Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 47 (2008) 903–915
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Grafting the molecular phylogenetic tree with morphological branches to reconstruct the evolutionary history of the genus *Zaprionus* (Diptera: Drosophilidae)

Amir Yassin [a,b,*], Luciana O. Araripe [c], Pierre Capy [a,b], Jean-Luc Da Lage [a,b], Louis Bernard Klaczko [c], Claude Maisonhaute [a,b], David Ogereau [a,b], Jean R. David [a,b]

[a] *Laboratoire Evolution, Génomes et Spéciation (LEGS), UPR 9034, bât. 13, CNRS, 91198 Gif-sur-Yvette Cedex, France*
[b] *Université Paris-Sud 11, 91405 Orsay Cedex, France*
[c] *Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas, Unicamp, CP 6109, 13083-9790 Campinas, SP, Brazil*

Received 6 April 2007; revised 13 November 2007; accepted 22 January 2008
Available online 21 February 2008

Abstract

A molecular phylogeny for the drosophilid genus *Zaprionus* was inferred using a mitochondrial (*CO-II*) and a nuclear (*Amyrel*) gene using 22 available species. The combined molecular tree does not support the current classification, dubbed phylogenetic, based entirely upon a morphocline of forefemoral ornamentation. For species for which DNA was not available, phylogenetic positioning was only assigned using morphological characters. In order to avoid conflict between DNA and morphology in the combined analyses (supermatrix method), we developed a new method in which few morphological characters were sampled according to an *a priori* homoplasy assessment on the consensus molecular tree. At each internal node of the tree, a number of synapomorphies was determined, and species with no molecular sequences were grafted thereon. Analogously to tree vocabulary, we called our method 'morphological grafting'. New species groups and complexes were then defined in the light of our findings. Further, divergence times were estimated under a relaxed molecular clock, and historical biogeography was reconstructed under a maximum likelihood model. *Zaprionus* appears to be of recent origin in the Oriental region during the Late Miocene (~10 MYA), and colonization of Africa started shortly after (~7 MYA) via the maritime route of the Indian Ocean Islands. Most of the morphological and ecological diversification took place, later, in Western Africa during the Quaternary cyclic climatic changes. Furthermore, some species became recent invaders, with one, *Zaprionus indianus*, has successfully invaded South and North America during the last decade.

Keywords: Africa paleobiogeography; *Amyrel*; *CO-II*; Drosophilidae; Morphological grafting; Supermatrix; Supertree; *Zaprionus*

1. Introduction

Zaprionus Coquillett, 1901 is a drosophilid genus characterized by the presence of longitudinal white stripes on the frons and the mesonotum. It contains 55 described species, divided into two geographically disjunctive subgenera: the subgenera *Zaprionus s.s.* (44 spp.) and *Anaprionus* (11 spp.) in the Afrotropical and Oriental biogeographic regions, respectively (Okada and Carson, 1983). In contrast to Oriental *Zaprionus*, Afrotropical species are very common and abundant in drosophilid communities. Tsacas et al. (1981) noted that it "is rare not to find one or more of these species in a trap put anywhere in Africa".

Chassagnard and Tsacas (1993) proposed a phylogenetic classification (hereafter CT93) of the subgenus *Zaprionus s.s.* based on a 'morphocline' of a single morphological character: the forefemoral ornamentation. According to CT93, species are classified into two groups: the *inermis* group containing species lacking stout spines on the

* Corresponding author. Address: Laboratoire Evolution, Génomes et Spéciation (LEGS), CNRS, Av. de la Terrasse, 91198 Gif-sur-Yvette Cedex, France. Fax: +33 1 69 82 37 36.
E-mail address: yassin@legs.cnrs-gif.fr (A. Yassin).

1055-7903/$ - see front matter © 2008 Published by Elsevier Inc.
doi:10.1016/j.ympev.2008.01.036

ventromedial margin of the forefemora, a character that identifies the two sexes in the other group, *armatus*. They have also divided the *armatus* species group to further three subgroups, also based on the same morphocline of forefemoral ornamentation: the *armatus* subgroup (with a series of simple spines), the *tuberculatus* (with one long simple spine borne on a salient wart) and the *vittiger* (with a series of spines, each with a stout bristle at its base).

In the present paper, a molecular phylogeny was inferred from the combined analysis of one mitochondrial (*CO-II*) and one nuclear (*Amyrel*) genes. This has been conducted for *Zaprionus s.s.* species for which DNA was available, i.e., 21 species, cultured or cryopreserved in Gif-sur-Yvette. Although the molecular phylogeny was in disagreement with CT93, it could not be used to systematically revise the subgenus due to incomplete taxon sampling. Nearly half of the species are available only as old, pinned museum specimens for which high-quality DNA can not be extractable. The use of morphological characters would hence appear inevitable.

Two questions may arise in conducting such an analysis. First, what is the precise role of morphology in phylogeny reconstruction in the post-genomic era? Morphological phylogenetics has been recently criticized (e.g., Baker and Gatesy, 2002; Scotland et al., 2003; Olmstead and Scotland, 2005). Most of the limitations result from character conceptualization, coding and *a posteriori* homoplasy assessment. Nonetheless, morphological characters are still relevant in cases where DNA sequences could not be obtained (e.g., fossils or old museum specimens; Jenner, 2004; Smith and Turner, 2005; Lee, 2006). This leads to the second question: how to compile molecular sequences and morphological traits to build an exhaustive phylogenetic tree when DNA could not be obtained from all the taxa within the studied group? The traditional approach, known as the 'combined analysis' or 'supermatrix method' (de Quieroz and Gatesy, 2007), involves the concatenation and the simultaneous analysis of separate character data sets. However, Wortley and Scotland (2006) reviewed the effect of combining morphological and molecular data and concluded that morphology usually does not increase neither the accuracy nor the support of the combined tree. Incongruence between DNA and morphology has also been shown in several phylogenetic studies of the Drosophilidae (e.g., DeSalle and Grimaldi, 1991; Kwiatowski and Ayala, 1999; Durando et al., 2000; Remsen and O'Grady, 2002). An alternative to the 'supermatrix method' is the 'supertree method' in which several data sets are analyzed separately, and then the tree derived from the independent analyses are used to produce a single, joint estimate of phylogeny (Bininda-Emonds, 2004). Unlike supermatrices, supertrees do not assume that all characters have experienced the same branching history (Crandall and Buhay, 2004). However, they have been criticized by loss of contact with the primary character data, which makes supertrees invalid as phylogenetic hypotheses. An intermediate approach combining the strengths of the two methods was thus needed.

The aims of this paper were: (1) to propose a novel approach in reconstructing large phylogenies from molecular and morphological data when morphology matters, an approach that we called 'morphological grafting'; (2) to revise the CT93 phylogeny of the subgenus *Zaprionus s.s.* with suggestions of new 'natural' taxonomic groupings; and (3) to infer the historical biogeography and the relationship between the Oriental subgenus *Anaprionus* and the Afrotropical subgenus *Zaprionus s.s.* in light of the new phylogeny.

2. Materials and methods

2.1. Sampling of species

Table 1 shows the list of species used in this study. A taxonomic hierarchy of species has been used as outgroup in relation to *Zaprionus s.s.* ingroup species: (1) two species of the Oriental subgenus *Anaprionus* (*Zaprionus multistriatus* and *Zaprionus dalagei* n. sp.) to cover the genus limit; (2) one *Samoaia* species (*Samoaia leonensis*) to cover the *Zaprionus* genus group limit (Grimaldi, 1990); (3) two *Drosophila* species (*Drosophila immigrans* and *Drosophila repletoides*) known to be related to *Zaprionus* genus group (Throckmorton, 1975; Da Lage et al., 2007); and (4) two *Scaptodrosophila* species, a basal group to the genus *Drosophila* (Robe et al., 2005; Da Lage et al., 2007). The simultaneous use of outgroup hierarchy enables testing the effect of long-branch attraction on tree topology (Sanderson and Shaffer, 2002). For the subgenus *Zaprionus s.s.*, 39 out of the 44 described species were included in this study. The remaining described species (*Zaprionus arduus* Collart, 1937; *Zaprionus badyi* Burla, 1954; *Zaprionus momorticus* Garber, 1957; *Zaprionus neglectus* Collart, 1937 and *Zaprionus niabu* Burla, 1954) all belonging to the *inermis* group, were not included in this study due to the lack of examined material or detailed, illustrated descriptions in the literature. In addition, three new species, *Zaprionus lachaisei* n. sp., *Zaprionus nigranus* n. sp. and *Zaprionus santomensis* n. sp. were added. All of these new species belong to the *vittiger* subgroup, and their description will be given in a future work.

