



**HAL**  
open science

# Phylogenetic analyses of termite post-embryonic sequences illuminate caste and developmental pathway evolution

Frédéric Legendre, Michael Whiting, Philippe Grandcolas

► **To cite this version:**

Frédéric Legendre, Michael Whiting, Philippe Grandcolas. Phylogenetic analyses of termite post-embryonic sequences illuminate caste and developmental pathway evolution. *Evolution & Development*, 2013, 15 (2), pp.146-157. 10.1111/ede.12023 . mnhn-02520984

**HAL Id: mnhn-02520984**

**<https://hal-mnhn.archives-ouvertes.fr/mnhn-02520984>**

Submitted on 19 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Phylogenetic analyses of termite post-embryonic sequences illuminate caste and developmental pathway evolution

Frédéric Legendre,<sup>a,\*</sup> Michael F. Whiting,<sup>b</sup> and Philippe Grandcolas<sup>a</sup>

<sup>a</sup> UMR 7205 CNRS, Origine, Structure et Evolution de la Biodiversité, Département Systématique et Evolution, Muséum national d'Histoire naturelle, CP 50, 45, rue Buffon, 75005, Paris, France

<sup>b</sup> Department of Biology and M.L. Bean Museum, Brigham Young University, 693 Widtsoe Building, Provo, UT 84602, USA

\*Author for correspondence (e-mail: legendre@mnhn.fr)

**SUMMARY** Termites are highly eusocial insects with a caste polyphenism (i.e., discontinuous morphological differences between castes) and elaborated behaviors. While the developmental pathways leading to caste occurrence are well-known in many species, the evolutionary origin of these pathways is still obscure. Recent molecular phylogenetic studies suggest multiple independent origins of sterile castes in termites, reviving a 30 years old debate. We demonstrate here that diploid sterile castes (“true” workers) evolved several

times independently in this group and that this caste was lost at least once in a lineage with developmentally more flexible workers called pseudergates or “false” workers. We also infer that flexibility in post-embryonic development was acquired multiple times independently during termite evolution. We suggest that focusing on detailed developmental pathways in phylogenetic analyses is essential for elucidating the origin of caste polyphenism in termites.

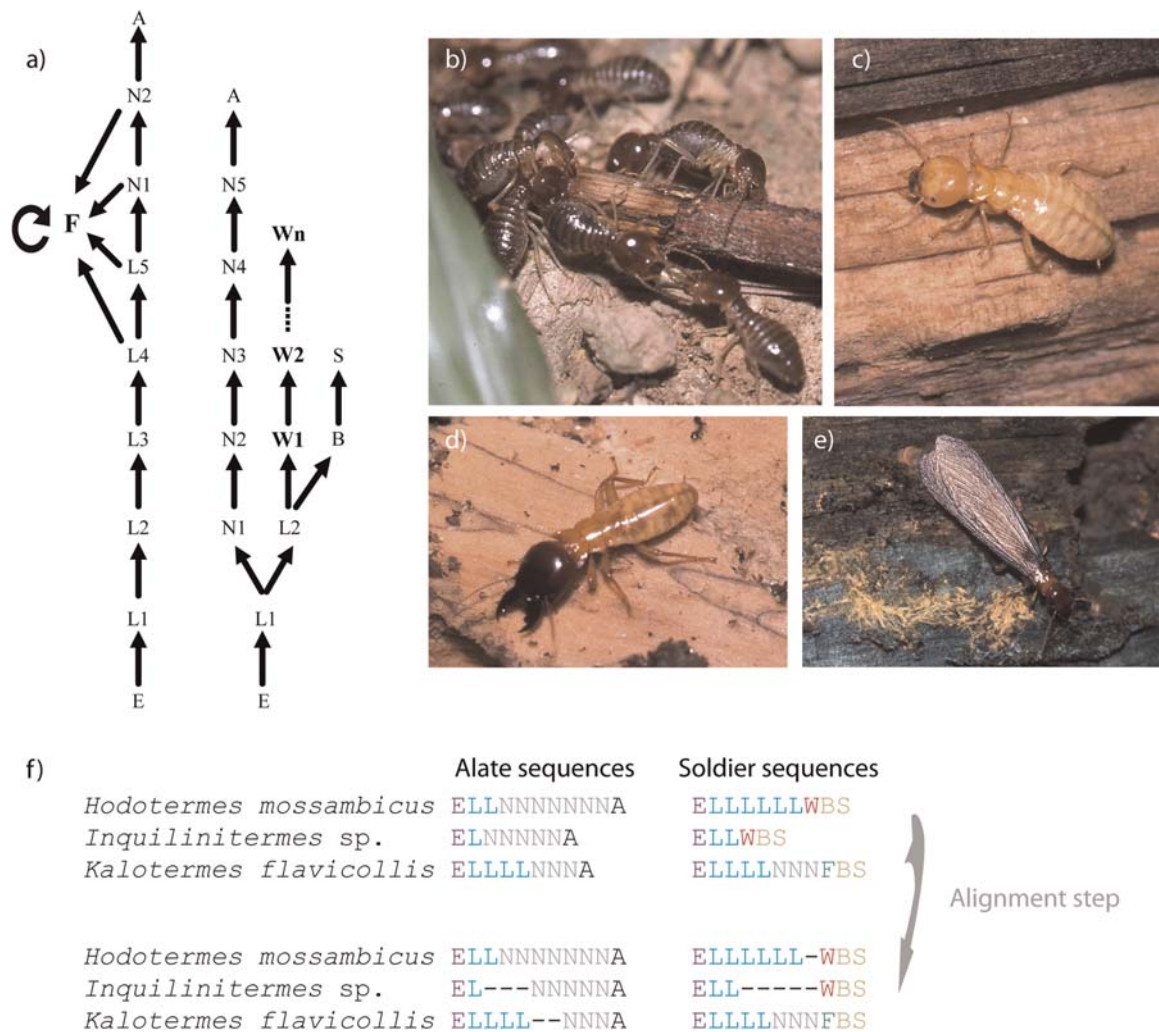
## INTRODUCTION

The evolution of complex and integrated social systems, Darwin’s “special difficulty” (Darwin 1859), still challenges modern evolutionary biologists (e.g., Hamilton 1964; Wilson 1971; Nowak et al. 2010; Nonacs 2011; Rousset and Lion 2011; Johnstone et al. 2012): How could sterile individuals, which are incapable of reproducing, have possibly evolved? How could they evolve independently in different lineages (Duffy and Macdonald 2010)? How does one explain the evolution of sociality in diploid organisms (Costa 2006), given that the relatedness component of Hamilton’s hypothesis does not play as strong a role as in haplodiploid organisms? Even though several studies have conceptually addressed the evolutionary origin of sterile caste in insects (e.g., Bartz 1979; Roisin 1994; Boomsma 2009; see Howard and Thorne 2011 for a recent review), the origin of these castes is not yet clear. While the developmental pathways leading to these castes are well known in many species, these pathways have yet to be thoroughly analyzed from a phylogenetic standpoint (see Korb and Hartfelder 2008; Nalepa 2010 for attempts in this direction). In order to understand the evolution of sociality, it is critical to identify through a comparative approach how developmental trajectories of sterile and sexual individuals evolved and diverged. Once key developmental events are characterized, new explanatory hypotheses about eusociality evolution emerge. Here, we perform such a comparative analysis on termites, a

group of hemimetabolous diploid insects. Because termites have an extensive caste system, they offer an excellent model for studying the sterile caste paradox and its relation to the evolution of development.

