A total evidence phylogeny for the processionary moths of the genus *Thaumetopoea* (Lepidoptera: Notodontidae: Thaumetopoeinae)

Andrea Basso[1], Enrico Negrisolo[2*], Alberto Zilli[3], Andrea Battisti[1], Pierfilippo Cerretti[4]

[1] Department of Agronomy Food Natural Resources Animals and Environment (DAFNAE), University of Padua, Agripolis, 35020, Legnaro, Padua, Italy;

[2] Department of Comparative Biomedicine and Food Science (BCA), University of Padua, Agripolis, 35020, Legnaro, Padua, Italy;

[3] Department Life Sciences - Division Insects, Natural History Museum, London, Cromwell Road, London SW7 5BD, UK;

[4] Department of Biology and Biotechnologies 'Charles Darwin', Sapienza University of Rome, Piazzale A. Moro 5, 00185, Rome, Italy.

* corresponding author,

Andrea Basso: andrea.basso.8@phd.unipd.it; andrea.basso.1988@gmail.com

Enrico Negrisolo: enrico.negrisolo@unipd.it

Alberto Zilli: a.zilli@nhm.ac.uk

Andrea Battisti: andrea.battisti@unipd.it

Pierfilippo Cerretti: pierfilippo.cerretti@uniroma1.it

Short running title: *Thaumetopoea* total evidence phylogeny

Abstract

Processionary moths belong to a group of about 100 species well known for their social behaviour and their urticating systems. The genus *Thaumetopoea s. lat.* includes about 15 species and has been divided into three genera (*Helianthocampa*, *Thaumetopoea s. str*., and *Traumatocampa*) in the last revision, based on key morphological features of the adults and on the host plants of the larvae. We performed a total evidence approach to resolve the phylogeny of the genus *Thaumetopoea s. lat.*, analysing all valid taxa included in this group, plus a broad array of close relatives. *Thaumetopoea* resulted monophyletic and supported by several apomorphies. Further subclades corroborated by synapomorphies were identified. Our phylogeny suggests that *Thaumetopoea* must be regarded as a single genus. The mapping of key life history traits on the total evidence tree allowed to sketch a plausible identikit of the *Thaumetopoea* ancestor and to track the evolution of the genus. The ancestor originated in the eastern Mediterranean area, and used broadleaves as host plants. Subsequently, a switch to conifers occurred, just once, in a large subclade. The ancestor pupated in the soil, like several current species, but in few taxa this trait was lost, together with the related morphological adaptations.

41 **Table of contents**

Introduction

Processionary moths belong to a group of insect herbivores well known since the ancient Greek and Roman times for their association with the host plants, their special behaviour, and their urticating and envenomation power (Roques and Battisti, 2015). All processionary moths are now included in Thaumetopoeinae, a clade of Notodontidae (Lepidoptera) (Miller, 1991; Zahiri *et al.*, 2011, 2013). In the past, this group was treated at family rank with three subfamilies: Thaumetopoeinae (mainly Palaearctic), Anaphinae (Afrotropical) and Epicominae (Australasian) (Kiriakoff, 1970). However, this view is now superseded by results obtained from morphological and molecular based phylogenetic analyses (Miller, 1991; Zahiri *et al.*, 2011, 2013). At present, Thaumetopoeinae account for 106 valid species split in 20 genera (Schintlmeister, 2013).

The processionary moths owe their name to the typical processions made by the larvae when they move in lines or groups to forage on trees or to pupation sites. A few species of processionary moths are plant pests and cause outbreaks on trees and shrubs of both broadleaves and coniferous trees and shrubs in Africa (Wagner and Cobbinah, 2013), Asia (Rahman and Chaudhry, 1992), Australia (Floater and Zalucki, 2000), and Europe (Jacquet *et al.*, 2012). All species are protected against vertebrate predators by urticating setae either as larvae or adults (Battisti *et al.*, 2017), and these setae may threaten animal and human health (Battisti *et al.*, 2011). In spite of these problems, in Africa some species are farmed for silk production or used as human food (van Huis, 2003; Schabel, 2006).

Most information on life history and taxonomy concerning Thaumetopoeinae relates to the genus *Thaumetopoea*. Hübner ([1820], in 1816-[1826]) erected *Thaumetopoea* to include two species, namely *Phalaena processionea* Linnaeus, 1758, the type species associated with oaks in central and southern Europe, and *Bombyx pityocampa* [Denis & Schiffermüller], 1775, associated with pine in southern Europe and the Mediterranean. More taxa were added later and currently fifteen species are included in this genus (Table 1) (Schintlmeister, 2013). The Palaearctic species of *Thaumetopoea* have been reviewed by Agenjo (1941) and de Freina and Witt (1985, 1987). Agenjo (1941) maintained the split of the genus into the two subgenera *Thaumetopoea* and *Traumatocampa* done by Wallengren (1871) according to the absence/presence of a toothed protuberance on the frontal part of the head (crest) and of a foretibia claw (spine) in the adults (*sensu* Hogue, 1963) (Table 2), whereas de Freina and Witt (1985) raised *Traumatocampa* Wallengren, 1871 to the status of distinct genus. Furthermore, they erected the new genus *Helianthocampa* to accommodate only *Bombyx herculeana* Rambur, 1840 a species feeding on another group of host plants, which was previously included in *Traumatocampa* (Agenjo, 1941) (Table 2). Both revisions did not use a phylogenetic approach and the suggested taxonomic changes were not based on rigorous analyses corroborating the monophyly of their genera. In spite of that, the three genera were retained as valid in the most recent catalogue

of Notodontidae (Schintlmeister, 2013). However, a recent phylogeny based on nucleotide sequence data on a subset of species indicated that the split of *Thaumetopoea* in three distinct genera is untenable, and provided evidence for inclusion of all the species in a single genus *Thaumetopoea s.lat.* until new revision (Simonato *et al.*, 2013).

We consider here all species of *Thaumetopoea s. lat.* occurring in the Palaearctic and Afrotropical regions (Table 1). They extend altogether over an area ranging from the Atlantic coast of the Iberian Peninsula to the Indian region of Jammu and Kashmir and from Scandinavia to southern Africa, where the species are associated with very diverse habitats and host plants. Our analysis included all taxa that were variously granted with independent status according to revisions and recent papers, for a total of 21 taxa. As outgroups we selected 16 species of other Thaumetopoeinae, encompassing a representative sample of the taxonomic diversity of the subfamily (Table 1, and Table S1) (Schintlmeister, 2013). Thus the final data set contained 37 taxa included in seven genera. We analysed morphological and molecular data separately and in combination. The major aim of our research was to produce a robust phylogeny of *Thaumetopoea s. lat.*, to lay out a framework for a thorough taxonomic revision of the genus. In addition, we aimed at tracking the origin of key life history traits of the group to draw a plausible profile of the common ancestor of processionary moths and to mark the main changes that led to the diversity currently shown by this group.

Materials and methods

Taxonomic sampling

The identification at species level of the specimens was done according to the dichotomous keys provided by Agenjo (1941) and Kiriakoff (1970). When possible, the specimens were compared with the types of various species. All the 15 species of *Thaumetopoea s. lat.* listed by Schintlmeister (2013) were analysed. The very recently described *Thaumetopoea loxostigma* Hacker, 2016 could not be included as the only extant specimen, viz. the holotype, was not available. According to the original description, *T. loxostigma* is closely related to the *Thaumetopoea apologetica – Thaumetopoea jordana* group (Hacker, 2016).

Five subspecies other than the nominotypical are currently recognised within *Thaumetopoea s. lat.* (Schintlmeister, 2013), namely *T. apologetica abyssinica*, *T. herculeana judaea*, *T. processionea pseudosolitaria*, *T. solitaria iranica*, and *T. pityocampa orana*. These taxa were included in our analysis because they show peculiar habitus and/or separate geographical distributions, as well as differences in habitat preference from the nominotypical subspecies (Kiriakoff, 1970). Finally, we included in the analysis the informal taxon named by Kerdelhué *et al.* (2009) as *Thaumetopoea pityocampa* ENA (Eastern-North African clade) as it markedly differs from nominal *T. pityocampa* on genetic grounds. Further molecular data, provided by Simonato *et al.* (2013), support the view that *T. pityocampa* ENA actually represents a distinct taxon.

The present study is based on the analysis of specimens deposited in the Department of Agronomy, Food, Natural resources, Animals and Environment of the University of Padua; Museo Civico di Zoologia of Rome; Museo di Zoologia - University of Rome 'Sapienza'; Natural History Museum, London; private collection of A. Schintlmeister (Dresden); Royal Belgian Institute of Natural Sciences of Brussels; Royal Museum of Central Africa of Tervuren; The Bavarian State Collection of Zoology of Munich and Witt Museum of Munich. The complete list of specimens analysed in this paper is provided in Table S1.

Specimens dissection and recording of morphological traits

The abdomen was removed from the specimens listed in Table S1 and processed according to the non-destructive dissection protocol described below, in order to prepare genitalia for successive examination. The abdomens were digested overnight in an incubator kept at 37 °C, immersed, within Eppendorf vial, in 180 µl of ATL buffer and 20 µl Proteinase K and mix. Before and after the incubation the abdomens were vortexed at 350 rpm for 15'. Successively, the abdomen and the solution were processed separately. The abdomens were dissected following the protocol given by Robinson (1976). All forceps and scissors were sterilized before cutting each abdomen. Pipet tips

were discarded and changed every time. The genitalia and the abdomens were stained in saturated chlorazol black (75% ethanol) for 30", soaked in absolute ethanol and eventually mounted on slides with Euparal medium. All slides were labelled according to Robinson (1976). Legs and palps were macerated in 10% KOH solution at 50°C for 5-10 minutes, cleaned in water, then stained and mounted on slides as above.

All morphological observations were taken from pinned specimens of adult moths. Pictures of the morphological characters were taken with a Canon Eos (600D) and a Lumix camera (DMC F200) equipped with additional lens (DMW-LC55) using a led light chamber (Figs 1-3 and S3). Composite focus-stacking images were produced from multiple images captured using AxioCam (MRc5), software Axiovision SE64, (v4.9.1) mounted on a Lumar.v12 Zeiss (Carl Zeiss Microscopy GmbH; Jena, Germany) stereomicroscope, and processed with Photoshop CS6 (v13.0) (Adobe System Incorporated, San Jose, CA, USA). CorelDraw X5 (v15.1) (Corel Corporation, Ottawa, ON, Canada) was used to create labels and drawings. Traits analysed in the present paper are listed in Figs 2, 3 and S3.

Molecular methods

Parallel to the abdomen dissection of single specimens, the total DNA was extracted with QIAGEN© DNeasy blood and tissue kit (Knölke *et al.*, 2005). This procedure did not provide always DNA samples of good quality because of the very different level of tissue conservation of the studied specimens. Thus, molecular analysis was restricted to the best preserved (for molecular purposes) samples.

The DNA extracted during dissection of the genitalia was amplified with universal primers (LCO1490/HCO2198) for the barcode portion of the mitochondrial gene *cox1* (Folmer *et al.*, 1994). Extracts were concentrated with Amicon® Ultra-0.5 Centrifugal Filter Devices, in order to bring all samples to a minimum detectable concentration for the amplification. Each Polymerase Chain Reaction (PCR) contained: 2 µl DNA template (≈15ng/µl), 11.7 µl molecular biology grade water, 4 µl buffer (25 mM), 1.4 µl $MgCl_2$ (5X), 0.2 µl dNTPs mix (2 mM), 0.25 µl forward primer (10 µM), 0.25 µl reverse primer (10 µM), and 0.2 µl GoTaq G2 DNA Polymerase (5 U/µl) (Promega, Fitchburg, WI, USA) in a total volume of 20 µl. The PCRs were performed on a Euroclone thermalcycler under the following conditions: (1) preheated lid at 105°C for 5 min; (2) 30 cycles: 94°C for 30 s, 48 ± 3°C for 45 s, and 72 °C for 45 s; (3) final extension at 72 °C for 5 min. The PCR products were visualised by 1.5% agarose gel electrophoresis. A DNA template to negative and positive control reaction was included in all experiments to test for contamination. PCR products were not obtained for some species due to the poor quality of the extracted DNA.

PCR products were purified with ExoSap enzymes (ExoSAP-IT® - USB corporation) and sequenced at the BMR Genomics company (Padua, Italy). The quality of the chromatograms was assessed with the Chromas Lite program (http://technelysium.com.au/wp/chromas/). The final consensus sequence, spanning about 650 base pairs, was assembled using the DNASTAR software (Lasergene® Madison, WI).

The data matrices

One hundred and sixty-five morphological traits were studied in adult moths, divided as follows: head (21), thorax (6), abdomen (9), male and female forewing (56), male and female hindwing (14), male genitalia (59) (Appendix 1). No female genitalia data were included in the present analysis. The hard decision to skip coding characters from the female genitalia was based on the circumstance that a quick screening over representatives of the principal assemblages of *Thaumetopoea* (viz. *T. processionea* and "*Traumatocampa*" *pityocampa*) essentially revealed differences in terms of relative sizes of soft structures or in the degree of sclerotisation of pieces, and did not allow to fix definite landmarks or boundaries. Consultation with two notodontid specialists (A. Schintlmeister and T.J. Witt, pers. comm.) confirmed this view and pointed out that traits identified by Agenjo (1941) were unreliable and largely a matter of artefact. It is also worth noting that Miller (1991), in his seminal phylogenetic reconstruction of the Notodontidae, was able to code only 11 female genital characters out of the whole family, whereas 23 were distinguished in the male ones. Further research, however, is being planned to check whether or not the female genitalia may increase resolution of the Thaumetopoeinae phylogeny, at least at the level of the most basal branches of the subfamily, which at this stage were not the major focus of our study.