2.2. Molecular analysis

DNA was extracted from single freshly killed or frozen flies, either using the QIAgen DNA extraction kit or following the method of Gloor and Engels (1992). *CO-II* and *Amyrel* primers are listed in Table 2, and their PCR amplification conditions were followed as described in Robe et al. (2005) and Da Lage et al. (2007), respectively. To reduce possible ambiguities in reading the sequence, single regions were generated either two times in one direction (one primer) or one time in both directions (using two primers). This has resulted in a 2164-bp-long sequence: 688 bp for *CO-II* and 1476 bp for *Amyrel*.

Table 1
List of species used in this study

Species	*CO-II*	*Amyrel*
Genus *Zaprionus*		
Subgenus *Anaprionus*		
Z. multistriatus Sturtevant, 1927	EF453720	AY736516[e]
Z. dalagei n. sp.	EU161099	AY736521[e]
Subgenus *Zaprionus* s.s.		
Group *armatus*		
Subgroup *armatus*		
Z. armatus Collart, 1937	—	—
Z. campestris Chassagnard, 1989	—	—
Z. enoplomerus Chassagnard, 1989	—	—
Z. fumipennis Séguy, 1938	—	—
Z. hoplophorus Tsacas and Chassagnard, 1990	—	—
Z. montanus Collart, 1937	—	—
Z. seguyi Tsacas and Chassagnard, 1990	—	—
Z. serratus Chassagnard, 1990	—	—
Z. spineus Tsacas and Chassagnard, 1990	—	—
Z. spinipes Tsacas and Chassagnard, 1990	—	—
Z. spinoarmatus Tsacas and Chassagnard, 1990	—	—
Z. spinosus Collart, 1937	—	—
Z. tuberarmatus Tsacas and Chassagnard, 1990	—	—
Z. vrydaghi Collart, 1937	—	—
Subgroup *tuberculatus*		
Z. mascariensis Tsacas and David, 1975	EF453714	AY736522[e]
Z. sepsoides Duda, 1939	EF453712	AY736523[e]
Z. tuberculatus Malloch, 1932	EF453719	AY736524[e]
Subgroup *vittiger*		
Z. beninensis Chassagnard and Tsacas, 1993	EF453700	EF458331
Z. camerounensis Chassagnard and Tsacas, 1993	EF453699	EF458332
Z. capensis Chassagnard and Tsacas, 1993	EF453705	EF458326
Z. davidi Chassagnard and Tsacas, 1993	EF453708	EF458323
Z. indianus Gupta, 1970	EF453709	EF458322
Z. koroleu Burla, 1954	—	—
Z. lachaisei n. sp.	EF453701	EF458321
Z. megalorchis Chassagnard and Tsacas, 1993	EF453710	EF458330
Z. multivittiger Chassagnard, 1996	—	—
Z. nigranus n. sp.	EF453698	EF458333
Z. ornatus Séguy, 1933	—	—
Z. proximus Collart, 1937	—	—
Z. taronus Chassagnard and Tsacas,1993	EF453707	EF458324
Z. santomensis n.sp.	EF453703	EF458328
Z. spinipilus Chassagnard and McEvey, 1992	EF453702	EF458329
Z. vittiger Coquillett, 1902	EF453704	EF458327
Group *inermis*		
Z. cercus Chassagnard and McEvey, 1992	EF453715	AY736517[e]
Z. ghesquierei Collart, 1937	EF453717	AY736518[e]
Z. inermis Collart, 1937	EF453716	AY736519[e]
Z. kolodkinae Chssagnard and Tsacas, 1987	EF453713	AY736520[e]
Z. litos Chassagnard and McEvey, 1992	—	—
Z. sexstriatus Chassagnard, 1996	EF453718	EF458320
Z. sexvittatus Collart, 1937	—	—
Z. simplex Chassagnard and McEvey, 1992	—	—
Z. verruca Chassagnard and McEvey, 1992	EF453711	AY736525[e]
Genus *Samoaia*		
S. leonensis	AF478438[c]	EU161100
Genus *Drosophila*		
Subgenus *Drosophila* s.s.		
Group *immigrans*		
D. immigrans Sturtevant, 1921	AF478424[c]	AF491632[b]
Group *tumiditarsus*		
D. repletoides Hsu, 1943	EU161098	AY736500[e]
Subgenus *Sophophora*		
Group *melanogaster*		

(*continued on next page*)

Table 1 (*continued*)

Species	*CO-II*	*Amyrel*
D. melanogaster Meigen, 1830	U37541[a]	AF022713[b]
Genus *Scaptodrosophila*		
Sc. latifasciaeformis (Duda, 1940)	AY847765[d]	—
Sc. finitima (Lamb, 1914)	—	AY736527[e]

Zaprionus classification follows CT93 (see text). GenBank Accession numbers for *CO-II* and *Amyrel* sequences are presented. Species with no accession numbers are those that were used only in morphological analysis.
Sequences obtained from previous studies: [a]de Bruijn (1983), [b]Da Lage et al. (1998), [c]Remsen and O'Grady (2002), [d]Robe et al. (2005), [e]Da Lage et al. (2007).

Table 2
Forward (F) and reverse (R) sequences for primers used in this study

Primer	Sequence
CO-II[a]	
F: TL2-J-3037	5′-ATGGCAGATTAGTGCAATGG-3′
R: TK-N-3785	5′-GTTTAAGAGACCAGTACTTG-3′
Amyrel[b]	
F: RELUDIR	5′-TGGATGCNGCCAAGCACATGGC-3′
R: RELAVBIS	5′-GCATTTGTACCGTTTGTGTCGTTATCG-3′
F: ZONE2BIS	5′-GTAAATNGGNNCCACGCGAAG-3′
R: RELREV+	5′-GTTCCCCAGCTCTGCAGCC-3′

[a] From Simon et al. (1994).
[b] From Da Lage et al. (2007).

2.3. Phylogenetic analysis

Nucleotide sequences were viewed and manually edited using MEGA ver. 3.1 (Kumar et al., 2004) and, then, aligned with ClustalW (Thompson et al., 1994) using the MEGA default parameters. MEGA was also used to reconstruct phylogenetic trees using two methods: (1) neighbor-joining (NJ) (Saitou and Nei, 1987) under Tamura and Nei (1993) model to correct for multiple hits; (2) maximum parsimony (MP) using a close-neighbor interchange (CNI) method (Nei and Kumar, 2000) with one search level and 10 repeats of random sequences addition. The maximum likelihood (ML) and the Bayesian inference (BI) trees were reconstructed using PHYML (Guindon and Pascual, 2003) and MrBayes ver. 3.1.1 (Ronquist and Huelsenbeck, 2003) programs, respectively. ML and BI reconstructions used the model proposed by the FindModel program (Tao, 2005) for the two genes: GTR + Γ (Tavaré, 1986). In all tree reconstruction methods, confidence level for monophyly was determined by the 50% cutoff of 1000 bootstrap replications (NJ, MP and ML) or posterior probability (BI) estimated after a run of 500,000 generations.