Termites have four castes resulting from the expression of alternative post-embryonic developmental trajectories. Alates or imagoes (Fig. 1e) follow a “normal” development (Noirot 1985a) with a post-embryonic developmental pathway similar to the one occurring in other, non-eusocial, hemimetabolous insects. Soldiers (Fig. 1d) share, in every termite species, a similar post-embryonic developmental pathway passing through a pre-soldier or white soldier stage (Noirot and Pasteels 1987) and likely have a single origin even though they express high interspecific morphological variability (Prestwich 1984). Two worker castes have been distinguished primarily based on a developmental criterion (Noirot and Pasteels 1987): “true” workers (Fig. 1b) defined “as individuals diverging early and irreversibly from the imaginal development” (Fig. 1a—right); and pseudergates (or “false” workers—Fig. 1c) defined “as individuals separating late from the imaginal line” (Fig. 1a—left).

Moreover, termites show a diversity of developmental flexibility, the evolution of which is closely intertwined with worker caste evolution. Pseudergates can develop from and into several different stages, including stages from the imaginal line. They have, by definition, a large spectrum of developmental possibilities so that species with pseudergates are



**Fig. 1.** Post-embryonic sequences and castes in termites. (a) Illustration of Noirot and Pasteels' ontogenetic criterion in worker caste recognition: simplified linear and bifurcated developmental pathways of *Kalotermes flavicollis* (left) and of *Microcerotermes* sp. (right), respectively. Each arrow represents a molt. Worker stages are in bold; (b) "true" workers of *Macrotermes annandalei*; (c)–(e) pseudergate, soldier and alate of *Hodotermopsis sjoestedti*, respectively; f) post-embryonic sequences of three termite species: left versus right = sequences ending in alates versus sequences ending in soldiers; top versus bottom = unaligned versus aligned sequences—this alignment is one putative homology hypothesis among several possibilities. A, alate; B, Pre-soldier; E, egg; F, pseudergate; L, larva; N, nymph; S, soldier; W, 'true' workers. Photos courtesy of C. Bordereau.

developmentally highly flexible. On the other hand, "true" workers, although capable of developing into soldier stages, do not have such a large spectrum of developmental potentialities, and usually display low developmental flexibility. High developmental flexibility has been considered either as an ancestral characteristic subsequently lost in some species, or as a derived one, acquired more recently (Noirot 1985b; Watson and Sewell 1985). The irreversible divergence from the imaginal line (referred to as "timing of differentiation") varies according to the species, and its ancestral condition (late or early) is unknown (Noirot 1985a; Watson and Sewell 1985).

At the macroevolutionary level, the evolution of sociality has been inferred by reconstruction of social and caste systems as

simple binary characters on molecular phylogenetic trees (Thompson et al. 2000, 2003; Grandcolas and D'Haese 2002, 2004; Inward et al., 2007; Pellens et al. 2007; Legendre et al. 2008a). Surprisingly, a comprehensive analysis of developmental stages from a phylogenetic standpoint is still lacking, as these previous attempts did not define castes systems with sufficient precision to adequately understand their evolution (Grandcolas and D'Haese 2003; Grandcolas et al. 2011). For instance, Termitidae and Mastotermitidae species each have a "true" worker caste, yet, in Mastotermitidae, "true" workers do not develop before the seventh instar while in Termitidae they can develop in the third instar, so these may not be homologous states.

The same problem is evident in the definition of pseudergate. As advocated recently (Korb and Hartfelder 2008; Roisin and Korb 2010), the original definition of pseudergate is impractical and does not accurately characterize the instars of *Prorhinotermes*, *Termitogeton* or *Glossotermes* species, so these authors proposed a broader definition, pseudergates *sensu lato*, which includes both larval and nymphal instars. However, combining larvae and nymphs into a single category may also result in nonhomologous states.

Combining the results of developmental processes into single, artificial categories obfuscates developmental differences which could be pivotal in deciphering caste evolution and that could only be revealed through the analysis of developmental sequences. Deciphering the succession of developmental stages in a phylogenetic analysis permits the inference of when developmental flexibility first originated and how the timing of differentiation has evolved.

One of the challenges in using developmental sequences in a phylogenetic framework has been the lack of computational tools to adequately deal with the data. Recent advances have resulted in more powerful phylogenetic tools, developed largely for the analysis of DNA sequence data (Thompson et al. 1994; Wheeler 1996; Edgar 2004) and developmental or behavioral sequences (Smith 1997; Velhagen 1997; Abbott and Tsay 2000; Bininda-Emonds et al. 2002; Jeffery et al. 2002, 2005; Hay et al. 2004; Schulmeister and Wheeler 2004; Robillard et al. 2006; Legendre et al. 2008b). Smith (1997) and Velhagen (1997) concomitantly proposed an “event-pairing” method to compare developmental sequences. But this method, which partitions developmental sequences in paired events, suffers from strong methodological inconsistencies and can lead to erroneous results when it comes to phylogenetic tree reconstruction or optimization of hypothetical ancestors (Schulmeister and Wheeler 2004). Schulmeister and Wheeler (2004) showed that whole developmental sequences could be used instead of paired events and proposed to analyze them as a single character (i.e., one developmental sequence is treated as a single multistate character in a search-based optimization procedure—Wheeler 2003). This method necessitates, however, a pool of predetermined developmental sequences to reconstruct hypothetical ancestral sequences, and thus requires untestable assumptions.

We rather suggest treating developmental sequences in a similar manner to DNA sequences during the alignment stage in phylogenetics (Fig. 1f). Just as DNA sequences are successions of nucleotides whose positional homology relative to another sequence is uncertain, post-embryonic developmental sequences are successions of instars or stages whose homology to stages in other species is uncertain.

For this analysis, we combine DNA sequences of seven genetic markers and developmental data (post-embryonic sequences from the egg to soldier or alate stages) to estimate phylogenetic relationships among termite families. We then use

this phylogeny to study caste and post-embryonic developmental pathway evolution through ancestral state reconstructions of developmental sequences. We also analyze developmental and molecular data separately to perform reconciliation analyses between both trees to address the evolution of developmental flexibility which provide new insights into the evolution of caste systems in termites.

## MATERIALS AND METHODS

### Sampling and molecular methods

Molecular data were primarily derived from a previous phylogenetic analysis (Legendre et al. 2008a) we conducted with seven markers (*12S rRNA*, *16S rRNA*, *18S rRNA*, *28S rRNA*, *cytochrome b*, and *cytochrome oxidase subunits I and II*; ~7500 bp) and 40 termite species representing the seven families currently recognized in termite classification. This sample put the emphasis on “lower” termites, the phylogenetic relationships of which are more controversial (Inward et al. 2007, Legendre et al. 2008a). We supplemented these data with sequences of *Glossotermes oculatus* (GenBank accession numbers JN647689–JN647692), the single Serritermitidae species for which developmental data are available (Bourguignon et al. 2009). Our molecular data set is thus comprised of seven partitions for 41 termite species. Primers and molecular techniques were described in Legendre et al. (2008a).

### Developmental data set

We have undertaken a comprehensive survey of the literature to document developmental data for termites and outgroup species (supporting information Table S1 and Appendix S1). Our developmental data set comprised two “kinds” of homologous developmental sequences: sequences ending in the alate stage and sequences ending in the soldier stage. We explain below the rationale for treating these as two developmental sequences (some supporting information details are given in Appendix S2). The alate stage constitutes the final stage in the imaginal line and every developmental sequence ending in an alate stage is homologous. Similarly, even if a few studies (e.g., Myles 1986) have shown that reproductive soldiers can occur in Termopsidae in exceptional circumstances, the soldier stage is a final stage in the development of an individual. Moreover, soldiers probably have a single origin (Noirot and Pasteels 1987). Every developmental sequence ending in a soldier stage is thus homologous. On the contrary, pseudergates and true workers are not final stages given that they can still differentiate into soldiers or other stages. In several species, most, if not all, of the soldiers derive from a worker stage (Noirot 1985b). Consequently, developmental sequences ending in any worker stage could not be used as a third “kind” of developmental sequence. As not every species has been studied ontogenetically, some chimeras were

built when developmental data were available only for congeneric species.