The traits included in the morphological matrix (hereafter named Thau.morph) were mostly those used to identify genera and species of Thaumetopoeinae (Agenjo, 1941; Miller, 1991). The traits were coded following the guidelines provided by Strong and Lipscomb (1999) and Sereno (2007). Thus, 'traits' and the 'variable of the traits' were coded as different characters, in order to split neomorphic from transformational characters, providing a hierarchical order, and maintain the highest phylogenetic signal (Lee and Bryant, 1999).

Morphological wings terminology mainly follows Heath and Emmet (1979) and Heppner (2008). The genitalia terminology follows Berio (1985) and Steinmann and Zombori (1999). For other morphological terminology refer to Fig.1.

The morphological matrix was built in Mesquite v3.04 (Maddison and Maddison, 2015) and exported as nexus files for subsequent phylogenetic analyses.

The newly determined sequences were combined with previously available mitochondrial molecular data (Simonato *et al.*, 2013) in order to create a single multigene data set. The alignment of the *cox1*

sequences was performed using the MAFFT program implemented in the pipeline TranslatorX (Abascal *et al.*, 2010). This alignment was concatenated with the mitochondrial sequences of Simonato *et al.* (2013) to produce a single multiple alignment (hereafter named Thau.DNA) spanning 6348 positions where unavailable genes were treated as missing data. The list of genes included in Thau.DNA is provided in Table S2.

A total evidence data matrix (hereafter named Thau.tot.comp) was produced by combining morphological and DNA data. The Thau.tot.comp set includes all 37 taxa analysed in this paper. In Thau.tot.comp combined morphological and molecular data are available for 18 species, only morphological traits are present for 18 taxa, and for *T. pityocampa* ENA only molecular data exist (Table S2). Unavailable characters were coded as missing data. Finally, a second total evidence data matrix (hereafter named Thau.tot.red) was created containing only the 18 taxa for which both DNA and morphological data were available.

Phylogenetic analysis

Parsimony analyses (MP) (Fig. S7) were performed with TNT software (v1.5-beta) (Goloboff *et al.*, 2008). The Thau.morph, Thau.DNA, Thau.tot.comp and Thau.tot.red sets were analysed according to the strategy described below. The 'traditional search' algorithm was activated with the following settings: General RAM of 1.0 Gbytes, memory set to hold 1,000,000 trees, setting 1,000 replicates with tree bisection-reconnection (Goloboff, 1999) branch swapping and saving 1,000 trees per replicate. Zero-length branches were collapsed. To evaluate clade support, Bremer support values (Bremer, 1994) were calculated in TNT from 10,000 trees up to 25 steps longer than the most parsimonious trees obtained from 'traditional search' using 'trees from RAM'. A Bootstrap resampling (bt) (Felsenstein, 1985) under equal weights was carried out with 'traditional search' producing 10,000 replicates each of 1,000 random taxa addition; sub-replicates applying TBR, branch swapping and saving 1,000 trees per replication; resampling percentiles were calculated as frequency differences. Consistency index and retention index were calculated in Mesquite v3.04 (Maddison and Maddison, 2015).

The trait changes were mapped in WinClada software (Nixon, 2002) using the following settings: (1) optimization with unambiguous changes only.

Bayesian analysis (BI) were performed using MrBayes (x64, 3.2.5) (Ronquist *et al.*, 2012). The Thau.morph dataset was analysed under Mk1 model (Lewis, 2001) with the following settings: rates= gamma, and coding= variable. The GTR+I+G model was applied to Thau.DNA, and a combination of Mk1 model and GTR+I+G model was used for Thau.tot.comp and Thau.tot.red. Four Markov Chain Monte Carlo (MCMC) chains were run for 10 million generations and sampled every 1,000th

generation, with the first 25% of trees discarded as burn-in. Stationarity was considered to be reached when the average standard deviation of split frequencies was less than 0.005.

Maximum likelihood analyses were done using the Iqtree v1.3.13 software (Nguyen *et al.*, 2015). Analyses on Thau.morph dataset were performed with Ordered and MK models (both +FQ+ASC+G4). The GTR+I+G model was applied to Thau.DNA. Finally, combinations of Ordered and MK models plus GTR+I+G were applied to Thau.tot.comp and Thau.tot.red sets.

In every analysis, 50 independent tree searches were performed to minimize the possibility to be entrapped in a local minimum. Ultrafast Bootstrap support (Minh *et al.*, 2013) was calculated for every topology generated with Iqtree v1.3.13 program. In all analyses 10,000 replicates were performed.

The following abbreviations are used in the Results section to describe statistical support to tree topologies: bt, standard bootstrap support, UFB= UltraFastBootstrap support, pp= posterior probability, and Bvs= Bremer value support.

To evaluate alternative phylogenetic hypotheses, alternative topology tests were performed according to the Weighted Shimodaira and Hasegawa test (WSH) (Shimodaira and Hasegawa, 1999) and the Expected Likelihood Weights (ELW) test (Strimmer and Rambaut, 2002). The tests were done with the Iqtree v1.3.13 software (Nguyen *et al.*, 2015).

Tracking the evolution of a selection of characters on the reference tree

The MP tree obtained from Thau.tot.comp set was used as reference topology to track the evolution of six selected characters of genus *Thaumetopoea*.

Firstly, the evolutionary history of five larval morphological, ecological and biological traits (see the Traits description appendix for the coding strategy) was investigated. These traits are: (1) presence of urticating setae on larva; (2) pupation site; (3) larval seasonal feeding activity; (4) host plant group; (5) host plant family (Appendix 1). The tracking of the transformation pathways was performed with Mesquite 3.04 (Maddison and Maddison, 2015). The character history was traced by the 'Ancestral reconstruction' option based both on maximum parsimony (unordered characters) and the maximum likelihood method of Lewis (Lewis, 2001).

The sixth trait investigated was the reconstruction of the biogeographical patterns of the *Thaumetopoea* species. This analysis was performed with the Statistical Dispersal-Vicariance (S-DIVA) (Yu *et al.*, 2010) (see Results, Fig. 6).

Results

Morphology-based phylogenetic analysis

The MP tree (Length = 534; CI = 0.537; RI = 0.797) obtained from Thau.morph set is provided in Fig. 4 (see Fig. S4 for the outgroups). Nodes are supported by a variable number of apomorphies and homoplasious changes (Fig. 4; Table 3). The genus *Thaumetopoea* is monophyletic with strong statistical corroboration. The BI and ML topologies are mostly congruent with the MP tree with two noteworthy exceptions listed below.

The arrangement of the species belonging to clade O is different in the BI and ML (ORDERED+FQ+ASC+G4 evolutionary model) trees (Fig. 4; Fig. S5 and S7). In this alternative topology *T. bonjeani* and *T. pinivora* are sister species (clade W), a relationship that receives a very robust statistical support. Moreover, *T. libanotica* is sister taxon of clade W with high, even if not conclusive ($p < 0.95$), statistical corroboration (clade V; BI-pp = 0.93) (Fig. S5). *Thaumetopoea herculeana* is sister taxon of the J clade in the ML analysis based on the MK+FQ+ASC+G4 evolutionary model, but this relationship does not receive statistical corroboration (clade X) (Fig. 4; Fig. S6).

Finally, in all analyses based on morphology the current concept of *T. processionea* turned out as non-monophyletic, as the taxon *T. processionea pseudosolitaria* was recovered as sister to *T. solitaria*, although with no apomorphies supporting this relationship (Fig. 4; Figs. S5-S7).

The test performed on alternative phylogenetic hypotheses provided the results described below. The BI and ML trees topologies obtained from Thau.morph dataset are not rejected (ELW and WSH $p > 0.14$). Their parsimony scores (BI, Length = 543, CI = 0.528, RI = 0.790) (ML, Length = 541, CI = 0.530, RI = 0.791) are comparable with those of MP tree. The cladogram implying the monophyly of *T. processionea* species is not rejected (ELW and WSH > 0.12). This alternative tree exhibits parsimony scores (Length = 535, CI = 0.536, RI = 0.796) almost identical to those of the MP tree (see above).

Conversely, the placement of *T. herculeana* as sister taxon of clade B of Fig. 4 is rejected by ELW test ($p = 0.0306$) but not by WSH ($p = 0.3205$). The parsimony scores for this topology are worse than previous ones (Length = 544, CI = 0.528, RI = 0.789). This latter phylogenetic arrangement was tested because the placement of *T. herculeana* within *Thaumetopoea* is controversial in different analyses (see below).

DNA-based phylogenetic analysis

Availability of molecular data was restricted to 11 species of *Thaumetopoea* (Table S2), due to the failure to obtain DNA for some taxa. The MP cladogram (Tree Length = 4491, CI = 0.492, RI = 0.492) obtained from Thau.DNA set (Fig. S8) shows phylogenetic relationships among *Thaumetopoea* species, which are almost completely mirrored by the BI/ML tree (Fig. S9). This latter topology identifies a statistically supported (1.00/95%) sister taxon relationship between *T. apologetica* and *T. dhofarensis* not observed in the MP tree. The general arrangement of *Thaumetopoea* species relationships in DNA trees broadly agrees with topologies obtained from Thau.morph set (for complete details see Figs. S8-S9). However, the position of *T. herculeana* differs from that obtained from Thau.morph dataset (see node L of Fig 4). In all the DNA-based trees *T. herculeana* belongs to a very well supported clade, which contains also *T. solitaria* and *T. processionea.* In particular, *T. herculeana* reveals a sister species relationship with *T. processionea* with good statistical corroboration (Figs. S8-S9).

The BI/ML tree is not rejected by alternative topology tests (WSH, p = 0.9868; ELW, p = 0.8529). The parsimony scores are Length = 4490, CI = 0.622 and RI = 0.492. The alternative topology implying the placement of *T. herculeana* as recovered from Thau.morph set (Fig. 4) is rejected by both ELW and WSH tests ($p < 0.0001$). MP scores for this tree (Tree Length = 4634, CI = 0.603, RI = 0.449) are also markedly worse than the MP tree for what concerns the tree length.

Total evidence phylogenetic analysis

The MP tree obtained from Thau.tot.comp is shown in Fig. 5 (Tree Length = 5046, CI = 0.611, RI = 0.569) (see Fig. S10 for outgroups). The ML and BI trees obtained from the same data set are identical (Fig. S11). Their topology differs from MP cladogram (Figs.5; Fig.S11) by the placement of *T. jordana* and *T. dhofarensis*. These species are sister in the MP cladogram without bt and Bremer support (clade J). Conversely, the alternative BI/ML arrangement receives ML-UFB support (clade U) (Figs. 5 and S11). The vast majority of MP tree nodes receives strong statistical support and are supported by apomorphies and/or homoplasious changes (Fig. 5). Therefore, both the genus *Thaumetopoea* and most of its subclades receive statistical corroboration, but the placement of *T. herculeana* is different from that observed in Fig. 4. Here this moth is sister to *T. processionea* and this relationship receives very robust statistical support. Furthermore, both subspecies of the latter cluster together to form a monophyletic group with strong statistical corroboration.

The analyses performed on Thau.tot.red, the total evidence data set without missing data, provide topologies fully consistent with results obtained from Thau.tot.comp and are not detailed here any further (see Figs. S12-S13).

The BI/ML tree obtained from Thau.tot.comp is not rejected by the topology statistical tests (ELW, p= 0.406; and WSH, p = 0.701). The cladogram implying the placement of *T. herculeana* as sister taxon of clade O of Fig. 5, which mirrors the placement of this species in the Thau.morph analysis (clade L, Fig. 4), is fully rejected by both ELW and WSH tests ($p < 0.0001$).

The evolution of life history traits of Thaumetopoea

The evolutionary pathways of six life history traits were tracked along the MP tree obtained from Thau.tot.comp (Fig. 5). The results of this analysis are summarized in Fig. 6 (see Figs. S14-S15 for details). The inspection of Fig. 6 allows to outline some features that could have characterized the common ancestor of current *Thaumetopoea* moths (clade A). According to Fig. 6, urticating setae already characterized the larval stage of the ancestor of *Thaumetopoea*, which was active in summer, and used angiosperms as host plants. However, it is unclear which group of angiosperms was the host. Our reconstruction also suggests that the shift from angiosperms to gymnosperms as host plants occurred only once (clade K). Pupation occurred preferably in soil. The ancestral area of the genus was likely the Palaearctic Region and successive colonization of the Afrotropics occurred once/twice in *T. dhofarensis* and *T. apologetica.* Finally, the shift of the larval feeding from summer to winter seems to have occurred in the common ancestor of clade L, and independently in *T. jordana*. The reconstruction for the ancestral state of larval feeding activity at nodes G and K is ambiguously resolved by both MP and ML methods. Thus it could be argued that the shift from summer to winter occurred at node G (slightly preferred by MK1 ML evolutionary model) and successively a secondary reversion to summer activity happened in most of the species of clade O. By considering the global distribution of this character among *Thaumetopoea* species and the effect due to missing data on the reconstruction, this latter interpretation is not very probable in our view.