2.4. Morphological analysis (and grafting)

Scotland et al. (2003) suggested that the use of fewer, rigorous and critical morphological characters in the context of a molecular phylogeny "is preferable than compiling larger data matrices of increasingly ambiguous and problematic morphological characters." In doing so we propose a new method that works as follows: (1) many (nearly 200) quantitative and qualitative characters are observed on species used in the molecular analysis. To avoid subjective homologization of characters, each character is binary coded (present/absent). (2) For each character, the two states are mapped on the combined molecular tree, and ancestral states are reconstructed under Mk_1 model using the MESQUITE ver. 1.12 program (Maddison and Maddison, 2006). Only characters that had evolved once (when present) and that can define particular internal molecular node are retained (Fig. 1a). (3) Characters that show a single secondarily loss are coded with a further derived state (i.e., two states represent the absence of the character) (Fig. 1b). (4) Characters that show high homoplasy are discarded (Fig. 1c). (5) This has resulted in 40 retained characters that were then scored for all of the 44 *Zaprionus s.s.* species (Appendix A). (6) A 'supermatrix' of 61 combined characters is built (Appendix B) consisting of two parts: the first part summarizes the combined molecular tree. In order to reduce the amount of missing molecular data that can result in upweighting the molecular set: only 21 fictional characters were retained *vs.* 2164 true nucleotide characters. The second part represents the 40 morphological characters. (7) An new phylogenetic tree including taxa that were missing in the original molecular tree was obtained using PAUP ver. 3.1 (Swofford, 2003) and MrBayes for MP and BI analyses, respectively. Within this new tree, the topology of the original molecular tree remained intact. Following this method, taxa with no molecular sequences were positioned on the molecular tree. Analogously to tree vocabulary, we called our method 'morphological grafting.'

2.5. Inference of the evolutionary history of Zaprionus

Divergence times were estimated using the *Drosophila* mutational clock with a rate of 1.1×10^{-8} substitutions per site per year per lineage (Tamura et al., 2004) using a relaxed clock under the UCLN model (Drummond et al., 2006). This was done using the BEAST package ver. 1.4.6. (Drummond and Rambaut, 2007) on the concatenated molecular sequences.

The current geographical distribution of the studied species (given in Appendix C) was used to reconstruct ancestral distributions on internal nodes of the phylogenetic tree. Geographical regions were treated like ordin-

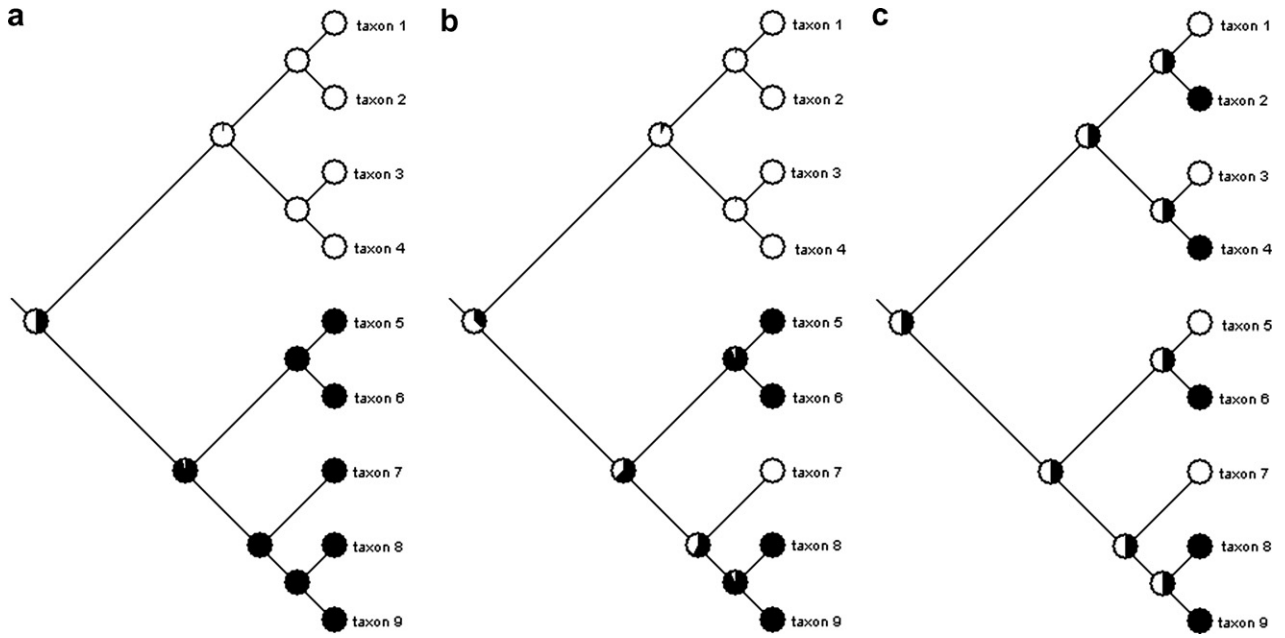

Fig. 1. Morphological character sampling and coding in the context of a molecular phylogeny as proposed by the 'morphological grafting' method. Lines indicate a hypothetical, molecular phylogenetic tree of nine taxa. Circles at terminal leaves (taxa) indicate a primarily coding of states of a morphological character as absent (white) or present (black). Circles at internal nodes indicate the maximum likelihood reconstruction of ancestral states under Mk_1 model. Color proportions at internal circles refer to the probability of the character state in the ancestor. Three cases are represented: (a) a character showing a very strong phylogenetic signal and which is going to be retained and coded with two states (absent or present); (b) a homoplasic character that is also going to be retained but recoded with three states (absent, present and secondarily lost); note that the state proportions at the ancestor of taxa 7, 8 and 9 indicate that the absence of the character in taxon 7 is more probably a reversal; (c) a highly homoplasic character that has to be discarded from the analysis.

ary characters, with the presence or absence of species within the particular region presented with binary codification. Nine geographical regions were considered: Oriental (O), Indian subcontinent (I), Middle-East (ME), Indian Ocean (IO), Southern Africa (SA), Eastern Africa (EA), Central Africa (CA), Western Africa (WA), and Americas (A). Maximum likelihood reconstruction of ancestral distributions was performed using the MESQUITE under the Mk_1 model. At each internal node, the likelihood estimated was compared among different geographical regions, and those with significantly highest probabilities were considered as the center of origin of different clades.

3. Results

3.1. Phylogenetic inference from combined molecular data

Phylogeny was inferred from combined data using the four reconstruction methods (NJ, MP, ML and BI). Fig. 2 shows the BI phylogram, while posterior probabilities for each clade (internal node) are given in Table 3, in comparison to bootstrap support values (after 1000 iterations) for the same clade under other reconstruction methods.

All analyses reconfirmed the paraphyly of the genus *Drosophila*, placing all *Zaprionus*, *Samoaia* and *Drosophila s.s.* species (i.e. *D. immigrans* and *D. repletoides*) within a single clade (with support values: NJ = 89, MP = 90, ML = 100, BI = 100; not given in Table 3), with *Drosophila (Sophophora) melanogaster* and *Scaptodrosophila* spp. being always the most distant taxa. The combination of *CO-II* sequences (and even the individual analysis of *CO-II*, not shown) reinforced the finding of Da Lage et al. (2007) using only *Amyrel* sequences that the *tumiditarsus* species group (i.e. *D. repletoides*) is the closest species group to the *Zaprionus* genus group (Fig. 2, node 1). Moreover, the previously presumed member of the *Zaprionus* genus group, the genus *Samoaia*, appeared to be slightly distant, and more related to *D. immigrans* (NJ = 39, ML = 57, BI = 91; with the exception of MP where it formed a clade with *D. repletoides* = 24).