We considered developmental flexibility (i.e., species polymorphism) in taking several terminals for the same species (89 terminals for 35 species; see supporting information Appendix S2). Relevant outgroups (Grandcolas et al. 2004), that is, *Locusta migratoria* and four cockroach species (including one species of the putative termites sister-genus *Cryptocercus*), and 30 termite species representing the seven currently recognized families were documented with developmental data, for a total combined sampling of 109 terminals.

The developmental data set comprised eight character states, so that standard molecular alignment techniques could not be used. Thus, we designed a specific alphabet for direct optimization procedures, using the “custom\_alphabet” feature of POY4 (Varón et al. 2010). This alphabet allowed us to consider the eight character states of termite developmental sequences in a phylogenetic framework, and to define costs associated with any change (substitutions and insertions/deletions) in these developmental sequences, as it is typically done with the four character states of DNA sequences.

## Phylogenetic analyses

We used parsimony and Bayesian methods to reconstruct phylogenetic trees and to conduct ancestral state reconstructions of developmental data. In parsimony, we analyzed developmental data as “dynamic” and “static” characters (i.e., developmental sequences either optimized during tree search with direct optimization procedures or aligned before tree reconstruction, respectively) to show that specific phylogenetic methods would not alter any of our main conclusions.

Phylogenetic analyses were conducted under direct optimization (DO; Wheeler 1996) with POY4.1.1 using the parsimony criterion. All the analyses were performed on a personal computer with 3.25 GB RAM and an Intel Xeon CPU E5410 at 2.33 GHz. For both molecular and developmental data, all events (i.e., substitutions and insertions/deletions) were equally weighted during the search procedure. Data sets were analyzed separately and in combination following search procedures including branch swapping, tree fusing and ratcheting algorithms. Molecular data outnumbered developmental data but analyzing developmental data alone was of prime importance for tree reconciliation analyses (see Tree Comparisons and Reconciliation Analyses Section). In the combined analysis, we did not provide differential weights for any character, meaning that each character (molecular or developmental) participated to the same degree to tree reconstruction. Bootstrap values were calculated in POY with 1000 replicates. In addition, we conducted “static” phylogenetic analyses in a Bayesian framework with a GTR + I +  $\Gamma$  model, running for 10 millions of generations. Details of all phylogenetic analyses are provided in supporting information Appendix S2.

## Ancestral state reconstruction (ASR) and timing of differentiation

ASR of developmental data was carried out in parsimony (under dynamic and static homology) and probabilistic frameworks. Under dynamic homology (parsimony criterion), we used POY 4.1.1 (Varón et al. 2010) and Mesquite (Maddison and Maddison 2009). Ancestral states were reconstructed both on the combined (i.e., developmental and molecular data) and on the molecular topologies. Under static homology (parsimony and Bayesian criteria), we first reconstructed multiple alignments of developmental sequences from clustalW as implemented in BioEdit 7.0.5.1 (Hall 1999). This resulted in a 27 characters long concatenated alignment (supporting information Appendix S3) with 13 positions involving pseudergates, “true” workers or both character states. We then performed parsimony and Bayesian ASR of these 13 developmental characters with Mesquite and BayesTraits v1.0 (Pagel et al. 2004) using a covarion model. In Bayesian analyses, ancestral states for each node of termite phylogeny were reconstructed using the topology found in Bayesian inference (BI). Details are provided in supporting information Appendix S2.

From the ASR, we estimated the evolution of the timing of differentiation between sexual and sterile lines. We used the occurrence of the first nymphal stage as a proxy to determine this differentiation as it is a pivotal stage in termite developmental pathways. Indeed, it is an obligatory stage on the alate (sexual) line but does not occur on the “true” worker (sterile) line. It is thus an accurate estimate of the timing of differentiation between sterile and sexual lines in species with “true” workers. As for species with pseudergates, an instar that can occur through regressive molts from nymph, it is a conservative estimate since timing of differentiation might be later but not sooner.

## Tree comparisons and reconciliation analyses

We assessed the congruence between molecular and developmental topologies using the quartet SD value ( $Q$ ) as implemented in Component 2.0 (Page 1993). We then compared this  $Q$ -value to a random distribution of SD values. To obtain this distribution, we generated 50 trees of 89 leaves and computed  $Q$  values for each comparison, resulting in a distribution of 1225 values ( $50 \times 49/2 = 1225$ ). To compute a  $Q$ -value between two trees, they must have the same terminals. Our molecular tree was, however, taxonomically more diverse than our developmental tree, because some taxa with molecular sequences lacked developmental data. Thus, we pruned the molecular tree so that only relationships between species with developmental data were displayed.

We also conducted reconciliation analyses between developmental and molecular trees. Association between “lineages” can be studied by “reconciling” trees of these “lineages” (Page and Charleston 1998). Reconciling two trees consists of superposing these trees. If both trees are perfectly congruent,

the reconciliation is trivial and implies only codivergence events; on the other hand, any disagreement between the trees will require at least another kind of event to be inferred to resolve this disagreement. These events are: duplication, horizontal transfer, and sorting event. These similar events occur in the different associations (e.g., organism–gene, host–parasite, etc.) but they have different “meanings” as they imply different processes. For instance, a codivergence event in a host–parasite association implies a cospeciation, while the same codivergence event implies a vicariance in an organism–area association. We suggest in Table 1 a developmental interpretation of each event in an organism–ontogeny association (bottom line in italics). We here used Tarzan v0.9 (Merkle and Middendorf 2005) with the following costs to reconcile developmental and molecular trees: cospeciation = 0, duplication = 2, sorting = 1, host switch = 2, and extinction = 2. Such costs are equivalent to the default options of TreeMap, whose jungles algorithms were not applicable for such a large data set (Charleston, pers. comm.).

## RESULTS

### Phylogenetic analyses

The most parsimonious tree, reconstructed from molecular and developmental data in DO ( $L = 17,773$ —Fig. 2), was similar to the most recent molecular hypotheses (Inward et al. 2007; Legendre et al. 2008a) with Isoptera, Hodotermitidae, Kalotermitidae, and Termitidae (neglecting *Sphaerotermites sphaerotherax* spurious position—see Discussion Section) each retrieved monophyletic with high bootstrap values (Fig. 2), whereas Rhinotermitidae and Termopsidae were paraphyletic. The deepest nodes, and notably the occurrence of *Mastotermes darwiniensis* and Kalotermitidae, respectively as first and second diverging lineages within the termites, were also very well supported. The analysis of only molecular data in DO (Fig. 3, left panel) and in a Bayesian framework ( $\ln L = -78354.88$ —supporting information Fig. S1) yielded very similar results. Indeed, Isoptera, Hodotermitidae, Kalotermitidae, and Termitidae were also monophyletic, while Rhinotermitidae and Termopsidae were paraphyletic. The main difference was that, in the Bayesian analysis, Kalotermitidae was not the

second diverging lineage within termites. This position was occupied with the clade (Termopsidae + Hodotermitidae).