Discussion

Taxonomic inconsistencies

The monophyly of *Thaumetopoea s. lat.* (clade A, Figs 4-5) is strongly supported in all our analyses, thus corroborating previous studies based on a limited number of taxa (cf. Simonato *et al.*, 2013 and references therein). The phylogenetic relationships within the genus are well resolved and generally consistent with the available classification schemes (Agenjo, 1941; Kiriakoff, 1970), albeit with some important exceptions. All our analyses, based on the combined dataset, identify a clade B (Fig. 5) including the species traditionally assigned to *Thaumetopoea s. str*. (i.e., *T. processionea* and *T. solitaria*), with the addition of *T. herculeana*. Thus, *Thaumetopoea s. str.* is retrieved as paraphyletic. The remaining species of *Thaumetopoea* (clade G), which share a crested frons and a spine on foretibia, form a group corresponding to the (sub)genus *Traumatocampa* (Agenjo, 1941) (Tab. 2.). *Thaumetopoea herculeana* is quite mobile across our phylogenetic reconstructions, depending upon the dataset used. This mirrors its variable taxonomic position, as it has been assigned either to *Traumatocampa* due to the possession of a crested frons and the spine on foretibia (Agenjo, 1941; de Freina and Witt, 1985, 1987) (see also Tab. 2.), or to its own, monotypic genus *Helianthocampa* based upon larval host plant and adult morphology (de Freina and Witt, 1985). However, removing *T. herculeana* from *Traumatocampa* would make this no longer monophyletic on morphological grounds (see above Fig. 4).

Circumscribing Thaumetopoea*,* Traumatocampa*, and* Helianthocampa*: where should the tree be cut?*

The simultaneous presence of frontal crest and foretibia spine supports the monophyly of *Traumatocampa sensu* Agenjo (i.e., clade E of Fig. 4) in our morphology-based analysis (Supplementary material Figs. S4-S7). Frontal crest and foretibia spine have a clear adaptive significance as these species spend the pupal stage deep into the soil (usually between 5 and 20 cm) (Battisti *et al.*, 2015), and newly emerged adults use such anatomical structures to break through to the surface (Démolin, 1969). In contrast, processionary moths that do not have frontal crest and foretibia spine pupate either inside the tent (*T. processionea*) or in the litter (*T. solitaria*), but never deep into the soil.

In contrast to morphology, both molecular and combined datasets recover *T. herculeana* as nested within *Thaumetopoea s. str.* (clade B, Fig. 5), suggesting that the phylogenetic reconstruction is influenced by a strong molecular signal. It must be noted here that for several taxa the DNA sequences are represented solely by the barcoding portion of *cox1* gene (i.e. about 650 base pairs).

In the present paper we analysed only mitochondrial genes. Thus it could be argued that the DNA-based placement of *T. herculeana* is biased by the origin of the selected markers and does not

represent a balanced molecular view. However, the same placement was obtained by Simonato *et al.* (2013), who worked on a DNA multi-genes alignment encompassing both mitochondrial and nuclear markers. Particularly important for our discussion, the positioning of *T. herculeana* (Fig. 5) was recovered also from a data set containing only the nuclear genes: *wingless*, *elongation-factor alpha*, and *photolyase* (Simonato *et al.*, 2013). This latter result strongly contradicts the hypothesis of the incorrect placement *T. herculeana* due to the mitochondrial origin of molecular markers. Thus, despite the limited gene sampling, compared to the 165 morphological traits, it appears that molecules alone, or in combination with morphology, support a close relationship between *T. herculeana* and *T. processionea*.

Considering the host plant associations, it seems likely that larvae of the ancestor of *Thaumetopoea* fed on angiosperms and that only later a drastic shift to gymnosperms occurred, as observed in clade K (Fig. 6). The reconstruction of the pupation strategy (Fig. 6) supports the view that the soil was probably the pupation site of the *Thaumetopoea* ancestor, even if the proportional likelihood is not conclusive (Fig. 6). It may thence be supposed that soil pupation made the processionary moths better suited to highly variable climatic conditions, such as for instance those occurring in the Mediterranean. Stiff frontal protuberances enabling emerging adults to dig themselves out of hardened soils, either associated with foretibia spines or not, are a common adaptation in several members of the allied family Noctuidae, particularly if living in arid environments, e.g. *Cardepia*, *Craterestra*, *Conicofrontia*, *Grotella*, *Aedophron* (e.g., Janse, 1939; Hogue, 1963; Berio, 1985; Matthews, 1991; Fibiger *et al.*, 2009), but none appears to be configured as those of *Thaumetopoea s. lat.* Thus, it is conceivable that early ancestors of *T. processionea* and *T. solitaria* independently lost both frontal crest and foretibia spine as a consequence of a change in their pupation strategy from soil to litter (*T. solitaria*) and to tent (*T. processionea*). To further substantiate this scenario, it is interesting to note that both *T. processionea* and *T. solitaria* show a pronounced median protuberance on the frons (Fig. 1), whereas in *Gazalina* and other genera of Thaumetopoeinae (e.g., *Anaphe*, *Epicoma*, *Hypsoides*, and *Ochrogaster*) the frons is flat. The same applies for the foretibia, as both *T. processionea* and *T. solitaria* show a pronounced distal edge (Fig. 1) recalling a "regressed" foretibia spine, which is pretty different from the flat distal edge shared by the remaining Thaumetopoeinae. This is suggestive of the absence of the toothed frontal crest in the two species as being non homologous to that of other genera examined and representing a secondary loss. The toothed crest would therefore be confirmed as a good autapomorphy defining the whole *Thaumetopoea*, not just the *Traumatocampa* branch as supposed by the early authors and corroborated by the morphological dataset alone. Maintained in *T. herculeana*, topology of the total evidence cladogram (Fig. 5) is in accordance with an independent regression of the crest in both *T. processionea* and *T. solitaria*. As a matter of fact, the crest of processionary moths is such a complex structure that following release of

456 the selection on the character it is more parsimonious to admit its loss as a result of some disruption
457 in the underlying developmental gene machinery (e.g., Griffiths *et al.*, 1999; Hottes *et al.*, 2013;
458 Stower, 2013) than assuming its independent evolution in *T. herculeana* and typical *Traumatocampa*.
459
460 *The conifer feeding taxa: summer vs. winter*
461 All our cladograms agree in reconstructing conifer feeding processionary moths (clade K, Fig. 6) as
462 monophyletic. Monophyly of this clade has been strongly supported also by previous molecular
463 studies (Simonato *et al.*, 2013). Conifer feeding processionary moths split into two subclades (L and
464 O), each being characterized by a number of apomorphic character states (Fig. 5). Species included
465 in clades L and O are also characterized by having different developmental strategies (Démolin, 1989)
466 and diverging sexual pheromone composition (Frérot and Démolin, 1993). Clade L includes taxa
467 whose larvae feed across winter ('winter species') while clade O those feeding across spring and
468 summer ('summer species'). Clades L and O have been recovered as monophyletic in all our analyses,
469 gaining strong statistical support. Clades differ in relation to host plants, as winter taxa usually feed
470 on *Pinus* while summer taxa on *Cedrus*, although notable exceptions exist. In fact, *T. pityocampa*
471 (Clade L) can develop on *Cedrus* as well as on other conifers (Battisti *et al.*, 2015) whereas *T. pinivora*
472 (Clade O) feeds solely on *Pinus*, probably as a result of a host shift from a *Cedrus*-associated ancestor
473 (Cassel-Lundhagen *et al.*, 2013).
474 The study of the evolution of seasonal feeding of *Thaumetopoea* larvae suggests that a change from
475 summer to winter would have characterized the developmental strategy of the ancestor of clade K,
476 thus triggering allochronic speciation events (e.g., Santos *et al.*, 2007). A winter feeding larval stage
477 has several major advantages in temperate regions, such as a comparatively enemy- and competition-
478 free space, although it can be constrained by winter cold (Battisti *et al.*, 2015). This adaptation, that
479 involves several biochemical and physiological modifications, is a key to understand the response of
480 *T. pityocampa* to former and current climate changes (Battisti *et al.*, 2005), which has likely
481 contributed to the genetic population structure of this taxon (Kerdelhué *et al.*, 2009). Interestingly,
482 under mild oceanic climatic conditions a winter population of *T. pityocampa* in coastal Portugal has
483 recently shifted back to summer feeding, indicating that the trait has great adaptive value (Santos *et*
484 *al.*, 2007, 2011).
485
486 *Evolutionary history and conclusions*
487 The study of the evolution of life history traits allows us to sketch an evolutionary scenario for
488 *Thaumetopoea*. The Western Palaearctic Region, and the East Mediterranean-Middle East sub-region
489 in particular, was likely the area from where this group started diversifying from an ancestor
490 developing on angiosperms. It is likely that the host shift from angiosperms to gymnosperms that

491 characterized the ancestor of clade K (Figs. 5-6) occurred during the Tertiary and was further affected
492 during the Quaternary glaciation events. Previous molecular work done on a subset of taxa indicated
493 that the Messinian period is when major genetic diversification of species and clades happened
494 (Salvato *et al.*, 2002; Simonato *et al.*, 2007; Kerdelhué *et al.*, 2009). Pines and cedar trees are
495 documented to occur throughout that period in the West Palaearctic (Richardson, 1998; Qiao *et al.*,
496 2007), mainly in the mountains. This could have promoted the isolation of the insect populations and
497 the speciation process, as it is evident for the cedar-feeding species in the summer clade (Basso *et al.*,
498 2016). As the host plant shift appears to be unique, it is not possible to reconstruct with precision its
499 timing and location, although pines are the most likely candidates because of their large predominance
500 in fossil remains of conifers (Richardson, 1998). The host shift drove speciation events on these
501 plants, resulting now in the so-called 'coniferous *Thaumetopoea*' (Simonato *et al.*, 2013) (clade K)
502 (Figs. 5-6), which comprise ten taxa spread mostly in the Mediterranean subregion. Conversely, the
503 ancestors of clades B and H continued to rely on a very diverse number of angiosperms as host plants.
504 A common trait shared by the latter groups is the colonization of dry areas of Middle East and
505 Afrotropical region, which could have been facilitated by the plasticity in the life history. This is
506 remarkably shown by *T. jordana*, which shifted to winter feeding, possibly driven by the necessity to
507 escape from harsh conditions provided by Jordan valley in summer (Trougth, 1954). The spread in
508 the Afrotropical region would ultimately have been achieved by clade H radiating from the Middle
509 East. It would be interesting to compare the special *Thaumetopoea* adaptation to climate and host
510 plants with those of several other genera of Thaumetopoeinae occurring in Africa and Australasia in
511 a number of different habitats, to come up with a comprehensive phylogenetic analysis of the whole
512 group that could help to better understand the evolution of *Thaumetopoea*.
513
514 **Acknowledgments**
515
516 The authors warmly acknowledge the help of following people who provided useful suggestions and
517 specimens for the analyses: M. Doganlar, P. Paolucci, A. Schintlmeister, M. Simonato, T. J. Witt.
518 Thanks also to the curators of the Bavarian State Collection of Zoology of Munich Germany, Museo
519 Civico di Zoologia of Rome Italy, Museo di Zoologia, 'Sapienza' University of Rome (Rome Italy),
520 Royal Belgian Institute of Natural Sciences of Brussels Belgium, Royal Museum of Central Africa
521 of Tervuren Belgium, Witt Museum (Munich Germany) for the opportunity to study their collections,
522 and using their facilities, and finally to M. Ströhle and V. Zolotuhin for the DNA sequences of
523 'African' *Thaumetopoea*. Finally, we thank two anonymous reviewers for providing helpful
524 comments to an early version of this paper. This work was supported by a grant to Enrico Negrisolo
525 from Padua University (ex 60% 2012) and by a PhD grant from Padua University to Andrea Basso.

526 **References**

527

528 Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: multiple alignment of nucleotide
529 sequences guided by amino acid translations. *Nucleic Acids Res*, **38**, W7-13.

530 Agenjo, R., 1941. Monografia de la familia Thaumetopoeidae (Lep.). *EOS*, **17**, 69-130.

531 Basso, A., Simonato, M., Cerretti, P., Paolucci, P., Battisti, A., 2016. A review of the "summer"
532 *Thaumetopoea* spp. (Lepidoptera: Notodontidae, Thaumetopoeinae) associated with *Cedrus*
533 and *Pinus*. *Turk J Forestry*, **17**, 31-39.

534 Battisti, A., Avcı, M., Avtzis, D.N., Ben Jamaa, M.L., Berardi, L., Berretima, W.A., Branco, M.,
535 Chakali, G., El Alaoui El Fels, M.A., Frérot, B., Hodar, J.H., Ionescu-Malancus, I., Ipekdal,
536 K., Larsson, S., Manole, T., Mendel, Z., Meurisse, N., Mirchev, P., Nemer, N., Paiva, M.-R.,
537 Pino, J., Protasov, A., Rahim, N., Rousselet, J., Santos, H., Sauvard, D., Schopf, A., Simonato,
538 M., Yart, A., Zamoum, M., 2015. Natural history of the processionary moths (*Thaumetopoea*
539 spp.): new insights in relation to climate change. In: *Processionary Moths and Climate*
540 *Change: An Update*, A. Roques, (Ed.). Springer, Dordrecht, pp: 15-80.

541 Battisti, A., Holm, G., Fagrell, B., Larsson, S., 2011. Urticating hairs in arthropods: their nature and
542 medical significance. *Annu Rev Entomol*, **56**, 203-220.

543 Battisti, A., Larsson, S., Roques, A., 2017. Processionary moths and associated urtication risk: global-
544 change driven effects. *Annu Rev Entomol*, **62**, 323-342.

545 Battisti, A., Stastny, M., Netherer, S., Robinet, C., Schopf, A., Roques, A., Larsson, S., 2005.
546 Expansion of geographic range in the pine processionary moth caused by increased winter
547 temperatures. *Ecol Appl*, **15**, 2084-2096.