All analyses supported the monophyly of the genus *Zaprionus* (node 2), as well as that of each of the two subgenera, *Anaprionus* and *Zaprionus s.s.*. However, only probabilistic methods (ML and BI) supported the subdivision of the subgenus *Zaprionus s.s.* into two sections: node 4 comprising species of the *inermis* group and of the *tuberculatus* subgroup, and node 12 comprising species of the *vittiger* subgroup. This result contradicts CT93 by illuminating the polyphyly of the group *armatus*. Both NJ and MP analyses did not contain node 3 which supports the *inermis* group rather there was a lack of resolution with a polytomy formed by *Zaprionus sexstriatus*, *Zaprionus ghesquierei*, and a clade containing

Fig. 2. Fifty percent majority-rule consensus tree from the BI analysis (500,000 generations) of combined molecular sequences (*CO-II* and *Amyrel*). Identical topologies were recovered from two distinct runs of MrBayes. Numbers above and below internal nodes indicate node number and posterior probability estimates, respectively. Shapes in front of species represent their taxonomic position according to CT93 (see text): blank circle (group *inermis*), solid triangle (group *armatus*: subgroup *tuberculatus*) and solid square (group *armatus*: subgroup *vittiger*) (see Table 1).

the remaining *inermis* species. The overlapping branching pattern between *inermis* and *tuberculatus* species (nodes 9–11) indicates the homoplasy of the character of a single chetiferous spine borne on a salient wart, once considered as a synapomorphy of species of the subgroup *tuberculatus* of the group *armatus*. Unarmed forefemora, which defined the *inermis* species group (*sensu* CT93), can be regarded as a symplesiomorphy shared by the outgroup subgenus *Anaprionus*.

The monophyly of species of the subgroup *vittiger* in the group *armatus* (*sensu* CT93) was supported whatever reconstruction method was used (node 12). These species were previously defined by the presence of a series of composite spines on their forefemora. This characteristic appears to be a good synapomorphy according to our analysis. NJ tree differed only in considering the most basal species and sister to the remaining species (*Zaprionus indianus* in NJ *vs. Zaprionus megalorchis* in all other methods).

3.2. Grafting morphological branches

Fig. 3 shows the BI tree obtained after grafting morphological branches (shown in black) on the consensus BI molecular phylogram given in Fig. 2 (shown in red in Fig. 3). Again, BI posterior probabilities (given below internal nodes) exceeded MP bootstrap values and gave higher resolution (not shown).

A revised phylogenetic classification with the new monophyletic species groups and complexes is also given in Fig. 3. *Zaprionus s.s.* is still divided into two main groups: *inermis* and *armatus*. The *inermis* group CT93 is divided into four species complexes: *sexvittatus* n. comp. Yassin (two species), *ghesquierei* n. comp. Yassin (one species), *inermis* n. comp. Yassin (two species), and *tuberculatus* Tsacas et al., 1977 n. comb. (five species). The *armatus* group CT93 is divided, in its turn, into species complexes: *litos* n. comp. Yassin (one species), *montanus* Chassagnard, 1989 (two species), *armatus* Chassagnard, 1989, *megalor-*

Table 3
Support values (1000 iterations bootstrap values for NJ, MP and ML and 500,000 generations posterior probability for BI) for clades (internal nodes) shown in Fig. 2 after analysis of combined molecular data

Node	Bootstrap			Posterior probability
	NJ	MP	ML	BI
1	62	—	57	67
2	90	72	86	100
3	100	99	100	100
4	—	—	46	89
5	48	—	43	62
6	100	100	100	100
7	100	100	100	100
8	88	57	56	71
9	100	100	100	100
10	—	—	59	70
11	99	98	100	100
12	100	99	100	100
13	—	70	97	92
14	99	92	39	100
15	—	—	47	89
16	—	—	100	99
17	100	100	100	100
18	99	95	68	100
19	51	45	69	99
20	—	69	100	100
21	—	56	48	84
22	99	97	100	100

(—) Incongruent nodes with the BI tree.

chis n. comp. Yassin (two species), *indianus* Yassin et al., in press (three species), *davidi* n. comp. Yassin (three species), *spinosus* n. comp. Yassin (five species), *vittiger* Tsacas, 1980 (three species), and *lachaisei* n. comp. Yassin (seven species). The CT93 subgroup category of the *armatus* group was not retained (Table 1), due to the inclusion of their *tuberculatus* subgroup as a new combined species complex in the *inermis* group, and to the paraphyly and polyphyly of their *armatus* subgroup (Fig. 3).

3.3. The evolutionary history of Zaprionus

For each internal node on the BI tree reconstructed from combined molecular data (Fig. 2, and shown in red above nodes in Fig. 3), the mean and the 95% confidence level (in MYA) and maximum likelihood distribution at every geographical region are given in Table 4. Historical biogeographical hypotheses, inferred from the results given in Table 4, are summarized on the geographical map shown in Fig. 4.

Okada (1981) was the first to propose an 'out-of-Asia' origin of the genus *Zaprionus*, in light of the Oriental distribution of its related genera (*Phorticella* and *Samoaia*), as well as of the *Drosophila immigrans* species group to which these genera are most allied (Throckmorton, 1975). Our analysis supports his hypothesis. Indeed, as shown in Table 4, the origin of the genus *Zaprionus* (node 1, Figs. 2 and 3) is significantly in the Oriental region ($P = 0.999$). Interestingly, this origin appears to be very recent (Middle Miocene, 13.81 ± 2.0 MYA), relative to the origin of the subgenus *Drosophila* of *Drosophila*, which was estimated to be during the Late Paleocene ~62.9 ± 12.4 MYA (Tamura et al., 2004).

Table 4 shows that the origin of the Afrotropical subgenus *Zaprionus s. str.* (nodes 2 and 3, Figs. 2 and 3) took place during the Late Miocene (from 10.59 ± 2.9 to 7.37 ± 0.66 MYA). This is in concordance with paleogeographic evidence stating that Africa was not in direct contact with other continents until the Early Miocene, when a definitive connection was formed with Eurasia (Gheerbrant and Rage, 2006). However, two scenarios may be proposed. On the one hand, a trans-Tethysian dispersal route via the Middle-East followed by an adaptive radiation of African *Zaprionus*, especially after the formation of the Great Rift Valley and the Red Sea at Late Miocene, acting as a geographical barrier. On the other hand, via an Indo-Malagasy route, which is in agreement with the origin of the heavy seasonal rains in Madagascar due to the initiation of Indian monsoons (~8 MYA) (Yoder and Nowak, 2006). A wet climate is conditional for tropical drosophilids. Our maximum likelihood reconstruction (Table 4 and Fig. 4) favored the second hypothesis ($P = 0.798$) although an East-African origin was not totally rejected ($P = 0.500$). This scenario is supported by the recent discovery of a *Phorticella* species, *Phorticella madagascariensis*, endemic to Madagascar (Chassagnard and McEvey, 1997). This genus, so far, has no representative in continental Africa.

Considering the current geographical distribution of species of the *inermis* group (nodes 4–11; Figs. 2 and 3), internal nodes had always higher probabilities at the Islands of the Indian Oceans (Table 4). This is mainly due to the fact that the four species (*Zaprionus cercus*, *Zaprionus mascariensis*, *Zaprionus kolodkinae* and *Zaprionus verruca*) are endemic to this region (Chassagnard and McEvey, 1992). Nonetheless, many species (*Z. sexstriatus*, *Z. sexvittatus*, *Z. inermis*) are found exclusively on continental Africa, as well as three others (*Z. ghesquierei*, *Z. sepsoides* and *Z. tuberculatus*) are also considered to be recent colonizers of Madagascar (Chassagnard and McEvey, 1992). This indicates that many independent trans-oceanic dispersals took place within this clade between mainland Africa and the Islands of the Indian Ocean, and vice versa, especially during the Pleistocene.