### Ancestral state reconstruction (ASR)

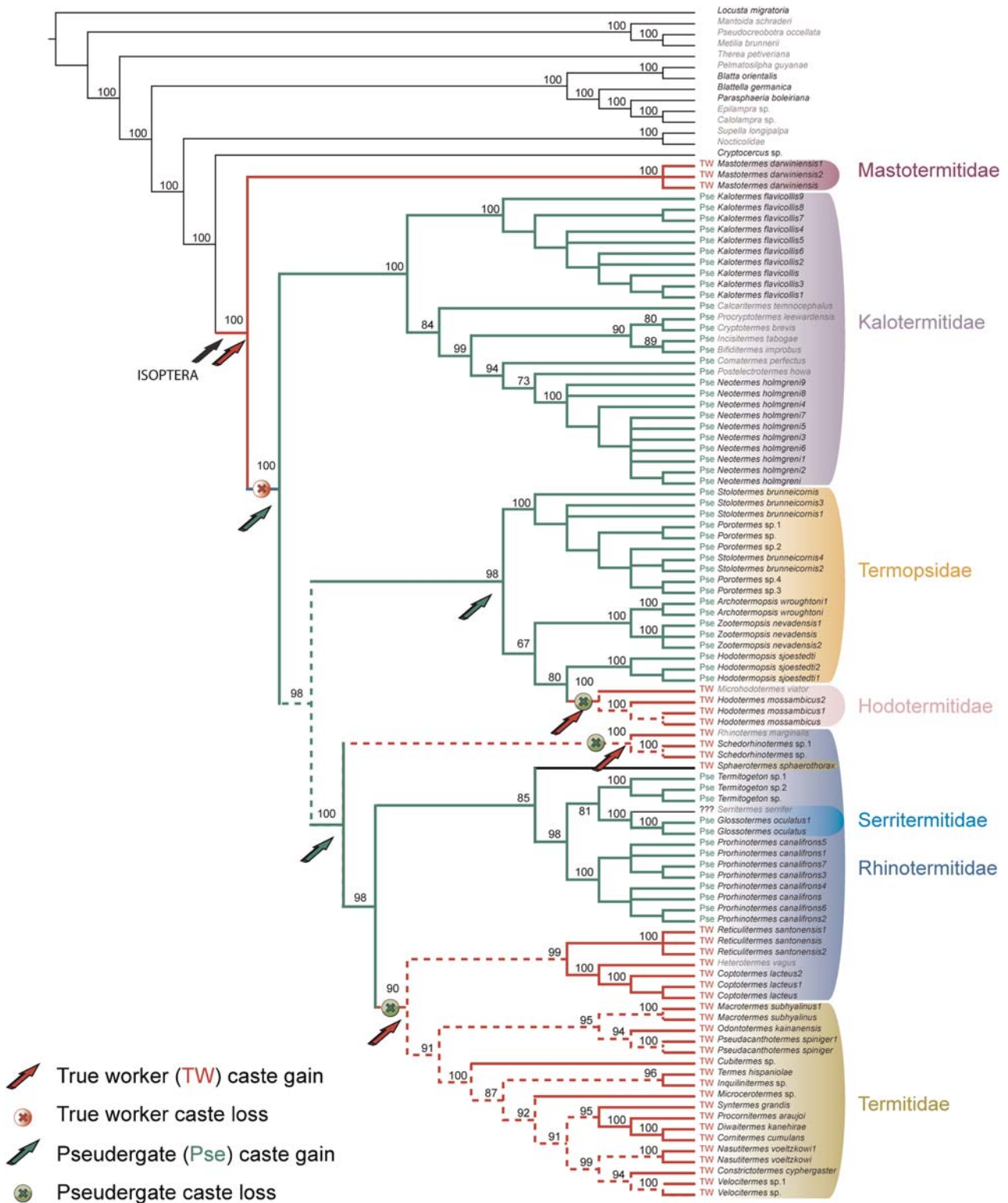
ASR of developmental sequences on the DO tree (Figs. 2 and 4) showed that the “true” worker caste evolved four times convergently: in the termites ancestor, in Hodotermitidae, and twice within the clade (Rhinotermitidae + “Serritermitidae” + Termitidae). This caste was also lost once in the most recent common ancestor of all termites but Mastotermitidae. More precisely, ASR suggested that different developmental events, such as insertion of a worker stage (termite ancestor), deletion of pseudergate stages (subsequent deepest nodes) or substitution from a pseudergate to a nymphal stage [((*Reticulitermes* + *Coptotermes*) + Termitidae)], occurred along developmental pathway evolution in termites (Fig. 4 and supporting information Fig. S2). ASR also revealed that a pseudergate caste evolved either once in the ancestor of all termite species except Mastotermitidae or that it evolved thrice independently. In each case the caste was lost three times: twice within the clade (Rhinotermitidae + Termitidae) and once in Hodotermitidae. Interestingly, our results suggested that pseudergates might have evolved from “true” workers ancestrally. ASR of whole developmental sequences also showed how timing of differentiation between sterile and sexual lines evolved in termites (Fig. 4) from an ancestrally late differentiation (6th molt—linear-like pathway, Fig. 1a—left) to a precocious one (1st and 2nd molt—bifurcated pathway, Fig. 1a—right) in Hodotermitidae, some Rhinotermitidae and Termitidae. Timing of differentiation was also delayed from the 6th to the 7th molt and from the 5th to the 6th molt in some Termopsidae and *Prorhinotermes*, respectively.

When developmental characters were analyzed as “static” characters in a Bayesian framework, we found that a “true” worker caste evolved thrice in termites, and that it was lost four times (supporting information Fig. S1). Similarly, although optimizations of most of the deepest nodes were ambiguous, “true” worker castes evolved at least twice when “static” developmental data were considered in parsimony (supporting information Fig. S3). As for pseudergates, the results were ambiguous. The number of origins and losses of this caste ranged

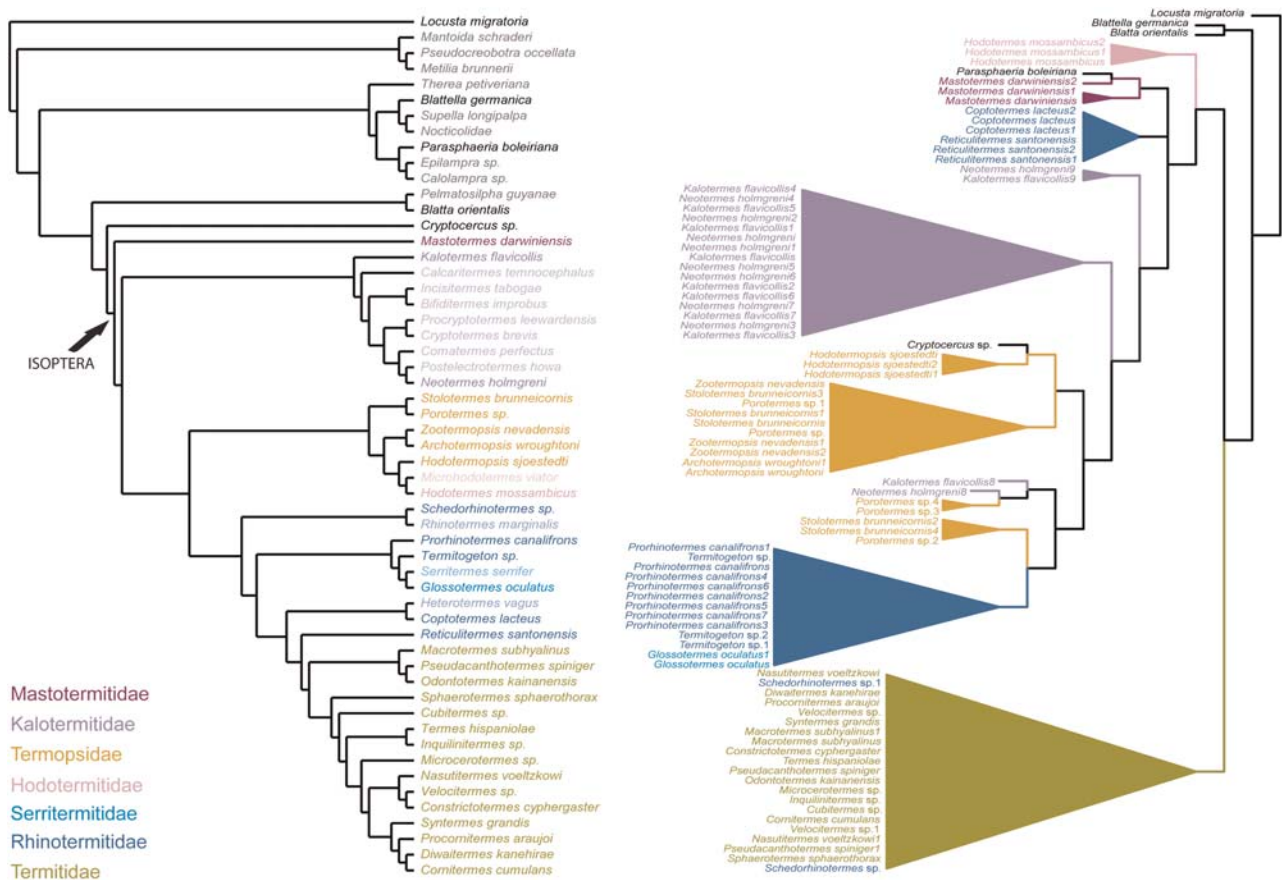
**Table 1. Meanings of reconciliation events between different kinds of trees (modified after Page and Charleston 1998)**