548 Berio, E., 1985. *Lepidoptera: Noctuidae. Generalità Hadeninae Cucullinae*. Fauna d'Italia, **22**.
549 Calderini, Bologna.

550 Bremer, K., 1994. Branch support and tree stability. *Cladistics*, **10**, 295-304.

551 Cassel-Lundhagen, A., Ronnas, C., Battisti, A., Wallen, J., Larsson, S., 2013. Stepping-stone
552 expansion and habitat loss explain a peculiar genetic structure and distribution of a forest
553 insect. *Mol Ecol*, **22**, 3362-3375.

554 Démolin, G., 1969. Comportement des adultes de *Thaumetopoea pityocampa* Schiff.: dispersion
555 spatiale, importance ecologique. *Ann Sci For*, **26**, 81-102.

556 Démolin, G., 1989. La processionnaire du Cedre: *Thaumetopoea bonjeani* Powell. Rapport
557 scientifique et rapport iconographique, 1986-87. *Intensification de la Protection*
558 *Phytosanitaire des Forets*, United Nations Program for Development; DP-FO-ALG/83/013:
559 110.

560 Felsenstein, J., 1985. Confidence-limits on phylogenies - an approach using the bootstrap. *Evolution*,
561 **39**, 783-791.
562 Fibiger, M., Ronkay, L., Steiner, A., Zilli, A., 2009. *Pantheinae-Bryophilinae*. Noctuidae Europaeae,
563 **11**. Entomological Press, Sorø.
564 Floater, G.J., Zalucki, M.P., 2000. Habitat structure and egg distributions in the processionary
565 caterpillar Ochrogaster lunifer: lessons for conservation and pest management. *J Appl Ecol*,
566 **37**, 87-99.
567 Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of
568 mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar*
569 *Biol Biotechnol*, **3**, 294-299.
570 Freina, J.J., de, Witt, T.J., 1985. Taxonomische Veränderungen bei den Bombyces und Sphinges
571 Europas und Norwestafrikas. *Helianthocampa* gen. nov.; *Traumatocampa galaica* (Palanca
572 Soler et al. 1982) comb. nov. et syn. nov. (Lepidoptera, Thaumetopoeidae, Thaumetopoeinae).
573 *Nota Lepid*, **8**, 175-183.
574 Freina, J.J., de, Witt, T.J., 1987. *Die Bombyces und Sphinges der Westpalaaerktis (Insecta,*
575 *Lepidoptera).* Forschung und Wissenschaft Verlag GmbH, München.
576 Frérot, B., Démolin, G., 1993. Sex pheromone of the processionary moths and biosystematics
577 considerations within the genus *Thaumetopoea* (Thaumetopoeidae, Thaumetopoeinae). *Boll*
578 *Zool Agrar Bachic*, **25**, 33-40.
579 Goloboff, P.A., 1999. Analyzing large data sets in reasonable times: solutions for composite optima.
580 *Cladistics*, **15**, 415-428.
581 Goloboff, P.A., Farris, J.S., Nixon, K.C., 2008. TNT, a free program for phylogenetic analysis.
582 *Cladistics*, **24**, 774-786.
583 Griffiths, A.J., Gelbart, W.M., Miller, J.H., Lewontin, R.C., 1999. *Modern Genetic Analysis*. W. H.
584 Freeman, New York.
585 Hacker, H., 2016. Systematic and illustrated catalogue of the Macroheterocera and superfamilies
586 Cossoidea Leach, [1815], Zygaenoidea Latreille, 1809, Thyridoidea Herrich-Schäffer, 1846
587 and Hyblaeoidea Hampson, 1903 of the Arabian Peninsula, with a survey of their distribution
588 (Lepidoptera). *Esperiana*, **20**, 147-149.
589 Heath, J., Emmet, A.M., 1979. *The Moths and butterflies of Great Britain and Ireland*, **9**. Curwen
590 Books, Oxford.
591 Heppner, J.B., 2008. Butterflies and Moths (Lepidoptera). In: *Encyclopedia of Entomology*, Capinera
592 John L., (Ed.). Springer Netherlands, pp: 690-747.
593 Hogue, C.L., 1963. A definition and classification of the tribe Stiriini (Lepidoptera: Noctuidae). *Contr*
594 *Sci*, **64**, 1-129.

595 Hottes, A.K., Freddolino, P.L., Khare, A., Donnell, Z.N., Liu, J.C., Tavazoie, S., 2013. Bacterial
596 adaptation through loss of function. *PLoS Genet*, **9**, e1003617.
597 Hübner, J., 1816-[1826]. *Verzeichniss bekannter Schmettlinge* [sic]. published by the Author,
598 Augsburg, pp: 185.
599 Huis, A., van, 2003. Insects as Food in sub-Saharan Africa. *Ins Sci Appl*, **23**, 163-185.
600 Jacquet, J.S., Orazio, C., Jactel, H., 2012. Defoliation by processionary moth significantly reduces
601 tree growth: a quantitative review. *Ann For Sci*, **69**, 857-866.
602 Janse, A.J.T., 1939. *The Moths of South Africa 3: Cymatophoridae, Callidulidae and Noctuidae*
603 *(partim)*. E. P. & Commercial Printing, Durban.
604 Kerdelhué, C., Zane, L., Simonato, M., Salvato, P., Rousselet, J., Roques, A., Battisti, A., 2009.
605 Quaternary history and contemporary patterns in a currently expanding species. *BMC Evol*
606 *Biol*, **9**.
607 Kiriakoff, S.G., 1970. *Lepidoptera Familia Thaumetopoeidae.* Genera Insectorum. SPRL, Anvers
608 Belgium.
609 Knölke, S., Erlacher, S., Hausmann, A., Miller, M.A., Segerer, A.H., 2005. A procedure for combined
610 genitalia dissection and DNA extraction in Lepidoptera. *Insect Syst Evol*, **35**, 401-409.
611 Lee, D.C., Bryant, H.N., 1999. A reconsideration of the coding of inapplicable characters:
612 Assumptions and problems. *Cladistics*, **15**, 373-378.
613 Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological
614 character data. *Syst Biol*, **50**, 913-925.
615 Maddison, W.P., Maddison, D.R., 2015. Mesquite: a modular system for evolutionary analysis.
616 Version 3.04 (http://mesquiteproject.org).
617 Matthews, M., 1991. Classification of the Heliothinae. *Nat Resour Inst Bull*, **44**, i-vi, 1-195.
618 Miller, J.S., 1991. Cladistics and classification of the Notodontidae (Lepidoptera, Noctuoidea) based
619 on larval and adult morphology. *B Am Mus Nat Hist*, **204**: 1-230.
620 Minh, B.Q., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast Approximation for Phylogenetic
621 Bootstrap. *Mol Biol Evol*, **30**, 1188-1195.
622 Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective
623 stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*, **32**, 268-
624 274.
625 Nixon, K.C., 2002. *Winclada, Version 1.00.08*. Access: November 2015. http://www.cladistics.com.
626 Qiao, C.Y., Ran, J.H., Li, Y., Wang, X.Q., 2007. Phylogeny and biogeography of *Cedrus* (Pinaceae)
627 inferred from sequences of seven paternal chloroplast and maternal mitochondrial DNA
628 regions. *Ann Bot*, **100**, 573-580.

629 Rahman, W.U., Chaudhry, M.I., 1992. Observations on outbreak and biology of oak defoliator,
630 *Gazalina chrysolopha* Koll. *Pak J For*, **42**, 134-137.
631 Richardson, D.M., 1998. *Ecology and Biogeography of Pinus*. Cambridge University Press,
632 Cambridge, UK.
633 Robinson, G.S., 1976. The preparation of slides of Lepidoptera genitalia with special reference to the
634 Microlepidoptera. *Entomol Gaz*, **27**, 127-132.
635 Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu,
636 L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic
637 inference and model choice across a large model space. *Syst Biol*, **61**, 539-542.
638 Roques, A., Battisti, A., 2015. Introduction. In: *Processionary Moths and Climate Change: An*
639 *Update*, A. Roques, (Ed.). Springer-Quae, Dordrecht-Versailles, pp: 1-14.
640 Salvato, P., Battisti, A., Concato, S., Masutti, L., Patarnello, T., Zane, L., 2002. Genetic
641 differentiation in the winter pine processionary moth (*Thaumetopoea pityocampa - wilkinsoni*
642 complex), inferred by AFLP and mitochondrial DNA markers. *Mol Ecol,* **11**, 2435-2444.
643 Santos, H., Burban, C., Rousselet, J., Rossi, J.P., Branco, M., Kerdelhue, C., 2011. Incipient
644 allochronic speciation in the pine processionary moth (*Thaumetopoea pityocampa*,
645 Lepidoptera, Notodontidae). *J Evol Biol*, **24**, 146-158.
646 Santos, H., Rousselet, J., Magnoux, E., Paiva, M.R., Branco, M., Kerdelhue, C., 2007. Genetic
647 isolation through time: allochronic differentiation of a phenologically atypical population of
648 the pine processionary moth. *Proc Biol Sci*, **274**, 935-941.
649 Schabel, H.G., 2006. *Forest Entomology in East Africa: Forest Insects of Tanzania*. Springer,
650 Dordrecht.
651 Schintlmeister, A., 2013. *Notodontidae & Oenosandridae (Lepidoptera)*. World catalogue of insects,
652 **11**. Brill, Leiden.
653 Sereno, P.C., 2007. Logical basis for morphological characters in phylogenetics. *Cladistics*, **23**, 565-
654 587.
655 Shimodaira, H., Hasegawa, M., 1999. Multiple Comparisons of Log-Likelihoods with Applications
656 to Phylogenetic Inference. *Mol Biol Evol*, **16**, 1114.
657 Simonato, M., Battisti, A., Kerdelhue, C., Burban, C., Lopez-Vaamonde, C., Pivotto, I., Salvato, P.,
658 Negrisolo, E., 2013. Host and phenology shifts in the evolution of the social moth genus
659 *Thaumetopoea*. *PloS one*, **8**, e57192.
660 Simonato, M., Mendel, Z., Kerdelhue, C., Rousselet, J., Magnoux, E., Salvato, P., Roques, A.,
661 Battisti, A., Zane, L., 2007. Phylogeography of the pine processionary moth *Thaumetopoea*
662 *wilkinsoni* in the Near East. *Mol Ecol*, **16**, 2273-2283.
663 Steinmann, H., Zombori, L., 1999. *Dictionary of Insect Morphology*. De Gruyter, Berlin.

Stower, H., 2013. Molecular evolution: Adaptation by loss of function. *Nat Rev Genet*, **14**, 596-597.

Strimmer, K., Rambaut, A., 2002. Inferring confidence sets of possibly misspecified gene trees. *P Roy Soc B-Biol Sci*, **269**, 137-142.

Strong, E.E., Lipscomb, D., 1999. Character coding and inapplicable data. *Cladistics*, **15**, 363-371.

Trougth, T., 1954. The life history of *Thaumetopoea jordana* Staudinger. *Entomol Rec J Var*, **66**, 188-191.

Wagner, M.R., Cobbinah, J.R., 2013. *Forest entomology in West Tropical Africa: Forest insects of Ghana*. Springer Netherlands.

Wallengren, H.D.J., 1871. *Spinnarne*. Skandinaviens Heterocer-Fjarilar, **2**. Chr. Bülow, Lund.

Yu, Y., Harris, A.J., He, X., 2010. S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool for inferring biogeographic histories. *Mol Phylogenet Evol*, **56**, 848-850.

Zahiri, R., Kitching, I.J., Lafontaine, J.D., Mutanen, M., Kaila, L., Holloway, J.D., Wahlberg, N., 2011. A new molecular phylogeny offers hope for a stable family level classification of the Noctuoidea (Lepidoptera). *Zool Scr*, **40**, 158-173.

Zahiri, R., Lafontaine, D., Schmidt, C., Holloway, J.D., Kitching, I.J., Mutanen, M., Wahlberg, N., 2013. Relationships among the basal lineages of Noctuidae (Lepidoptera, Noctuoidea) based on eight gene regions. *Zool Scr*, **42**, 488-507.

682 **Legend to the Tables**

683

684 **Table. 1.** List of Thaumetopoeinae taxa considered in this work.

685

686 **Table. 2.** Association of *Thaumetopoea s. lat.* species-group taxa to genera or subgenera by various
687 authors.

688

689 **Table. 3.** Indication of apomorphic characters and homoplasious changes of total evidence tree
690 showed in Fig. 4.

691

692 **Table. 4.** Indication of apomorphic characters and homoplasious changes of total evidence tree
693 showed in Fig. 5.

694

695

Table. 1.