In contrast to the *inermis* group, with the exception of the problematic species *Zaprionus litos* and *Zaprionus simplex* with unarmed forefemora, there are no species of the *armatus* group endemic to Madagascar (Table 4 and Fig. 4) with the highest probabilities for occurrence being in mainland Africa. We do not know much about the divergence times of the *armatus* group due to the lack of molecular sequences. However, considering the *vittiger* group (node 12; Figs. 2 and 3), it appears to have originated in Central Africa ($P = 0.709$) during Early Pliocene (4.37 ± 0.99 MYA).

Major divergence events took place during the Pleistocene (nodes 14–22; Figs. 2 and 3) with the probabilities of historical geographical distribution oscillating between Eastern and Central Africa. This is in concordance with episodic glaciations periods which were responsible for the fragmentation and the re-expansion of Afrotropical rainforests during the Pleistocene, that had a major influence on the diversification of African drosophilids (Tsacas et al., 1981; Cobb et al., 2000).

For the more derived species complexes of the *vittiger* subgroup, two diversification patterns could be observed. The first is the sympatric mode of diversification in the

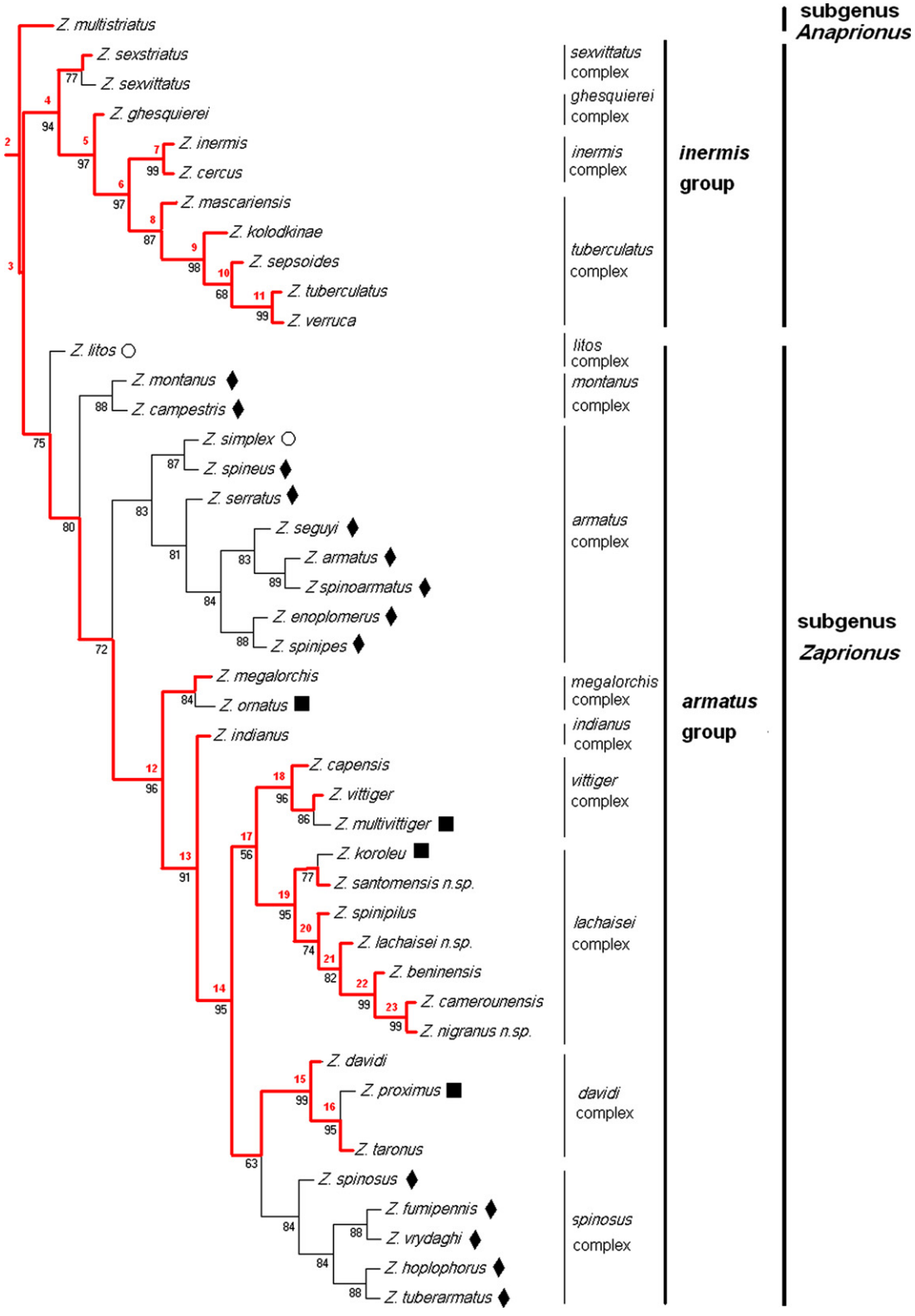

Table 4
Time and likelihood for geographical distribution at each internal node on the molecular tree given in Figs. 2 and 3

Node	Time (in MYA)			Biogeographical region (ML)							
	Mean	95% conf.		Asia				Africa			
		Min	Max	O	I	ME	IO	EA	CA	SA	WA
1	13.81	10.87	14.9	**0.999**	0	0	0	0	0	0	0
2	10.59	7.68	12.05	**0.976**	0.046	0	0.789	0.5	0.15	0.016	0.056
3	7.37	6.71	9.01	0.012	0.001	0	**0.798**	0.5	0.301	0.044	0.096
4	6.98	6.98	6.98	0	0	0.001	**0.712**	0.5	0.476	0.181	0.165
5	6.13	5.52	9.15	0	0	0.037	**0.931**	0.5	0.636	0.29	0.459
6	3.93	3.6	5.92	0	0	0.001	**0.943**	0.5	0.444	0.065	0.333
7	1.95	1.56	3.05	0	0	0	**0.835**	0.5	0.471	0.011	0.399
8	2.99	2.93	18	0	0	0	**0.989**	0.5	0.227	0.037	0.162
9	1.68	1.19	2.46	0	0	0	**0.996**	0.5	0.259	0.096	0.208
10	1.4	1.12	2.3	0	0	0.001	**0.996**	0.5	0.606	0.546	0.592
11	1.07	0.61	1.57	0	0.001	0.036	**0.997**	0.5	0.555	0.534	0.555
12	4.37	3.36	5.3	0	0	0.001	0.486	0.5	**0.709**	0.213	0.635
13	3.15	2.61	4.14	0	0	0.037	0.505	0.5	**0.811**	0.336	0.643
14	2.29	1.9	2.72	0	0	0	0.006	0.5	**0.719**	0.116	0.237
15	2.21	1.39	2.44	0	0	0	0.004	0.5	**0.835**	0.021	0.078
16	2.12	1.69	2.45	0	0	0	0.008	**0.5**	0.479	0.112	0.132
17	0.85	0.39	1.13	0	0	0	0.004	**0.5**	0.166	0.303	0.048
18	1.97	1.37	2.02	0	0	0	0.028	0.5	**0.54**	0.022	0.149
19	1.58	1.2	1.8	0	0	0	0.159	0.5	**0.619**	0.007	0.075
20	1.58	1.02	1.59	0	0	0	0.024	**0.5**	0.321	0.005	0.1
21	1.41	0.9	1.42	0	0	0	0.006	**0.5**	0.364	0.004	0.293
22	1.04	0.57	1.04	0	0	0	0.003	0.5	**0.835**	0.004	0.069

Divergence times (in MYA) were estimated under UCLN model. For each geographical region, maximum likelihood value was estimated under Mk_1 model (see text for abbreviations). Geographical regions with highest likelihood are bold faced.

two sister species complexes of *davidi* and *spinosus* (between nodes 14 and 15, Fig. 3). Species of both complexes are endemic to Central Africa (Appendix C) and both show a certain degree of differentiation of forefemoral spines. Moreover, species of the *spinosus* complex are known or putative anthophilic, (Tsacas and Chassagnard, 1990), which may suggest an ecological role in the diversification between these two complexes.