Biological “field”	“Host” tree	“Associate” tree	Codivergence	Duplication	Horizontal transfer	Sorting event
Molecular biology	Organism	Gene	Interspecific coalescence	Gene duplication	Gene transfer	Gene loss
Coevolution	Host	Parasite	Cospeciation	Within host speciation	Host switch	Parasite extinction
Biogeography	Organism	Area	Vicariance	Sympatry	Dispersal	Extinction
<i>Development</i>	<i>Organism</i>	<i>Ontogeny</i>	<i>Inheritance</i>	<i>Increase in plasticity</i>	<i>Convergence, Parallelism</i>	<i>Decrease in plasticity</i>

We suggest here a developmental interpretation of each event in an organism–ontogeny association (bottom line in italics). See text for more details.



**Fig. 2.** Termite phylogeny with worker caste ancestral states reconstructed from direct optimization of whole developmental sequences. These reconstructions imply four independent gains of the “true” worker caste. The phylogenetic relationships in termites are derived from a data set combining seven molecular markers and developmental data ( $L = 17,773$  steps). Species with developmental data are in black. Dashed lines are used when worker stages are not present in the ancestral sequences reconstructed but are suspected in the corresponding lineages (e.g., some lineages wherein soldiers derived from larval stages, like in *Schedorhinotermes* sp.) Bootstrap values above 50 are displayed at each corresponding node.



**Fig. 3.** Comparison of molecular and developmental trees. (Left) Most parsimonious tree reconstructed from seven molecular markers in Direct Optimization ( $L = 17,616$  steps). Taxa with developmental data are in bold. (Right) Strict consensus of the two most parsimonious trees obtained from developmental data alone ( $L = 109$  steps). The high congruence (Quartet SD value  $Q = 0.384$  versus  $Q = 0.667$  for 1225 random comparisons) at the familial level between this tree and the molecular topology indicate the large phylogenetic signal conveyed by developmental data in termites. Polyphyly of Kalotermitidae and Termopsidae species suggest, for each lineage, an ancestral flexibility, a result confirmed by reconciliation analyses.

from one to three in parsimony, whereas ASR in BI suggested three origins. We summarized the results of the different ASR in Table 2.

### Tree comparisons and reconciliation analyses

To assess the phylogenetic signal conveyed by developmental data, we compared the topologies obtained in DO from the molecular and developmental data sets with the quartet SD value  $Q$ . This statistics revealed that, despite a low resolution at the intra-familial level, the developmental tree was more congruent with the molecular one than by chance alone (quartet SD value  $Q = 0.384$  vs.  $Q = 0.667$  for 1225 random comparisons), and thus that developmental sequences conveyed a strong phylogenetic signal, at least at the familial level.

We also conducted reconciliation analyses between developmental and molecular trees. Multiple optimal solutions (cost = 142) were found with 22–23 cospeciation events, 10–12 sorting

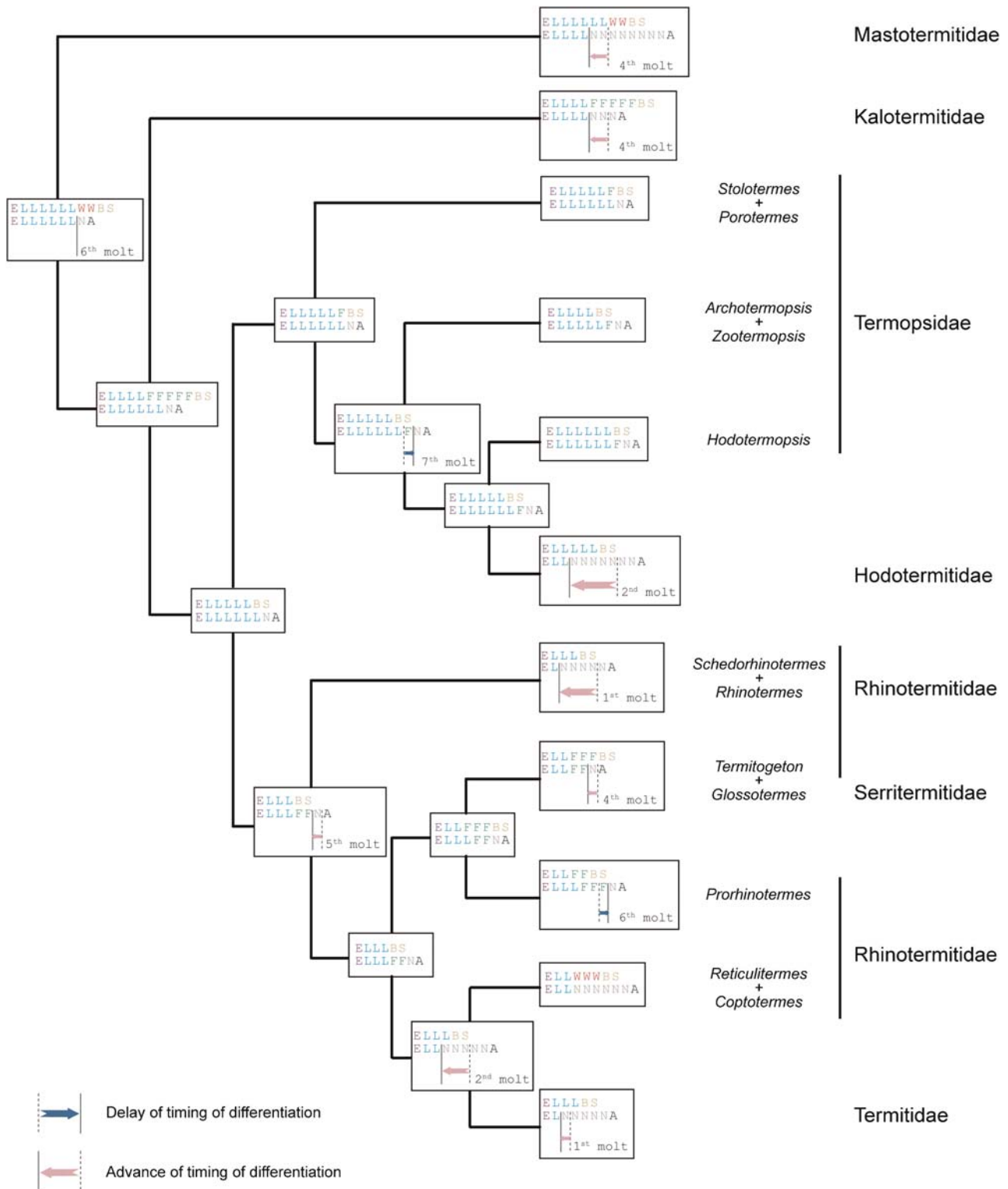
events, 20–25 duplications, and 41–46 switches. One optimal representative solution with 22 cospeciation and 20 duplication events was randomly chosen to be displayed in Table 3. Most duplication events occurred in Kalotermitidae, Termopsidae, and *Prorhinotermes* clades. Such results signified that developmental plasticity increased independently in these lineages, and not ancestrally to termites as previously suggested by Noirot (1985a). The latter hypothesis would have implied that multiple duplication events would have occurred ancestrally to all termites, a prediction we did not find here.