Taxon	Author	Distribution	Host plants
Thaumetopoea processionea processionea	(Linnaeus, 1758)	Europe, Middle East	*Quercus* spp.
Thaumetopoea processionea pseudosolitaria	Daniel, 1951	South - Eastern Europe, Middle East	*Quercus* spp.
Thaumetopoea solitaria iranica	Agenjo, 1941	Middle East	*Pistacia* spp.
Thaumetopoea solitaria solitaria	(Freyer, 1838)	South - Eastern Europe, Middle East	*Pistacia* spp.
Thaumetopoea herculeana herculeana	(Rambur, 1837)	South - Western Europe, Northern Africa, Middle East	*Cistus* spp., *Erodium* spp., *Helianthemum* spp.
Thaumetopoea herculeana judaea	Bang-Haas, 1910	Middle East	unknown
Thaumetopoea apologetica abyssinica	Strand, 1911	Eastern Africa	*Maerua* spp.
Thaumetopoea apologetica apologetica	Strand, 1909	Southern and Eastern Africa	*Maerua* spp.
Thaumetopoea dhofarensis	Wiltshire, 1980	Middle East	unknown
Thaumetopoea jordana	Staudinger, 1887	Middle East	*Rhus tripartita*
Thaumetopoea cheela	Moore, 1883	Southern Asia	*Pinus* spp., or *Cedrus* spp.
Thaumetopoea bonjeani	Powell, 1922	North - Western Africa	*Cedrus atlantica*
Thaumetopoea pinivora	(Treitschke, 1834)	Central, Northern and Eastern Europe	*Pinus* spp.
Thaumetopoea libanotica	Kiriakoff & Talhouk, 1975	Middle East	*Cedrus libani*
Thaumetopoea ispartaensis	Doganlar & Avci, 2001	Middle East	*Cedrus libani*
Thaumetopoea sedirica	(Doganlar, 2005)	Middle East	*Cedrus libani*
Thaumetopoea torosica	(Doganlar, 2005)	Middle East	*Pinus brutia*
Thaumetopoea pityocampa pityocampa	([Denis & Schiffermüller], 1775)	Central Europe, Mediterranean region	*Cedrus* spp., *Pinus* spp.
Thaumetopoea pityocampa orana	(Staudinger, 1901)	North - Western Africa	*Pinus* spp.
Thaumetopoea pityocampa ENA	—	North - Eastern Africa	*Cedrus atlantica, Pinus halepensis*
Thaumetopoea wilkinsoni	Tams, 1925	Middle East	*Cedrus libani, Pinus* spp.
Anaphe panda panda	(Boisduval, 1847)	Central - Southern Africa	Polyphagous
Anaphe panda infracta	Walsingham, 1885	Central - Southern Africa	Polyphagous
Anaphe venata	Butler, 1878	Central - Southern Africa	Polyphagous
Anaphe etiennei	Schouteden, 1912	Central - Southern Africa	Polyphagous
Epanaphe nigricincta	(Hulstaert, 1924)	Central - Southern Africa	Polyphagous
Epanaphe subsordida	(Holland, 1893)	Central - Southern Africa	Polyphagous
Epanaphe moloneyi	(Druce, 1887)	Central - Southern Africa	Polyphagous
Epanaphe carteri	(Walsingham, 1855)	Central - Southern Africa	Polyphagous
Gazalina apsara	(Moore, 1859)	Southern Asia	Betulaceae, Fagaceae
Gazalina chrysolopha	(Kollar, 1844)	Southern Asia	Betulaceae, Fagaceae
Gazalina transversa	Moore, 1879	Southern Asia	Betulaceae, Fagaceae
Hypsoides antsianakana	(Oberthür, 1922)	Eastern Africa	Gentianaceae
Hypsoides placidus	(Oberthür, 1923)	Eastern Africa	Gentianaceae
Ochrogaster lunifer	Herrich-Schäffer, 1855	Australia	*Acacia* spp., *Corymbia* spp., *Eucalyptus* spp.
Paradrallia punctigera	Hulstaert, 1924	Central Africa	Fabaceae
Paradrallia rhodesi	Bethune-Baker, 1908	Central Africa	Fabaceae

Table. 2.

Taxon	Agenjo, 1941	de Freina & Witt, 1987	Simonato, 2013	Schintlmeister, 2013
processionea processionea	Thaumetopoea	Thaumetopoea	Thaumetopoea	Thaumetopoea
processionea pseudosolitaria	Thaumetopoea	—	Thaumetopoea	Thaumetopoea
solitaria solitaria	Thaumetopoea	Thaumetopoea	Thaumetopoea	Thaumetopoea
solitaria iranica	Thaumetopoea	—	Thaumetopoea	Thaumetopoea
herculeana herculeana	Traumatocampa	Helianthocampa	Thaumetopoea	Helianthocampa
herculeana judaea	Traumatocampa	—	Thaumetopoea	Helianthocampa
apologetica abyssinica	—	—	Thaumetopoea	Thaumetopoea
apologetica apologetica	—	—	Thaumetopoea	Thaumetopoea
dhofarensis	—	—	Thaumetopoea	Thaumetopoea
jordana	—	—	Thaumetopoea	Traumatocampa
cheela	—	—	Thaumetopoea	Thaumetopoea
bonjeani	Traumatocampa	Traumatocampa	Thaumetopoea	Traumatocampa
pinivora	Traumatocampa	Traumatocampa	Thaumetopoea	Traumatocampa
libanotica	Traumatocampa	—	Thaumetopoea	Traumatocampa
ispartaensis	—	—	Thaumetopoea	Traumatocampa
sedirica	—	—	Thaumetopoea	Traumatocampa
torosica	—	—	Thaumetopoea	Traumatocampa
pityocampa pityocampa	Traumatocampa	Traumatocampa	Thaumetopoea	Traumatocampa
pityocampa orana	Traumatocampa	—	Thaumetopoea	Traumatocampa
pityocampa ENA	—	—	Thaumetopoea	—
wilkinsoni	Traumatocampa	—	Thaumetopoea	Traumatocampa

696
697

Table. 3.

Node	Apomorphies	Homoplasious changes
A	17:0; 30:0; 36:0; 91:0; 102:1; 111:1; 129:1	37:3; 38:1;74:0
B	29:3; 123:1	18:2; 24:2; 85:1
C	-	113:2; 115:2; 121:1; 133:1; 146:1; 150:1
D	28:3; 41:1; 86:1	37:2; 49:0; 64:0; 76:1; 79:1; 80:0; 82:0; 97:1; 163:1
E	2:0; 11:0; 26:0; 27:0; 112:0; 113:0	49:0; 64:0; 132:0
F	39:1; 42:1; 47:1	28:1; 31:1; 52:0; 60:1; 68:0; 81:3; 84:1; 104:1; 105:1; 135:1; 136:0; 157:2
G	-	132:2
H	8:2	74:0; 115:0
I	22:3; 101:1	106:0
J	29:1	24:0; 35:0; 89:1; 115:0; 133:1
K	-	82:0
L	50:1; 66:1; 112:1	33:1; 82:2; 100:1; 121:1
M	7:1; 8:0; 40:0; 48:0; 61:0; 94:1	12:0; 22:0; 29:0; 52:0; 68:0; 76:1; 96:1; 99:1; 103:1; 104:1; 115:2; 151:2
N	5:1; 28:2; 43:1; 49:2; 134:1	45:1; 140:1
O	138:1; 149:2	13:1; 103:1; 114:1; 150:1; 163:1
P	-	11:1; 12:2; 34:2; 88:1; 89:1
Q	-	10:0; 18:3; 37:3; 86:2; 103:1
R	71:1	9:0; 13:0; 44:0; 79:1; 81:1; 89:1; 147:3; 149:0; 151:0
S	-	18:1

698
699

Table. 4.

Node	Apomorphies	Homoplasious changes
A	15:1; 17:0; 23:1, 2; 30:0; 36:0; 64:0; 91:0; 129:1; 145:1	22:2,3; 38:1; 74:1; 107:0
B	29:3	14:0; 113:2; 115:2; 128:1
C	28:3	80:0; 82:0; 128:0; 133:1; 146:1; 150:1; 163:1
D	96:1	28:0; 99:1; 100:1
E	13:3	37:3; 49:1; 64:1; 81:1; 155:1
F	-	-
G	112:0; 113:0	2:0; 11:0; 26:0; 27:0; 78:2
H	4:0; 39:1; 42:1; 47:1; 151:0	18:3; 19:2; 52:0; 60:1; 68:0; 78:1; 81:3; 104:1; 135:1; 136:0; 137:3; 157:2
I	-	132:2
J	8:2	74:0; 115:0; 132:0
K	10:1	18:1; 23:0; 101:1; 131:0; 137:0
L	29:1	13:0; 24:0; 28:0; 35:0; 89:1; 97:1; 115:0; 161:1; 128:1; 133:1
M	10:2	34:2; 37:3; 82:0; 132:0
N	-	-
O	1:2; 15:2; 28:2; 43:1; 134:1; 145:0; 147:3	14:1; 33:1; 49:2; 50:1; 66:1; 100:1; 140:1
P	138:1	114:1; 128:4; 130:1; 163:1
Q	148:0	18:0; 101:0
R	-	10:0; 37:3; 86:2; 103:1
S	71:1	9:0; 13:0; 44:0; 79:1; 89:1
T	-	-

Legend to the Figures

Fig. 1. Traits of *Thaumetopoea* spp.. a) Dorsal habitus and veins of wings; b) head traits. Frontal protuberance or crest: c) *T. herculeana*, d) *T. apologetica*, e) *T. dhofarensis* / *T. jordana*, f) *T. pityocampa*, g) *T. wilkinsoni*, h) *T. bonjeani*, i) *T. cheela*, j) *T. libanotica*, k) *T. ispartaensis* / *T. sedirica*, l) *T. pinivora,* m) *T. processionea* / *T. solitaria*. Legs silhouette: n) foreleg of crested *Thaumetopoea,* o) foreleg of non-crested *Thaumetopoea*, p) midleg / hindleg of *Thaumetopoea*. Genitalia of *Thaumetopoea* spp.: q) front view, and r) lateral view.

Fig. 2. Taxa studied in this work. (a), *Thaumetopoea herculeana herculeana* - Algeria (BMHN(E)_1378597); (b), *Thaumetopoea processionea* - France (BMHN(E)_1378609); (c), *Thaumetopoea solitaria solitaria* - Cyprus (BMHN(E)_1378618); (d), *Thaumetopoea apologetica apologetica* - Kenya (BMHN(E)_1378606); (e), *Thaumetopoea dhofarensis* - Oman (WITT_TH73); (f), *Thaumetopoea jordana* - Palestine (ZSM_TH72); (g), *Thaumetopoea pityocampa pityocampa* - Greece (BMHN(E)_1378583); (h), *Thaumetopoea wilkinsoni* - Cyprus (BMHN(E)_1378572). The collection code associated to the specimen is provided in brackets (see Table S1). Scale bar = 1 cm.

Fig. 3. Taxa studied in this work. (a), *Thaumetopoea bonjeani* - Morocco (BMHN(E)_1378573); (b), *Thaumetopoea cheela* ST - India (BMHN(E)_1378635); (c), *Thaumetopoea ispartaensis* - Turkey (DAFNAE_TH30); (d), *Thaumetopoea libanotica* - Lebanon (BMNH(E)_1378640); (e), *Thaumetopoea pinivora* - France (BMHN(E)_1378577); (f), *Thaumetopoea sedirica* PT - Turkey (DAFNAE_TH29); (g), *Thaumetopoea torosica* PT - Turkey (DAFNAE_TH64). The collection code associated to the specimen is provided in brackets (see Table S1). Scale bar = 1 cm.

Fig. 4. The most parsimonious tree (Length = 534, CI = 0.537, RI = 0.797) obtained from Thau.morph dataset. Optimization: only unambiguous changes are mapped. Black circles, apomorphic characters; white circles, homoplasious changes. Partial topologies show the differences observed in BI and ML trees. MP-Bremer, Bremer support to the node; MP-bt, bootstrap support to the node; ML-ord-UFB, ultrafast bootstrap support to the node (ordered model); ML-mk-UFB, ultrafast bootstrap support to the node (unordered model); BI-pp, posterior probability support to the node.

Fig. 5. The most parsimonious tree (Length = 504, CI = 0.611, RI = 0.569) obtained from Thau.tot.comp dataset. Optimization: only unambiguous changes are mapped. Black circles, apomorphic characters; white circles, homoplasious changes. Partial topology shows the difference observed BI and ML trees. MP-Bremer, Bremer support to the node; MP-bt, bootstrap support to the node; ML-ord-UFB, ultrafast bootstrap support to the node (ordered model); ML-mk-UFB, ultrafast bootstrap support to the node (unordered model); BI-pp, posterior probability support to the node.

Fig. 6. The evolution of life history traits tracked of the MP tree obtained from Thau.tot.comp dataset. The transformational pathways are provided for all terminal nodes plus selected internal nodes (for full details see Figs. S14-S15). Maximum parsimony reconstructions are figured on the branch leading to the selected node. Maximum likelihood reconstructions are provide only for ambiguous results below the reference branch. Traits 1-5 were analysed with the Mesquite software. Trait 6 was mapped with S-DIVA software.