The second pattern is the allopatric diversification of the two other sister species complexes: *vittiger* and *lachaisei* (node 16, Figs. 2 and 3). The former complex is known only from Southern and Eastern Africa (node 17, Figs 2 and 3), whereas the later (node 18, Figs. 2 and 3) is of Central-Western African origin (Appendix C). The *lachaisei* complex contains two new species (*Z. santomensis* and *Z. nigranus*) which are endemic to the Atlantic island of São Tomé. Each is a sibling to another continental species: *Zaprionus koroleu* and *Zaprionus camerounensis*, respectively. The Cameroon volcanic line (CVL) might, thus, play a role in the insular speciation within this complex, as in other drosophilids (e.g., Lachaise et al., 2000; Cariou et al., 2001).

Recently, three distant Afrotropical species (*Z. indianus*, *Z. tuberculatus* and *Z. ghesquierei*) have acquired invasive capacities and were collected from the Palearctic region (Chassagnard and Kraaijeveld, 1991). *Z. indianus* is the most widespread *Zaprionus* species, found equally on three continents: India and the Middle-East, Africa and the Americas (Appendix C). However, this great expansion has been estimated using mtDNA to be recent, only during the second part of the 20th century (Yassin et al., in press). This recent expansion did not affect our estimates of ancestral distribution.

4. Discussion

4.1. Morphological grafting: an intermediate between supermatrix and supertree methods

The first aim of this paper was to present a new method to reduce the problems of missing data in reconstructing large phylogenetic trees from different data sets with overlapping taxa. Traditionally, two methods are usually used: supermatrix and supertree (see Section 1). De Queiroz and

Fig. 3. Fifty percent majority-rule consensus tree from the BI analysis (500,000 generations) of combined molecular and morphological data (on the left) reconstructed using the 'morphological grafting' technique (see text). Identical topologies were recovered from two distinct runs of MrBayes, with numbers below nodes (in black) indicate posterior probability estimates. Red branches represent the molecular phylogenetic tree given in Fig. 2, with nodes numerated in red according to the node number in Fig. 2. Thin black branches are those grafted using morphological characters for species of which no molecular sequences were available. For these species, their previous taxonomic position according to CT93 is as follows: blank circle (group *inermis*), solid diamond (group *armatus*: subgroup *armatus*) and solid square (group *armatus*: subgroup *vittiger*). New taxonomic designations are made (on the right) which modify the species composition from CT93.

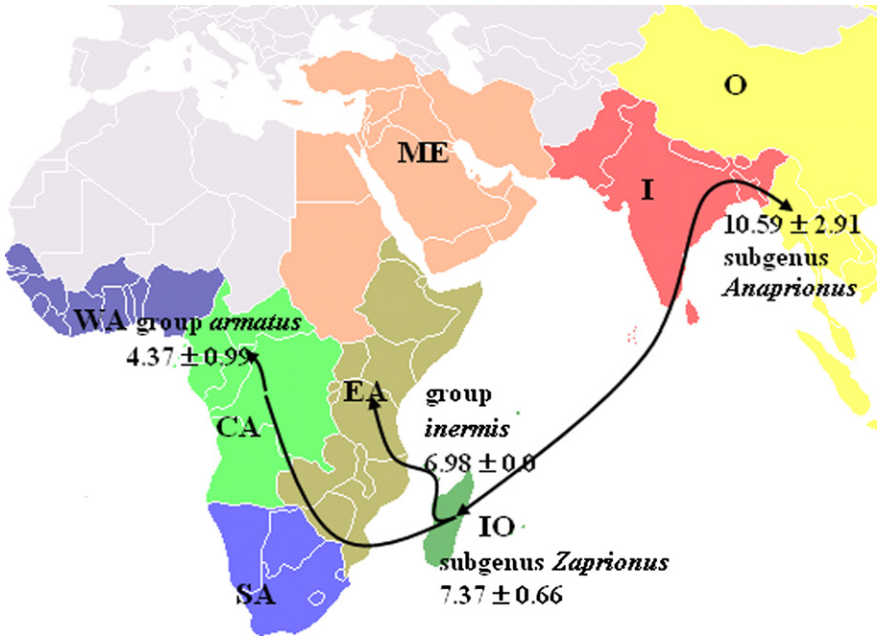

Fig. 4. Hypothetical reconstruction of the historical biogeography of the genus *Zaprionus* from the results shown in Table 3. Major geographical regions are colored and abbreviated as follows: O, Oriental (yellow); I, India (red); ME, Middle-East (orange); IO, Islands of the Indian Ocean (green); SA, South Africa (blue); EA, Eastern Africa (brown); CA, Central Africa (light green) and WA, Western Africa (indigo). Ages of colonization (in MYA) are indicated for the two subgenera and the two groups of *Zaprionus s.s.* (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

Gatesy (2007) compared the two methods and privileged the supermatrix method for using full character evidence, which may be lost when summarized as trees in the supertree method, and different sets of characters (e.g., morphology). However, supermatrix approach implicitly assumes that all characters have experienced the same branching history, which is not always valid (Crandall and Buhay, 2004).

We proposed here a new method intermediate between the two previous ones, that we have called 'morphological grafting'. In this method, a supermatrix is built of two parts: the first part is a matrix summarizing a robust tree obtained from a molecular supermatrix. The second part includes all taxa, even those with missing DNA sequences (and thus not present in the molecular tree), with a matrix of their morphological characters. The novelty in this method is the dependency of the second part on the first one. The molecular tree is used as an external hypothesis to *ad hoc* assess the homoplasy of the morphological structures. However, the procedure is still laborious and complicated without a single software compilation of the different steps. There is thus a strong need to create a computer program with algorithms capable to recode morphological characters in the molecular phylogenetic context. In doing so, this program will be able to estimate probabilistic evolutionary models of character state transformations that can be utilized later in building morphological phylogenies using likelihood (Lewis, 2001) or Bayesian methods (Ronquist, 2004). The lack of consistent evolutionary models has long been a deep critique to morphological phylogenetics, that has limited it to maximum parsimony (Buckley, 2002). Although we were limited to the Mk model in our Bayesian analysis, a computerized morphological grafting method may provide a statistical future for morphological phylogenetics beyond this simplest model.

4.2. Molecular and morphological phylogenetics of the Drosophilidae

Our method may be of major significance in revising the large and most morphologically and ecologically diversified Muscomorphan family, the Drosophilidae. There are 3939 described drosophilid species, of which only 560 species (almost 14%) possess sequences in the GenBank (as of August 1st, 2007). Among the sequenced species, 84% belong to the genus *Drosophila*, only one of the 73 genera of the family. This shows that in spite of the greater advantage of molecular sequences in phylogeny reconstruction, morphological structures still and will remain a very rich character source for reconstructing an explicit phylogeny of the Drosophilidae. Nonetheless, previous attempts to include DNA and morphology in Drosophilidae resulted in conflicting trees (DeSalle and Grimaldi, 1991; Kwiatowski and Ayala, 1999; Remsen and O'Grady, 2002). All these studies, however, followed the same supermatrix approach. Explicit morphological phylogenetic studies for other drosophilid genera have been conducted on other drosophilid genera (*Zygothrica*, Grimaldi, 1987; *Colocasiomyia*, Grimaldi, 1992; Sultana et al., 2006; *Cladochaeta*, Grimaldi and Nguyen, 1999; *Lordiphosa*, Hu and Toda, 2001; *Amiota*,

Chen and Toda, 2001; *Dichaetophora*, Hu and Toda, 2002; *Pseudostegana*, Chen et al., 2005). A molecular refinement of these studies should be a good start for the application of morphological grafting in the phylogenetics of the Drosophilidae.