## DISCUSSION

### Termites phylogenetic relationships

Our phylogenetic hypothesis, reconstructed from molecular and developmental data, is similar to the most recent molecular hypotheses (Inward et al. 2007; Legendre et al. 2008a) with





**Fig. 4.** Developmental pathways evolution. The timing of differentiation between sexual and sterile lines was ancestrally late (6th molt—linear-like pathway) and occurred much earlier (pink arrows—bifurcated pathway) convergently in Hodotermitidae and in the clade (Rhinotermitidae + “Serritermitidae” + Termitidae). A later timing of differentiation (blue arrows) occurred in Termopsidae and *Prorhinotermes*, two lineages with pseudergates. Ancestral sequences have been reconstructed on the combined topology and are presented here on a simplified tree. Both sequences ending in soldier (above) and alate stages (below) are shown. In each box, a plane line indicates the current state of timing of differentiation, whereas a dashed line represents the ancestral timing of differentiation (i.e., the state at the immediate previous node). The molt at which the timing of differentiation occurs is specified ancestrally and for each node where a change is postulated; this timing was hypothesized using the nymphal stage as a proxy (see text for further explanations). Alphabet as in Fig. 1.

**Table 2. Summary of the ancestral state reconstructions of both worker castes**

Evolutionary events	Ancestral state reconstruction methods		
	Direct optimization	Parsimony-static	Bayes
True workers			
# of origins	4	2–4	3
# of losses	1	0–3	4
Pseudergates			
# of origins	1–3	1–3	3
# of losses	3	1–3	3

Multiple origins of the “true” worker caste are inferred whatever methods are used.

Isoptera, Hodotermitidae, and Kalotermitidae found monophyletic with high support values, whereas Rhinotermitidae and Termopsidae are paraphyletic. Nevertheless, three unexpected results deserve to be discussed: the position of *S. sphaerotherax* and the placement of Serritermitidae and Kalotermitidae.

In a previous study (Legendre et al. 2008a), we found two alternate positions for the Serritermitidae according to analysis parameters. Either Serritermitidae was sister group to (Rhinotermitidae + Termitidae), or it was nested within Rhinotermitidae as sister taxon to *Termitogeton*. Sequencing *G. oculatus*, a species hypothesized to belong to Serritermitidae (Cancello and DeSouza 2004), we bring here new data supporting the latter hypothesis. This result contradicts the family status of the Serritermitidae but supports a single origin of a pseudergate caste for *Glossotermes*, *Termitogeton*, and *Prorhinotermes*, which is of prime importance for our understanding of caste evolution in termites. We now have at hand a whole clade to study pseudergate evolution and to compare with Kalotermitidae and Termopsidae (Fig. 2).

The second unexpected result refers to the position of *S. sphaerotherax*. Even though the hypothesis that *S. sphaer-*

*otherax* belongs to Macrotermitinae has been doubted (Aanen et al. 2002), all known morphological and behavioral characters support its inclusion within Termitidae so that its placement in some of the present analyses within Rhinotermitidae as a sister-taxon to the clade ((*Termitogeton* sp., “Serritermitidae”), *Prorhinotermes canalifrons*) is doubtful. Our previous molecular analyses (Legendre et al. 2008a) and the present Bayesian result (supporting information Fig. S1), in which *S. sphaerotherax* is nested within Termitidae, also validate this interpretation. So far, *S. sphaerotherax* phylogenetic placement is, however, unstable, which can probably be explained by a lack of data. Indeed, ~80% of the molecular data are missing for this taxon. To complete the sampling of the genetic loci for *S. sphaerotherax* is thus required before drawing any conclusion about its phylogenetic position.

Finally, as already found previously (Legendre et al. 2008a), our phylogenetic result contradicts Inward et al.’s work (2007), as to the position of Kalotermitidae. Inward et al. (2007) and Legendre et al. (2008a) studies differ in taxonomic and character samplings, and in phylogenetic reconstruction methods as well. It is therefore not so surprising that they have ended up with different conclusions. But it should be noted that, first, both studies are compatible with convergent origins of “true” worker castes. Second, both topologies are poorly supported when it comes to the phylogenetic position of Kalotermitidae [Bremer value of 2 (Inward et al. 2007), or position methodology-dependent (Legendre et al. 2008a)]. Thus, the next taxonomic sampling effort in termite phylogenetics should focus on Kalotermitidae and Termopsidae species to decipher inter-familial relationships.

### Caste and ontogenetic pathways evolution

By analyzing entire developmental sequences instead of mapping binary characters, we develop here a more detailed pattern of caste evolution than in previous studies. Whereas it was impossible with binary characters to reconstruct an unambiguous scenario for the evolution of the “true” worker

**Table 3. Reconciliation analyses between species and ontogenetic trees**

Clade	# of cospeciation events (inheritance)	# of duplication events (increase in plasticity)
Outgroups	2	0
<i>Mastotermes</i> lineage	0	1
Kalotermitidae	4	5
(Termopsidae + Hodotermitidae + Rhinotermitidae + Serritermitidae + Termitidae)	1	0
Termopsidae	6	4
Hodotermitidae	0	2
(Serritermitidae + <i>Prorhinotermes</i> + <i>Termitogeton</i> )	2	5
( <i>Heterotermes</i> + <i>Coptotermes</i> + <i>Reticulitermes</i> + Termitidae)	7	3