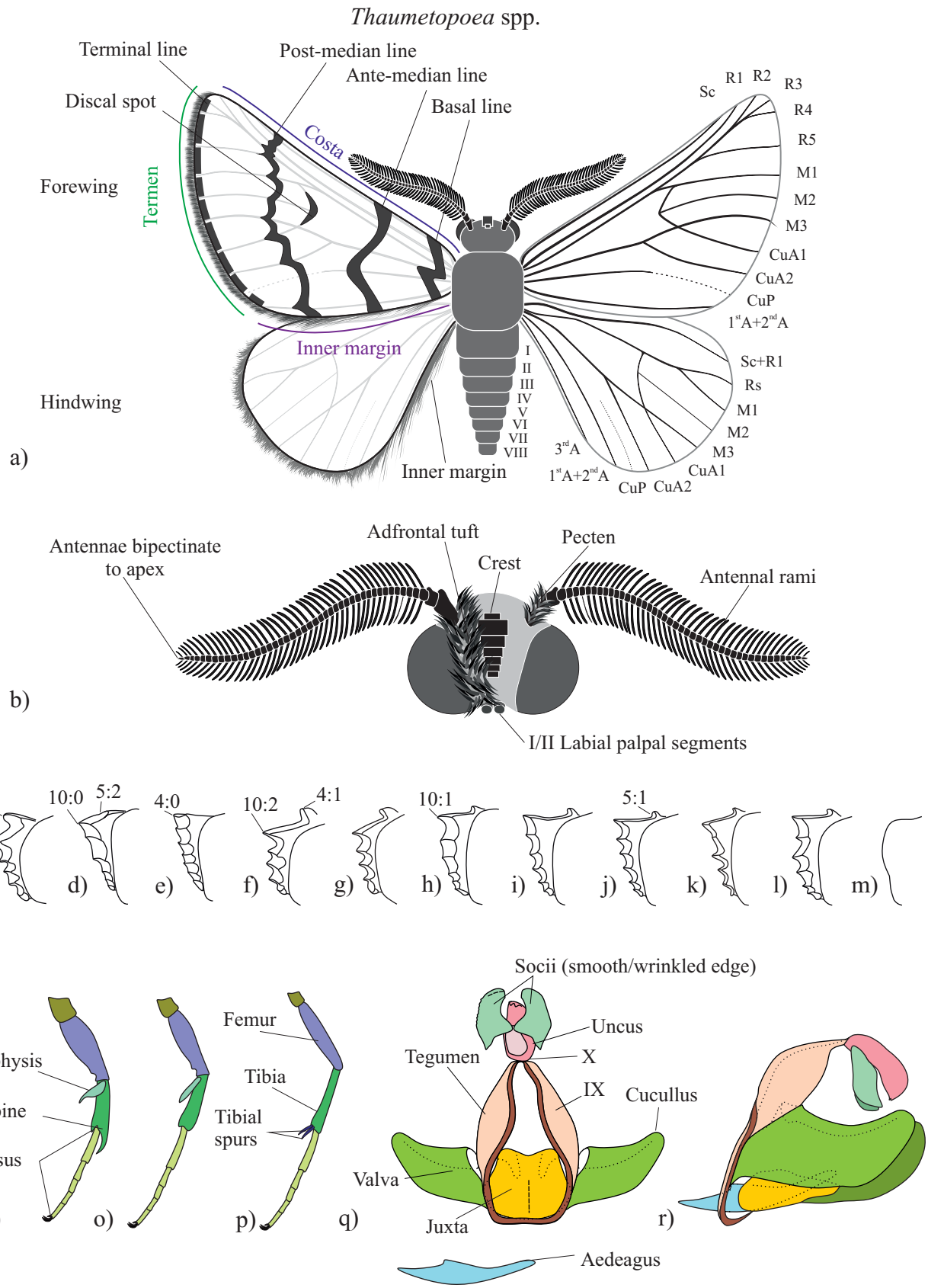

Thaumetopoea spp.
Terminal line
Post-median line
Ante-median line
Discal spot
Basal line
Costa
Forewing
Termen
Inner margin
Hindwing
a)
Sc
R1
R2
R3
R4
R5
M1
M2
M3
CuA1
CuA2
CuP
1st A+2nd A
Sc+R1
Rs
M1
M2
M3
CuA1
CuP
CuA2
3rd A
1st A+2nd A
I
II
III
IV
V
VI
VII
VIII
Inner margin
Antennae bipectinate to apex
Adfrontal tuft
Crest
Pecten
Antennal rami
I/II Labial palpal segments
b)
5:0
10:0
5:2
4:0
10:2
4:1
10:1
5:1
c)
d)
e)
f)
g)
h)
i)
j)
k)
l)
m)
Epiphysis
Spine
Tarsus
Femur
Tibia
Tibial spurs
n)
o)
p)
Socii (smooth/wrinkled edge)
Uncus
Tegumen
X
IX
Cucullus
Valva
Juxta
Aedeagus
q)
r)

Figure 1

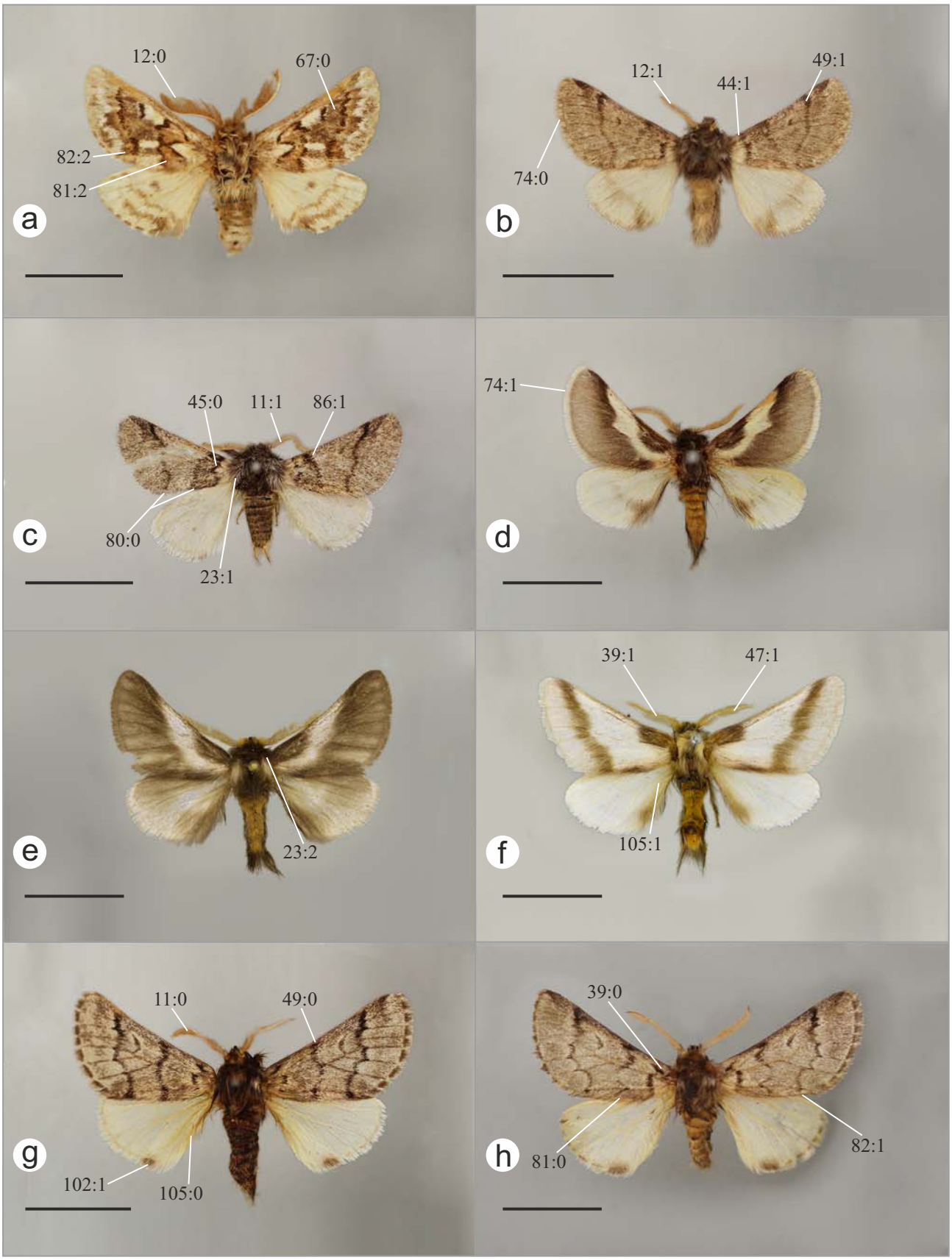

12:0
67:0
82:2
81:2
a
12:1
44:1
49:1
74:0
b
45:0
11:1
86:1
80:0
23:1
c
74:1
d
23:2
e
39:1
47:1
105:1
f
11:0
49:0
102:1
105:0
g
39:0
81:0
82:1
h

Figure 2

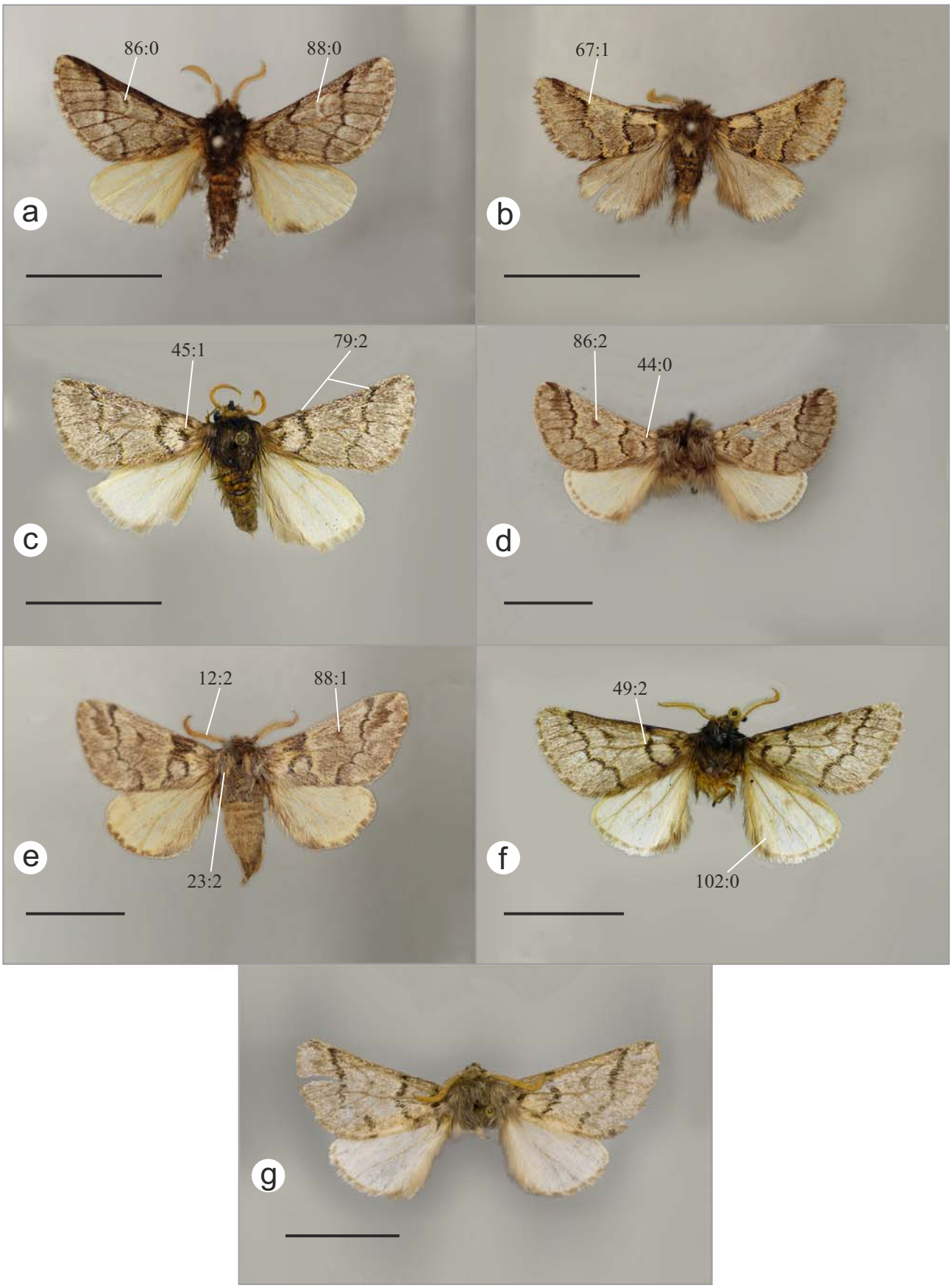
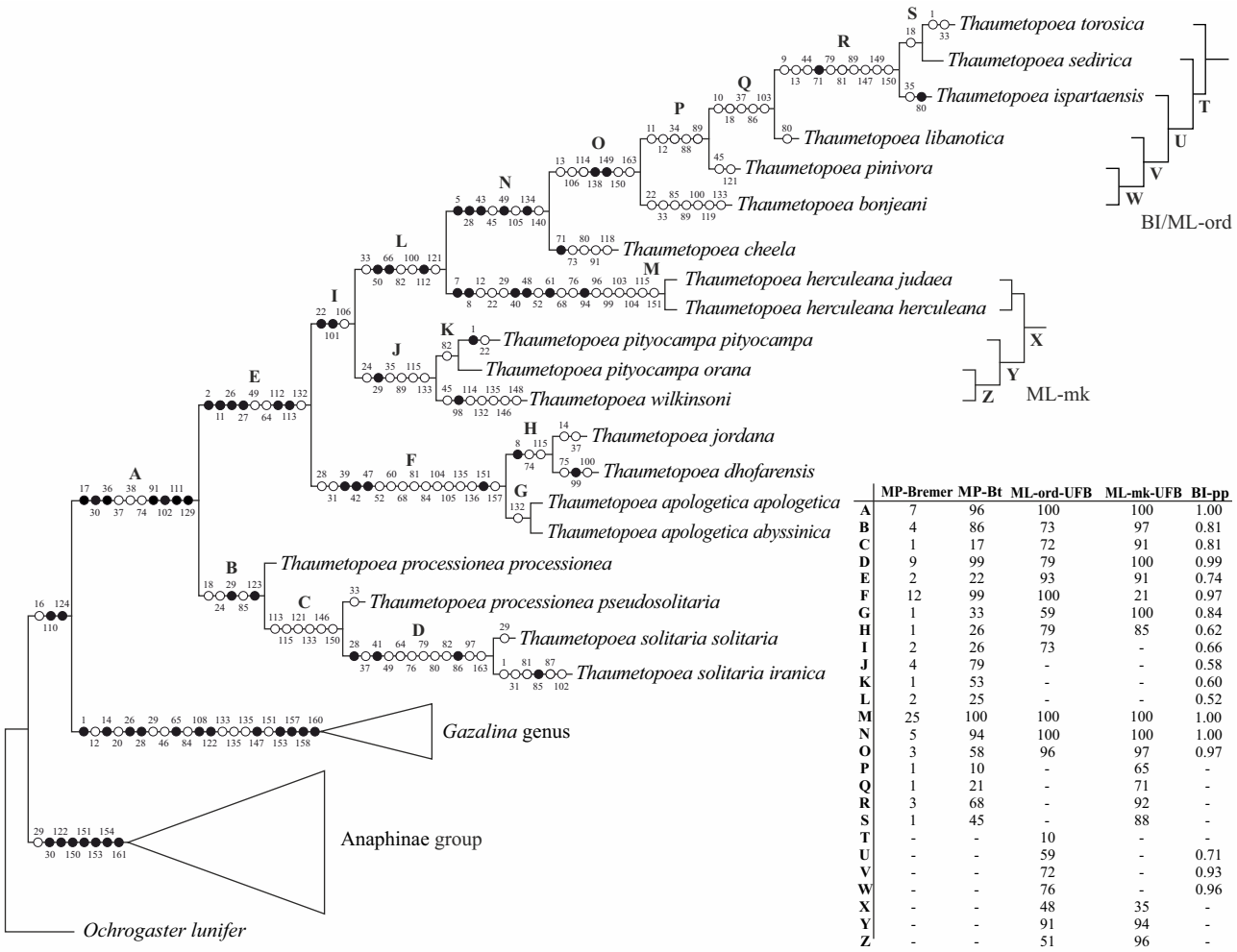