4.3. Zaprionus: a new model drosophilid clade

The second aim of this paper was to investigate the evolutionary history of the drosophilid genus *Zaprionus*, and to present it as a new model clade in evolutionary studies. DeSalle and Grimaldi (1991) blamed the community of *Drosophila* biologists to focus only on a single genus of drosophilids: the genus *Drosophila*. The most classical drosophilid clade in evolutionary biology is the *Drosophila melanogaster* species subgroup (David et al., in press). Despite of the large amount of data available from this subgroup (especially from the two sibling species *D. melanogaster* and *D. simulans*), its number of species is only 9, with 2 species (*Drosophila erecta* and *Drosophila sechellia*) known to have specialized ecological niches. *Zaprionus*, by contrast, is a very rich genus containing about 60 species, and interestingly it appears to share the same age and geographic origin as the *melanogaster* subgroup, which too has originated during Middle- to Early-Miocene in the Oriental region and then diversified in Tropical Africal (Lachaise et al., 2004). Ecological and morphological diversity of *Zaprionus* species, however, far exceeds that of the *melanogaster* subgroup, with half of them (about 30 species) being easily grown under laboratory conditions, and with many species known to be anthophilic. This may allow a number of comparative genetics studies with a statistically sufficient number of close species of known phylogenetic relationships that address questions about the evolutionary significance of biological diversity in a clade (e.g., karyotypes, morphology, ecophysiology, behavior, geographical distribution, etc.).

Another interesting *Drosophila* clade for such comparative studies is the *obscura* species group, containing 41 species with an estimated age of origin about 18 MYA (Tamura et al., 2004). Moréteau et al. (2003) and Huey et al. (2006) investigated the quantitative evolution of body size in 20 species of this clade in a phylogenetic context. Nonetheless, the phylogeny used in these studies was arbitrarily reconstructed from different allozyme and DNA studies (i.e., they lacked relative branch lengths). This gives an advantage for *Zaprionus* species cultured in Gif-sur-Yvette, of which we possess now, thanks to this study, a robust molecular phylogeny. Indeed, several statistical analyses of the phylogenetic comparative method require branch lengths (Garland et al., 2005). In addition, most species of the *obscura* group are holarctic, with one species, *Drosophila subobscura*, became invasive and extended its geographical borders to the south. *D. subobscura* was the first invasive drosophilid to be used as an evolutionary tool for the study of adaptation (Ayala et al., 1989; Huey et al., 2005). An interesting difference between this model and the invasive species of the genus *Zaprionus*, *Z. indianus*, is the opposing direction of invasion. *Z. indianus*, by contrast, is a tropical species that is currently expanding its northern borders to the temperate regions. This will imply opposite patterns of adaptation to temperature (and thus latitudinal cline formation) in the two species that are worth investigation.

Acknowledgments

We are most grateful to Drs. Masanori Toda (Hokkaido University, Japan), Didier Casane (CNRS, France), Stéphane Prigent (Academia Sinica, Taiwan), and to two anonymous reviewers for their fruitful comments and detailed suggestions. We also thank Drs. Daniel Lachaise, Stéphane Dupas (CNRS, France) and Joseph Vouidibio (Université de Brazzaville, Congo) for collecting the examined material. We are also indebted to Dr. Christophe Daugeron (MNHN, France) for lending museum specimens. This work was supported by the Fondation des Treilles, the Agence Universitaire de la Francophonie (AUF), the Comité Française d'Evaluation de la Coopération avec le Brésil (COFECUB) and the Coordençaõ de Aperfeicoamento de Pessaoal de Nivel Superior (CAPES).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2008.01.036.

References

Ayala, F.J., Serra, L., Prevosti, A., 1989. A grand experiment in evolution: the *Drosophila subobscura* colonization of the Americas. Genome 31, 246–255.

Baker, R.H., Gatesy, J., 2002. Is morphology still relevant? In: DeSalle, R., Giribet, G., Wheeler, W. (Eds.), Molecular Systematics and Evolution: Theory and Practice. Birkhäuser Verlag, Basel, Switzerland, pp. 163–174.

Bininda-Emonds, O.R.P., 2004. The evolution of supertrees. TREE 19, 315–322.

Buckley, T.R., 2002. A statistical future for morphological phylogenetics? TREE 17, 161.

Cariou, M.-L., Silvain, J.-F., Daubin, V., Da Lage, J.-L., Lachaise, D., 2001. Divergence between *Drosophila santomea* and allopatric or sympatric populations of *D. yakuba* using paralogous amylase genes and migration scenarios along the Cameroon volcanic line. Mol. Ecol. 20, 649–660.

Chassagnard, M.-T., Kraaijeveld, A.R., 1991. The occurrence of *Zaprionus sensu stricto* in the Palearctic region (Diptera: Drosophilidae). Ann. Soc. Entomol. Fr. 27, 495–496.

Chassagnard, M.-T., McEvey, S.F., 1992. The *Zaprionus* of Madagascar, with descriptions of five new species (Diptera: Drosophilidae). Ann. Soc. Entomol. Fr. 28, 317–335.

Chassagnard, M.-T., McEvey, S.F., 1997. Le genre *Phorticella* Duda de la région afrotropicale et de Sumatra. Description de deux nouvelles espèces (Diptera: Drosophilidae). Ann. Soc. Entomol. Fr. 33, 81–87.

Chassagnard, M.-T., Tsacas, L., 1993. Le sous-genre *Zaprionus s.str.* Definition de groupes d'espèces et revision du sous-groupe *vittiger* (Diptera: Drosophilidae). Ann. Soc. Entomol. Fr. 29, 173–194.

Chen, H.-W., Toda, M.J., 2001. A revision of the Asian and European species in the subgenus *Amiota* Loew (Diptera, Drosophilidae) and the establishment of species-groups based on phylogenetic analysis. J. Nat. Hist. 35, 1517–1563.

Chen, H.-W., Toda, M.J., Wang, B.-C., 2005. A taxonomic revision of the genus *Pseudostegana* Okada, 1978 (Diptera, Drosophilidae). Insect Syst. Evol. 36, 407–442.

Cobb, M., Huet, M., Lachaise, D., Veuille, M., 2000. Fragmented forests, evolving flies: molecular variation in African populations of *Drosophila teissieri*. Mol. Ecol. 9, 1591–1597.

Crandall, K.A., Buhay, J.E., 2004. Genomic databases and the tree of life. Science 306, 1144–1145.

Da Lage, J.-L., Renard, E., Chartois, F., Lemeunier, F., Cariou, M.-L., 1998. *Amyrel*, a paralogous gene of the *amylase* gene family in *Drosophila melanogaster* and the *Sophophora* subgenus. Proc. Natl. Acad. Sci. USA 95, 6848–6853.

DaLage, J.-L., Kergoat, G.J., Maczkowiak, F., Silvain, J.-F., Cariou, M.-L., Lachaise, D., 2007. A phylogeny of Drosophilidae using the *amyrel* gene: questioning the *Drosophila melanogaster* species group boundaries. J. Zool. Syst. Evol. Res. 45, 47–63.

David, J.R., Lemeunier, F., Tsacas, L., Yassin, A., in press. The historical discovery of species of the *Drosophila melanogaster* subgroup. Genetics.

de Bruijn, M.H., 1983. *Drosophila melanogaster* mitochondrial DNA, a novel organization and genetic code. Nature 304, 234–241.

DeSalle, R., Grimaldi, D.A., 1991. Morphological and molecular systematics of the Drosophilidae. Ann. Rev. Ecol. Syst. Evol. 22, 447–475.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Ranbaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Durando, C.M., Baker, R.H., Etges, W.J., Heed, W.B., Wasserman, M., DeSalle, R., 2000. Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. Mol. Phylogenet. Evol. 16, 296–307.

Garland, T., Bennett, A.F., Rezende, E.L., 2005. Phylogenetic approaches in comparative physiology. J. Exp. Biol. 208, 3015–3035.