caste (Thompson et al. 2000, 2003; Grandcolas and D’Haese 2002, 2004; Inward et al. 2007), we show here that “true” workers, individuals belonging to a very specialized caste involving altruism, evolved multiple times in termites (from two to four independent origins). For pseudergates, results are less straightforward as both a single origin and multiple origins are postulated according to the analyses. The latter hypothesis with convergent origins is congruent with Parmentier and Roisin’s work (2003), which emphasized the developmental differences between *Termitogeton planus* on one hand, and Kalotermitidae and Termopsidae on the other hand. These differences would be explained by different evolutionary origins and through different developmental processes, a conclusion that would have been impossible to reach if binary characters had been used. Moreover, our results suggest that this caste might have evolved from “true” workers (ancestrally or within Rhinotermitidae according to ASR methods), an unexpected result contradicting Cope’s rule of specialization, wherein the acquisition of a complex and highly specialized character is considered an evolutionary dead end (Cope 1896). Again, this would be congruent with the study on *T. planus*, whose second instar larvae have wing buds, but not the subsequent instars. The occurrence of wing buds in the second instar would be reminiscent of a typical “true” worker development (like the one found in Termitidae and some Rhinotermitidae species) with an early differentiation of nymphs, which are characterized by wing buds. Additionally, reconciliation analyses suggest that the developmental data set has a strong phylogenetic signal (Fig. 3) and that developmental plasticity is not an ancestral feature to termites but evolved convergently in different lineages (e.g., Kalotermitidae, *Prorethotermitidae*—Table 3). Thus, our results mostly confirm Noirot’s hypothesis (1985a, b—multiple origins of the “true” worker caste, ancestral high flexibility and linear-like pathway), with multiple origins of a “true” worker caste in termites, which would have evolved from an ancestor with a weakly flexible and linear developmental pathway (Fig. 4). Our results about ancestral flexibility are, however, more congruent with Watson and Sewell’s hypothesis (1985—a single origin, ancestral medium flexibility and bifurcated pathway). High flexibility would have been acquired convergently mostly in Kalotermitidae, Termopsidae, and (*Termitogeton* + Serritermitidae + *Prorethotermitidae*), three lineages in which species do not build “true” nests but live within their relatively ephemeral food source, the dead wood. Once their food is almost consumed, some individuals (alates) need to disperse to ensure colony survival, selecting therefore for a high flexibility and a late timing of differentiation. Conversely, Hodotermitidae, Termitidae, and some Rhinotermitidae species (Grassé 1986) live in a more perennial environment as they build “true” nests (e.g., cathedral mounds), and they have weakly flexible post-embryonic developmental pathways. This corroborates a classical evo-devo hypothesis (Nijhout 2003), wherein environmental cues [here, dependent of species foraging type

sensu Abe (1987)] are directly related to developmental flexibility. These results suggest that acquisition of stable conditions has been a prime mover for the evolution of irreversibly sterile caste. This is congruent with recent hypotheses and models about eusociality emphasizing the role of ecological pressures as a key phenomenon to drive evolution (Johns et al. 2009; Lion and Gandon 2010).

While we are still puzzled by the underlying processes responsible for sterile caste evolution in diploid organisms under natural selection, the phylogenetic analysis of developmental sequences reveals that such a Darwinian paradox occurred repeatedly in termites. These results will help identify new model termite species for evo-devo studies, which is pivotal for any major advances in the field (Jenner and Wills 2007). For instance, *Hodotermitidae* *mossambicus*, with its convergent “true” workers origin and timing of differentiation, appears as an adequate taxon to compare with Termitidae species. Similarly, *P. canalifrons* should be a valuable model species to start with on pseudergate evolution and to compare to Kalotermitidae species. In addition, we show here that obligatory sterile castes do not appear necessarily as an evolutionary dead-end, with transitions from “true” worker to pseudergate, contradicting Cope’s rule of specialization (1986). These results reveal that caste and developmental pathway evolution in termites is versatile.

## ACKNOWLEDGMENTS

We thank Andres Varon and Ward Wheeler who gave us advice with POY4, especially with the “custom\_alphabet” command; Dan Janies offered us the possibility to run several memory- and time-consuming preliminary analyses at the Ohio Supercomputer Center; Christian Bordereau provided comments on several previous drafts; Yves Roisin commented previous drafts and sent us fresh material of *G. oculatus*. André Nel, Patricia Nel and Tony Robillard commented the last draft. We also warmly thank an anonymous reviewer for its useful and constructive comments on the manuscript. F.L. was successively supported by the French Ministère de la Recherche (PhD grant), by the Muséum national d’Histoire naturelle of Paris (Programme Pluriformation “Etat et structure phylogénétique de la biodiversité actuelle et fossile,” P. Janvier) and by a bourse Lavoisier (Ministère des Affaires Etrangères et Européennes).

## REFERENCES

- Aanen, D. K., Eggleton, P., Rouland-Lefevre, C., Guldborg-Froslev, T., Rosendahl, S., and Boomsma, J. J. 2002. The evolution of fungus-growing termites and their mutualistic fungal symbionts. *Proc. Natl. Acad. Sci. USA* 99: 14887–14892.
- Abbott, A., and Tsay, A. 2000. Sequence analysis and optimal matching methods in sociology. Review and prospect. *Sociol. Method. Res.* 29: 3–33.
- Abe, T. 1987. Evolution of life types in termites. In S. Kawano, J. H. Connell, and T. Hidaka (eds.). *Evolution and Coadaptation in Biotic Communities*. University of Tokyo Press, Tokyo, pp. 125–148.
- Bartz, S. H. 1979. Evolution of eusociality in termites. *Proc. Natl. Acad. Sci.* 76: 5764–5768.