Figure 3



	MP-Bremer	MP-Bt	ML-ord-UFB	ML-mk-UFB	BI-pp
A	7	96	100	100	1.00
B	4	86	73	97	0.81
C	1	17	72	91	0.81
D	9	99	79	100	0.99
E	2	22	93	91	0.74
F	12	99	100	21	0.97
G	1	33	59	100	0.84
H	1	26	79	85	0.62
I	2	26	73	-	0.66
J	4	79	-	-	0.58
K	1	53	-	-	0.60
L	2	25	-	-	0.52
M	25	100	100	100	1.00
N	5	94	100	100	1.00
O	3	58	96	97	0.97
P	1	10	-	65	-
Q	1	21	-	71	-
R	3	68	-	92	-
S	1	45	-	88	-
T	-	-	10	-	-
U	-	-	59	-	0.71
V	-	-	72	-	0.93
W	-	-	76	-	0.96
X	-	-	48	35	-
Y	-	-	91	94	-
Z	-	-	51	96	-

Figure 4

	MP-Bremer	MP-Bt	ML-UFB	BI-pp
A	9	94	100	1.00
B	9	86	100	1.00
C	8	99	100	1.00
D	1	6	52	0.62
E	1	99	94	1.00
F	25	100	100	1.00
G	2	41	88	0.98
H	11	97	100	1.00
I	-	-	100	1.00
J	1	31	-	-
K	1	55	97	0.99
L	19	100	100	1.00
M	21	86	99	1.00
N	-	-	100	1.00
O	14	97	100	1.00
P	4	65	90	0.88
Q	6	58	69	0.59
R	3	72	95	0.96
S	2	66	99	0.98
T	1	45	96	0.95
U	-	-	90	-

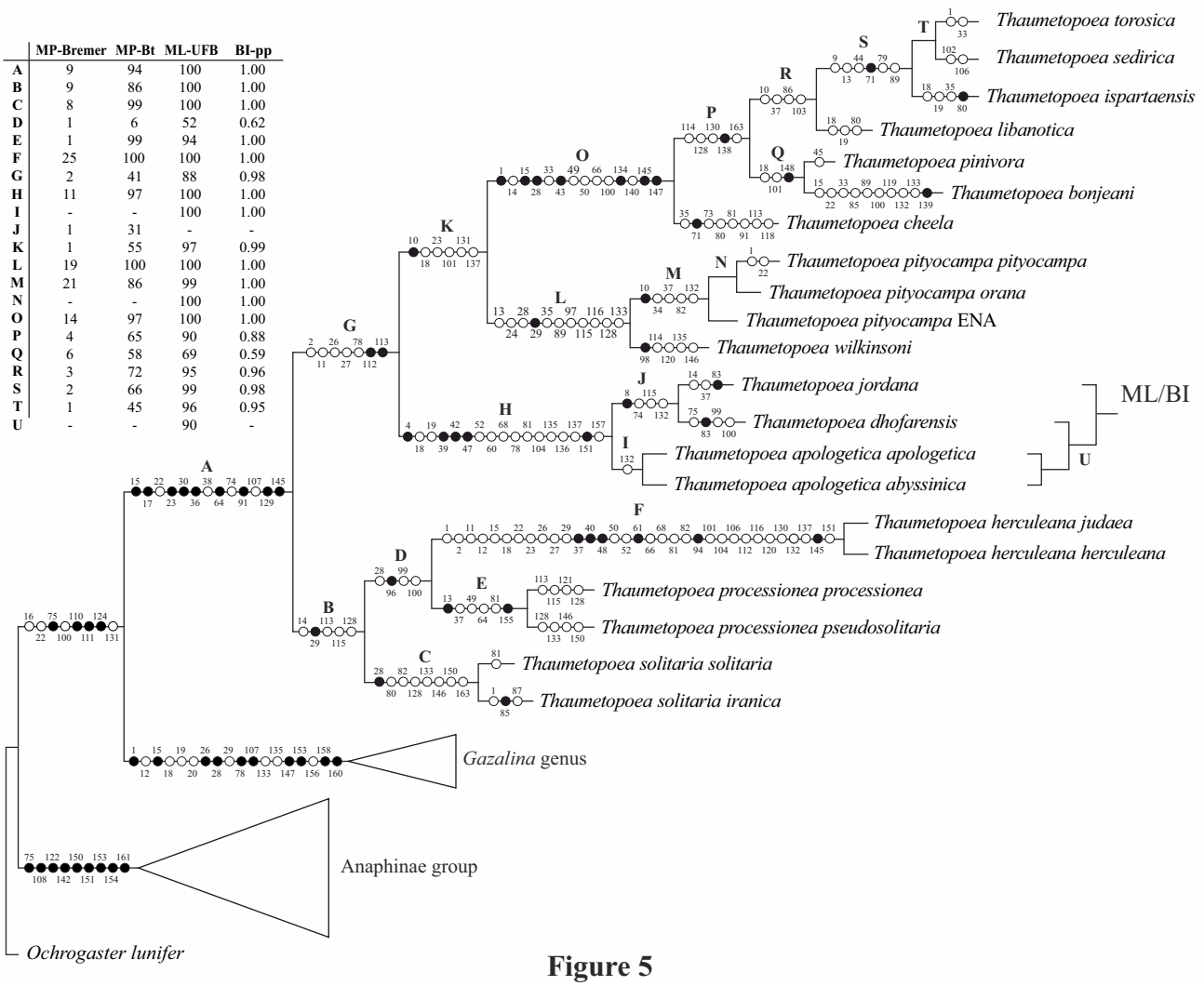


Figure 5

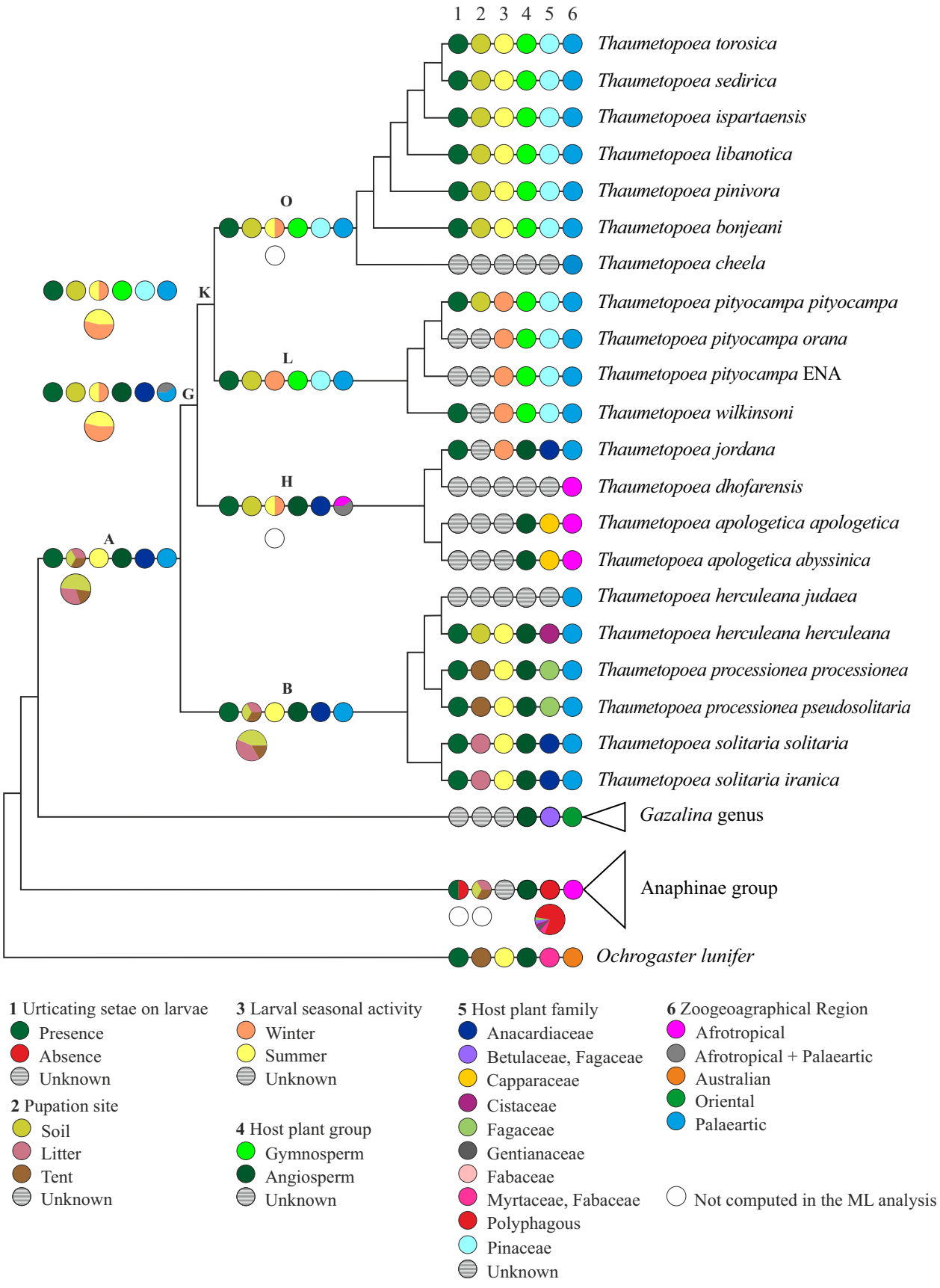

1 2 3 4 5 6
Thaumetopoea torosica
Thaumetopoea sedirica
Thaumetopoea ispartaensis
Thaumetopoea libanotica
Thaumetopoea pinivora
Thaumetopoea bonjeani
Thaumetopoea cheela
Thaumetopoea pityocampa pityocampa
Thaumetopoea pityocampa orana
Thaumetopoea pityocampa ENA
Thaumetopoea wilkinsoni
Thaumetopoea jordana
Thaumetopoea dhofarensis
Thaumetopoea apologetica apologetica
Thaumetopoea apologetica abyssinica
Thaumetopoea herculeana judaea
Thaumetopoea herculeana herculeana
Thaumetopoea processionea processionea
Thaumetopoea processionea pseudosolitaria
Thaumetopoea solitaria solitaria
Thaumetopoea solitaria iranica
Gazalina genus
Anaphinae group
Ochrogaster lunifer
O
K
L
G
H
A
B
1 Urticating setae on larvae
Presence
Absence
Unknown
2 Pupation site
Soil
Litter
Tent
Unknown
3 Larval seasonal activity
Winter
Summer
Unknown
4 Host plant group
Gymnosperm
Angiosperm
Unknown
5 Host plant family
Anacardiaceae
Betulaceae, Fagaceae
Capparaceae
Cistaceae
Fagaceae
Gentianaceae
Fabaceae
Myrtaceae, Fabaceae
Polyphagous
Pinaceae
Unknown
6 Zoogeoagraphical Region
Afrotropical
Afrotropical + Palaeartic
Australian
Oriental
Palaeartic
Not computed in the ML analysis

Figure 6

Figure S1. Species studied in this work. (a), *Thaumetopoea herculeana herculeana* - Algeria (BMHN(E)_1378597); (b), *Thaumetopoea processionea* - France (BMHN(E)_1378609); (c), *Thaumetopoea solitaria solitaria* - Cyprus (BMHN(E)_1378618); (d), *Thaumetopoea apologetica apologetica* - Kenya (BMHN(E)_1378606); (e), *Thaumetopoea dhofarensis* - Oman (WITT_TH73); (f), *Thaumetopoea jordana* - Palestine (ZSM_TH72); (g), *Thaumetopoea pityocampa pityocampa*- Greece (BMHN(E)_1378583); (h), *Thaumetopoea wilkinsoni* - Cyprus (BMHN(E)_1378572). The collection code associated to the specimen is provided in brackets (see Table S1).