Gheerbrant, E., Rage, J.-C., 2006. Paleobiogeography of Africa: how distinct from Gondwana and Laurasia? Paleogeogr. Paleoclimatol. Paleoecol. 241, 224–246.

Gloor, G., Engels, W., 1992. Single fly DNA preps for PCR. Dros. Inf. Serv. 71, 148–149.

Grimaldi, D., 1992. Systematics of the genus *Colocasiomyia* de Meijere (Diptera: Drosophilidae): cladistics, a new generic synonym, new records, and a new species from Nepal. Entomol. Scand. 22, 417–426.

Grimaldi, D., Nguyen, T., 1999. Monograph on the spittlebug flies, genus *Cladochaeta* (Diptera: Drosophilidae: Cladochaetini). Bull. Am. Mus. Nat. Hist. 241, 1–326.

Grimaldi, D.A., 1987. Phylogenetics and taxonomy of *Zygothrica* (Diptera: Drosophilidae). Bull. Am. Mus. Nat. Hist. 186, 103–268.

Grimaldi, D.A., 1990. A phylogenetic, revised classification of genera in the Drosophilidae (Diptera). Bull. Am. Mus. Nat. Hist. 197, 1–139.

Guindon, S., Pascual, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Hu, Y.-G., Toda, M.J., 2001. Polyphyly of *Lordiphosa* and its relationships in Drosophilinae (Diptera: Drosophilidae). Syst. Entomol. 26, 15–31.

Hu, Y.-G., Toda, M.J., 2002. Cladistic analysis of the genus *Dichaetophora* Duda (Diptera: Drosophilidae) and a revised classification. Insect Syst. Evol. 33, 91–102.

Huey, R.B., Gilchrist, G.W., Hendry, A.P., 2005. Using invasive species to study evolution: case studies with *Drosophila* and salmon. In: Sax, D.F., Stachowicz, J.J., Gaines, S.D. (Eds.), Species Invasions: Insights into Ecology, Evolution, and Biogeography. Sinauer Associates, Inc., Sunderland, MA, USA, pp. 139–164.

Huey, R.B., Moréteau, B., Moréteau, J.C., Gibert, P., Gilchrist, G.W., Ives, A.R., Garland Jr., T., David, J.R., 2006. Sexual dimorphism in a *Drosophila* clade, the *D. obscura* group. Zoology 1090, 318–330.

Jenner, R.A., 2004. Accepting partnership by submission? Morphological phylogenetics in a molecular millennium. Syst. Biol. 53, 333–342.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefs Bioinform. 5, 150–163.

Kwiatowski, J., Ayala, F.J., 1999. Phylogeny of *Drosophila* and related genera: conflict between molecular and anatomical analyses. Mol. Phylogenet. Evol. 13, 319–328.

Lachaise, D., Harry, M., Solignac, M., Lemeunier, F., Benassi, V., Cariou, M.-L., 2000. Evolutionary novelties in islands: *Drosophila santomea*, a new *melanogaster* sister species from São Tomé. Proc. R. Soc. Biol. Sci. 267, 1487–1495.

Lachaise, D., Capy, P., Cariou, M.L., Joly, D., Lemeunier, F., David, J.R., 2004. Nine relatives from one African ancestor: population biology and evolution of the *Drosophila melanogaster* subgroup species. In: Singh, R.S., Uyenoyama, M.K. (Eds.), The Evolution of Population Biology. Cambridge University Press, New York, USA.

Lee, M.S.Y., 2006. Morphological phylogenetics and the universe of useful characters. Taxon 55, 5–7.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925.

Maddison, W.P., Maddison, D.R., 2006. Mesquite: a Modular System for Evolutionary Analysis, version 1.12. <http://mesquiteproject.org>.

Moréteau, B., Gibert, P., Pétavy, G., Moréteau, J.C., Huey, R.B., David, J.R., 2003. Morphometric evolution in a *Drosophila* clade: the *Drosophila obscura* group. J. Zool. Syst. Evol. Res. 41, 64–71.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Okada, T., 1981. Oriental species, including New Guinea. In: Ashburner, M., Carson, H.L., Thompson, J.N. (Eds.), The Genetics and Biology of *Drosophila*, vol. 3a. Academic Press, New York, pp. 261–289.

Okada, T., Carson, H.L., 1983. The genera *Phorticella* Duda and *Zaprionus* Coquillett (Diptera, Drosophilidae) of the Oriental region and New Guinea. Kontyu 51, 539–553.

Olmstead, R.G., Scotland, R.W., 2005. Molecular and morphological data sets. Taxon 54, 7–8.

Remsen, J., O'Grady, P., 2002. Phylogeny of Drosophilinae (Diptera: Drosophilidae), with comments on combined analysis and character support. Mol. Phylogenet. Evol. 24, 249–264.

Robe, L.J., Valente, V.L.S., Budnik, M., Loreto, E.L.S., 2005. Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: a nuclear versus mitochondrial gene approach. Mol. Phylogenet. Evol. 36, 623–640.

Ronquist, F., 2004. Bayesian inference of character evolution. TREE 19, 475–481.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 9, 406–425.

Sanderson, M.J., Shaffer, H.B., 2002. Troubleshooting molecular phylogenetic analyses. Ann. Rev. Ecol. Syst. 33, 49–72.

Scotland, R.W., Olmstead, R.G., Bennett, J.R., 2003. Phylogeny reconstruction: the role of morphology. Syst. Biol. 52, 539–548.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., 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.

Smith, N.D., Turner, A.H., 2005. Morphology's role in phylogeny reconstruction: perspectives from paleontology. Syst. Biol. 52, 539–548.

Sultana, F., Hu, Y.-G., Toda, M.J., Takenaka, K., Yafuso, M., 2006. Phylogeny and classification of *Colocasiomyia* (Diptera, Drosophilidae), and its evolution of pollination mutualism with aroid plants. Syst. Entomol. 31, 684–702.

Swofford, D.L., 2003. PAUP: Phylogenetic Analysis Using Parsimony, version 4.0. Computer program distributed by Illinois Natural History Survey, Illinois, USA.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526.

Tamura, K., Subramanian, S., Kumar, S., 2004. Temporal patterns of fruitfly (*Drosophila*) evolution revealed by mutation clocks. Mol. Biol. Evol. 21, 36–44.

Tavaré, S., 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Lect. Math. Life Sci. 17, 57–86, 17, 368–376.

Tao, N. 2005. FINDMODEL: a Tool to Select the Best-fit Model of Nucleotide Substitution. M.Sc. Thesis, University of New Mexico, New Mexico.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Throckmorton, L.H., 1975. The phylogeny, ecology, and geography of *Drosophila*. In: King, R.C. (Ed.), Handbook of Genetics, vol. 3. Plenum Press, New York, pp. 421–469.

Tsacas, L., Chassagnard, M.-T., 1990. Les espèces du genre *Zaprionus* à fémurs antérieurs spinuleux (Diptera: Drosophilidae). Ann. Soc. Entomol. Fr. 26, 461–487.

Tsacas, L., Lachaise, D., David, J.R., 1981. Composition and biogeography of the Afrotropical drosophilid fauna. In: Ashburner, M., Carson, H.L., Thompson, J.N. (Eds.), The Genetics and Biology of *Drosophila*, vol. 3a. Academic Press, New York, pp. 197–259.

Wortley, A.H., Scotland, R.W., 2006. The effect of combining molecular and morphological data in published phylogenetic analyses. Syst. Biol. 55, 677–685.

Yassin, A., Capy, P., Madi-Ravazzi, L., Ogereau, D., David J.R., in press. DNA barcode discovers two cryptic species and two geographical radiations in the invasive drosophilid *Zaprionus indianus*. Mol. Ecol. Notes.

Yoder, A.D., Nowak, M.D., 2006. Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. Annu. Rev. Ecol. Syst. 37, 405–431.