- Bininda-Emonds, O. R. P., Jeffery, J. E., Coates, M. I., and Richardson, M. K. 2002. From Haeckel to event-pairing: the evolution of developmental sequences. *Theory Biosci.* 121: 297–320.
- Boomsma, J. J. 2009. Lifetime monogamy and the evolution of eusociality. *Phil. Trans. R. Soc. Lond. B* 364: 1–17.
- Bourguignon, T., Sobotnik, J., Hanus, R., and Roisin, Y. 2009. Developmental pathways of *Glossotermes oculatus* (Isoptera, Serritermitidae): at the cross-roads of worker caste evolution in termites. *Evol. Dev.* 11: 659–668.
- Canello, E. M., and DeSouza, O. 2004. A new species of *Glossotermes* (Isoptera): reappraisal of the generic status with transfer from the Rhinotermitidae to the Serritermitidae. *Sociobiology* 44: 1–19.
- Cope, E. D. 1896 *The Primary Factors of Organic Evolution*. Open Court Publishing, Chicago.
- Costa, J. T. 2006. *The other Insect societies*. The Belknap Press of Harvard University Press, Cambridge, Massachusetts, and London, England.
- Darwin, C. 1859 *On the Origin of Species by Means of Natural Selection, or Preservation of Favored Races in the Struggle for Life*. Murray, London.
- Duffy, J. E., and Macdonald, K. S. 2010. Kin structure, ecology and the evolution of social organization in shrimp: a comparative analysis. *Proc. R. Soc. Lond. B* 277: 575–584.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792–1797.
- Grandcolas, P., and D’Haese, C. 2002. The origin of a “true” worker caste in termites: phylogenetic evidence is not decisive. *J. Evol. Biol.* 15: 885–888.
- Grandcolas, P., and D’Haese, C. 2003. Testing adaptation with phylogeny: How to account for phylogenetic pattern and selective value together? *Zool. Scr.* 32: 483–490.
- Grandcolas, P., and D’Haese, C. 2004. The origin of a ‘true’ worker caste in termites: mapping the real world on the phylogenetic tree. *J. Evol. Biol.* 17: 461–463.
- Grandcolas, P., Guilbert, E., Robillard, T., D’Haese, C., Murienne, J., and Legendre, F. 2004. Mapping characters on a tree with or without the outgroups. *Cladistics* 20: 579–582.
- Grandcolas, P., Nattier, R., Legendre, F., and Pellens, R. 2011. Mapping extrinsic traits such as extinction risks or modelled bioclimatic niches on phylogenies: does it make sense at all? *Cladistics* 27: 181–185.
- Grassé, P. P. 1986. *Termitologie: Comportement, Socialité, Ecologie, Evolution, Systématique*. Masson, Paris: 716 p.
- Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41: 95–98.
- Hamilton, W. D. 1964. The genetical evolution of social behaviour. I and II. *J. Theor. Biol.* 7: 1–52.
- Hay, B., Wets, G., and Vanhoof, K. 2004. Mining navigation patterns using a sequence alignment method. *Knowl. Inf. Syst.* 6: 150–163.
- Howard K. J., and Thorne, B. L. 2011. Eusocial evolution in termites and Hymenoptera. In D. E. Bignell, Y. Roisin, and N. Lo (eds.). *Biology of Termites: A Modern Synthesis*. Springer, London, pp. 97–132.
- Inward, D., Vogler, A. P., and Eggleton, P. 2007. A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspects of their evolutionary biology. *Mol. Phylogen. Evol.* 44: 953–967.
- Jeffery, J. E., Bininda-Emonds, O. R. P., Coates, M. I., and Richardson, M. K. 2005. A new technique for identifying sequence heterochrony. *Syst. Biol.* 54: 230–240.
- Jeffery, J. E., Richardson, M. K., Coates, M. I., and Bininda-Emonds, O. R. P. 2002. Analyzing developmental sequences within a phylogenetic framework. *Syst. Biol.* 51: 478–491.
- Jenner R. A., and Wills, M. A. 2007. The choice of model organisms in evo-devo. *Nat. Rev. Genet.* 8: 311–314.
- Johns, P. M., Howard, K. J., Breisch, N. L., Rivera, A., and Thorne, B. L. 2009. Nonrelatives inherit colony resources in a primitive termite. *Proc. Natl. Acad. Sci. USA* 106: 17452–17456.
- Johnstone, R. A., Cant, M. A., and Field, J. 2012. Sex-biased dispersal, haplodiploidy and the evolution of helping in social insects. *Proc. R. Soc. Lond. B* 279: 787–793.
- Korb, J., and Hartfelder, K. 2008. Life history and development—a framework for understanding developmental plasticity in lower termites. *Biol. Rev.* 83: 295–313.
- Legendre, F., Robillard, T., Desutter-Grandcolas, L., Whiting, M. F., and Grandcolas, P. 2008b. Phylogenetic analysis of non-stereotyped behavioural sequences with a successive event-pairing method. *Biol. J. Linn. Soc.* 94: 853–867.
- Legendre, F., Whiting, M. F., Bordereau, C., Canello, E. M., Evans, T. A., and Grandcolas, P. 2008a. The phylogeny of termites (Dictyoptera: Isoptera) based on mitochondrial and nuclear markers: Implications for the evolution of the worker and pseudergate castes, and foraging behaviors. *Mol. Phylogen. Evol.* 48: 615–627.
- Lion S., and Gandon, S. 2010. Life history, habitat saturation and the evolution of fecundity and survival altruism. *Evolution* 64: 1594–1606.
- Maddison, W. P., and Maddison, D. R. 2009. Mesquite: a modular system for evolutionary analysis. Version 2.71. <http://mesquiteproject.org>
- Merkle, D., and Middendorf, M. 2005. Reconstruction of the cophylogenetic history of related phylogenetic trees with divergence timing information. *Theory Biosci.* 123: 277–299.
- Myles, T. G. 1986. Reproductive soldiers in the Termopsidae (Isoptera). *Pan-Pac. Entomol.* 62: 293–299.
- Nalepa, C. A. 2010. Altricial development in subsocial cockroach ancestors: foundation for the evolution of phenotypic plasticity in termites. *Evol. Dev.* 12: 95–105.
- Nijhout, F. H. 2003. Development and evolution of adaptive polyphenisms. *Evol. Dev.* 5: 9–18.
- Noirot, C. 1985a. The caste system in higher termites. In J. A. L. Watson, B. M. Okot-Kotber, and C. Noirot, (eds.). *Caste Differentiation in Social Insects*. Pergamon, Oxford: pp. 75–86.
- Noirot, C. 1985b. Pathways of caste development in the lower termites. In J. A. L. Watson, B. M. Okot-Kotber, and C. Noirot, (eds.). *Caste Differentiation in Social Insects*. Pergamon, Oxford: pp. 75–86.
- Noirot, C., and Pasteels, J. M. 1987. Ontogenetic development and evolution of the worker caste in termites. *Experientia* 43: 851–952.
- Nonacs, P. 2011. Kinship, greenbeards, and runaway social selection in the evolution of social insect cooperation. *Proc. Natl. Acad. Sci. USA* 108: 10808–10815.
- Nowak, M. A., Tarnita, C. E., and Wilson, E. O. 2010. The evolution of eusociality. *Nature* 466: 1057–1062.
- Page, R. D. M. 1993. *Component, Tree Comparisons software for Microsoft Windows, Version 2.0*. The Natural History Museum, London.
- Page, R. D. M., and Charleston, M. A. 1998. Trees within trees: phylogeny and historical associations. *Trends Ecol. Evol.* 13: 356–359.
- Pagel, M., Meade, A., and Barker, D. 2004. Bayesian estimation of ancestral character states on phylogenies. *Syst. Biol.* 53: 673–684.
- Parmentier, D., and Roisin, Y. 2003. Caste morphology and development in *Termitogeton* nr. *planus* (insecta, isopectera, rhinotermitidae). *J. Morphol.* 255: 69–79.
- Pellens, R., et al. 2007. The evolutionary transition from subsocial to eusocial behavior: phylogenetic and ecological evidence for modification of the “shift-in-dependent-care” hypothesis with a new prototermite model. *Mol. Phylogen. Evol.* 43: 616–626.
- Prestwich, G. D. 1984. Defense mechanisms of termites. *Annu. Rev. Entomol.* 29: 201–232.
- Robillard, T., Legendre, F., Desutter-Grandcolas, L., and Grandcolas, P. 2006. Phylogenetic analysis and alignment of behavioral sequences by direct optimization. *Cladistics* 22: 602–633.
- Roisin, Y. 1994. Intragroup conflicts and the evolution of sterile caste in termites. *Am. Nat.* 143: 751–765.
- Roisin, Y., and Korb, J. 2010. Social organisation and the status of workers in termites. In D. E. Bignell, Y. Roisin, and N. Lo, (eds.). *Biology of Termites: A Modern Synthesis*. Springer, London, pp. 133–164.
- Rousset, F., and Lion, S. 2011. Much ado about nothing: Nowak et al.’s charge against inclusive fitness theory. *J. Evol. Biol.* 24: 1386–1392.
- Schulmeister, S., and Wheeler, W. C. 2004. Comparative and phylogenetic analysis of developmental sequences. *Evol. Dev.* 6: 50–57.
- Smith, K. K. 1997. Comparative patterns of craniofacial development in eutherian and metatherian mammals. *Evolution* 51: 1663–1678.
- Thompson, G. J., Kitade, O., Lo, N., and Crozier, R. H. 2000. Phylogenetic evidence for a single, ancestral origin of a ‘true’ worker caste in termites. *J. Evol. Biol.* 13: 869–881.

- Thompson, G. J., Kitade, O., Lo, N., and Crozier, R. H. 2003. On the origin of termite workers: weighing up the phylogenetic evidence. *J. Evol. Biol.* 17: 217–220.
- Thompson, J. D., Higgins, D. G., and 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.
- Varón, A., Vinh, L. S., and Wheeler, W. C. 2010. POY version 4: phylogenetic analysis using dynamic homologies. *Cladistics* 26: 72–85.
- Velhagen, W. A. 1997. Analyzing developmental sequences using sequence units. *Syst. Biol.* 46: 204–210.
- Watson, J. A. L., and Sewell, J. J. 1985. Caste development in *Mastotermes* and *Kaloterme*s: which is primitive? In J. A. L. Watson, B. M. Okot-Kotber, and C. Noirot (eds.). *Caste Differentiation in Social Insects*. Pergamon, Oxford, pp. 27–40.
- Wheeler, W. 1996. Optimization alignment: the end of multiple sequence alignments in phylogenetics? *Cladistics* 12: 1–9.
- Wheeler, W. C. 2003. Search-based optimization. *Cladistics* 19: 348–355.
- Wilson, E. O. 1971. *The Insect Societies*. The Belknap Press of the Harvard University Press, Cambridge, Massachusetts.