transitions among nucleotide substitutions: new methods of estimating transition bias underscore its significance. *Trends Ecol. Evol.*, **11**, 158-163.

WARE, A. D. AND OPELL, B. D. 1989. A test of the mechanical isolation hypothesis in two similar spider species. *J. Arachnol.*, **17**, 149-162.

WEBB, T. AND BARTLEIN, P. J. 1992. Global changes during the last 3 million years: climatic controls and biotic responses. *Annu. Rev. Ecol. Syst.*, **23**, 141-173.

WHITE, E. B. 1952. *Charlotte's Web.* Puffin Books, London.

WILCOX, T. P., HUGG, L., ZEH, J. A. AND ZEH, D. W. 1997. Mitochondrial DNA sequencing reveals extreme genetic differentiation in a cryptic species complex of neotropical pseudoscorpions. *Mol. Phylogenet. Evol.* **7**(2), 208-216.

WINDIG, J. J. 1998. Evolutionary genetics of fluctuating asymmetry in the peacock butterfly (*Inachis io*). *Heredity*, **80**, 382-392.

WISE, D. H. 1993. Spiders in Ecological Webs. Cambridge University Press.

WITT, P. N. 1975. The web as a means of communication. *Biosci. Commun.* **1**, 7-23.

WRIGHT, S. 1951. The genetical structure of populations. *Ann. Eugen.*, **15**, 323-354.

WRIGHT, S. 1978. *Evolution and Genetics of Populations, vol. 4. Variability Within and Among Natural Populations*. University of Chicago Press, Chicago.

ZHANG, D. -X. AND HEWITT, G. M. 1996. Nuclear integrations: challenges for mitochondrial DNA markers. *Trends Ecol. Evol.*, **11**, 247-251.