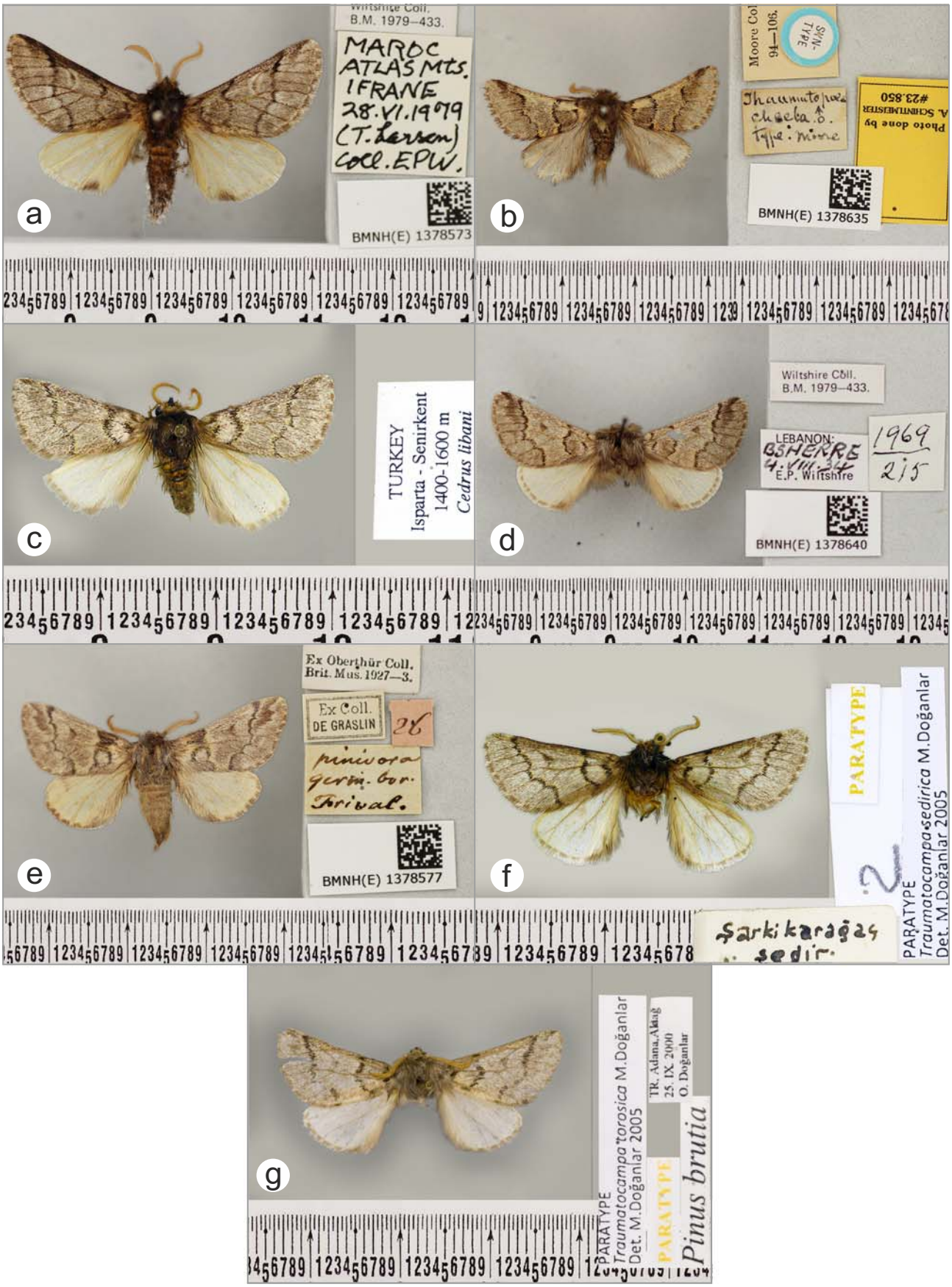


Figure S2. Species studied in this work. (a), *Thaumetopoea bonjeani* - Morocco (BMHN(E)_1378573); (b), *Thaumetopoea cheela* ST - India (BMHN(E)_1378635); (c), *Thaumetopoea ispartaensis* - Turkey (DAFNAE_TH30); (d), *Thaumetopoea libanotica* - Lebanon (BMNH(E)_1378640); (e), *Thaumetopoea pinivora* - France (BMHN(E)_1378577); (f), *Thaumetopoea sedirica* PT - Turkey (DAFNAE_TH29); (g), *Thaumetopoea torosica* PT - Turkey (DAFNAE_TH64). The collection code associated to the specimen is provided in brackets (see Table S1).

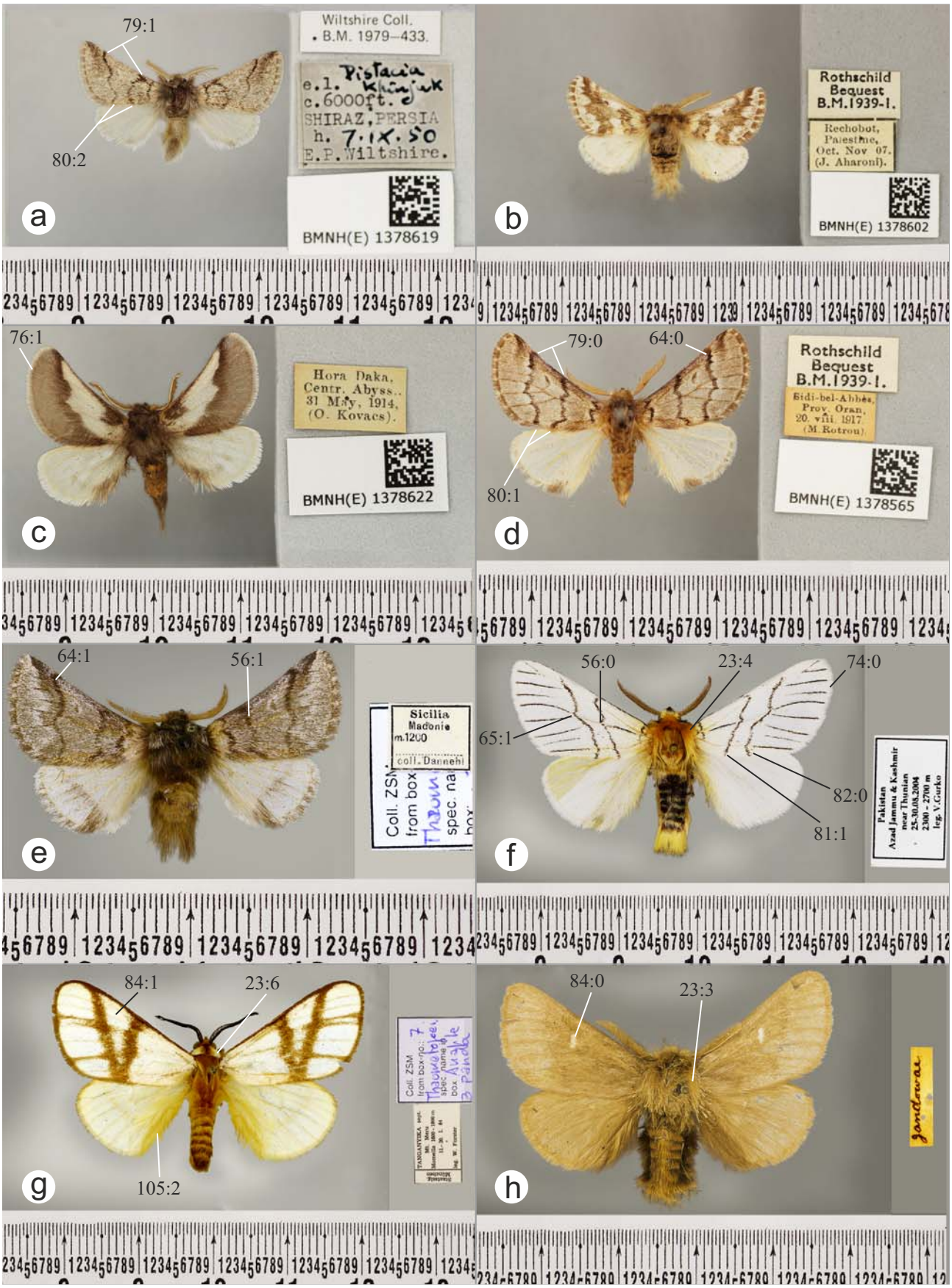


Figure S3. Species studied in this work. (a), *Thaumetopoea apologetica abyssinica* - Abyssinia (BMHN(E)_1378622); (b), *Thaumetopoea herculeana judaea* - Palestine (BMHN(E)_1378602); (c), *Thaumetopoea pityocampa orana* - Algeria (BMHN(E)_1378565); (d), *Thaumetopoea solitaria iranica* - Persia (BMHN(E)_1378619); (e), *Thaumetopoea processionea pseudosolitaria* - Italy (ZSM_TH66); (f), *Gazalina apsara* - Pakistan (DAFNAE_TH77); (g), *Anaphe panda panda* - Tanzania (DAFNAE_TH78); (h), *Ochrogaster lunifer* - Australia (DAFNAE_TH85). The collection code associated to the specimen is provided in brackets (see Table S1).

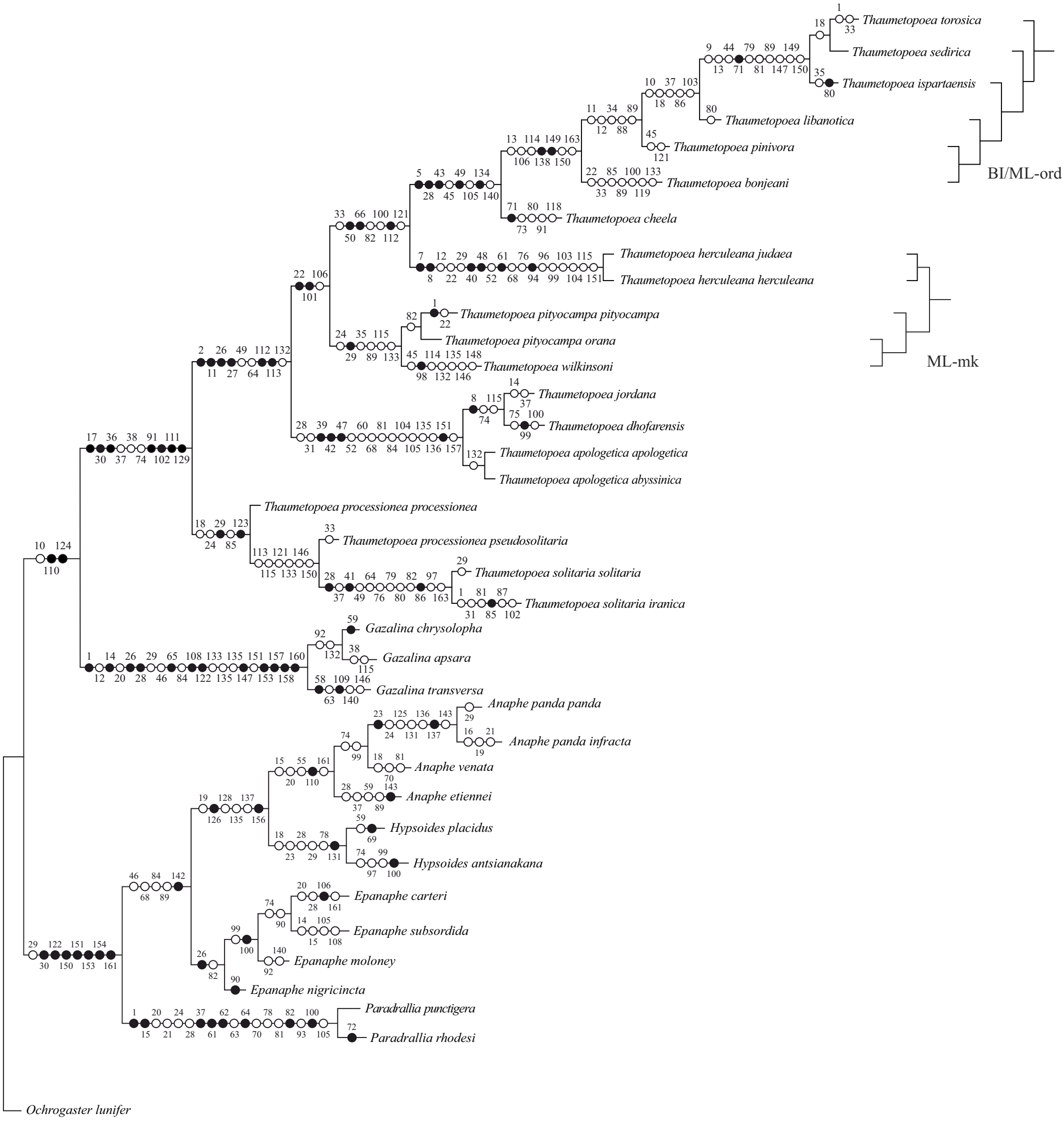


Figure S4. The most parsimonious tree (Length= 534 , CI= 0.537, RI= 0.797) obtained from Thau.morph dataset. Optimization: only unambiguous changes are mapped. Black circles, synapomorphic characters; white circles, homoplasious changes. Partial topology shows the difference with ML and BI trees. MP-Bremer, Bremer support to the node; MP-bt, bootstrap support to the node; ML-ord-UFB, ultrafast bootstrap support to the node, with ordered model; ML-mk-UFB, ultrafast bootstrap support to the node, with unordered model; BI-pp, posterior probability support to the node.

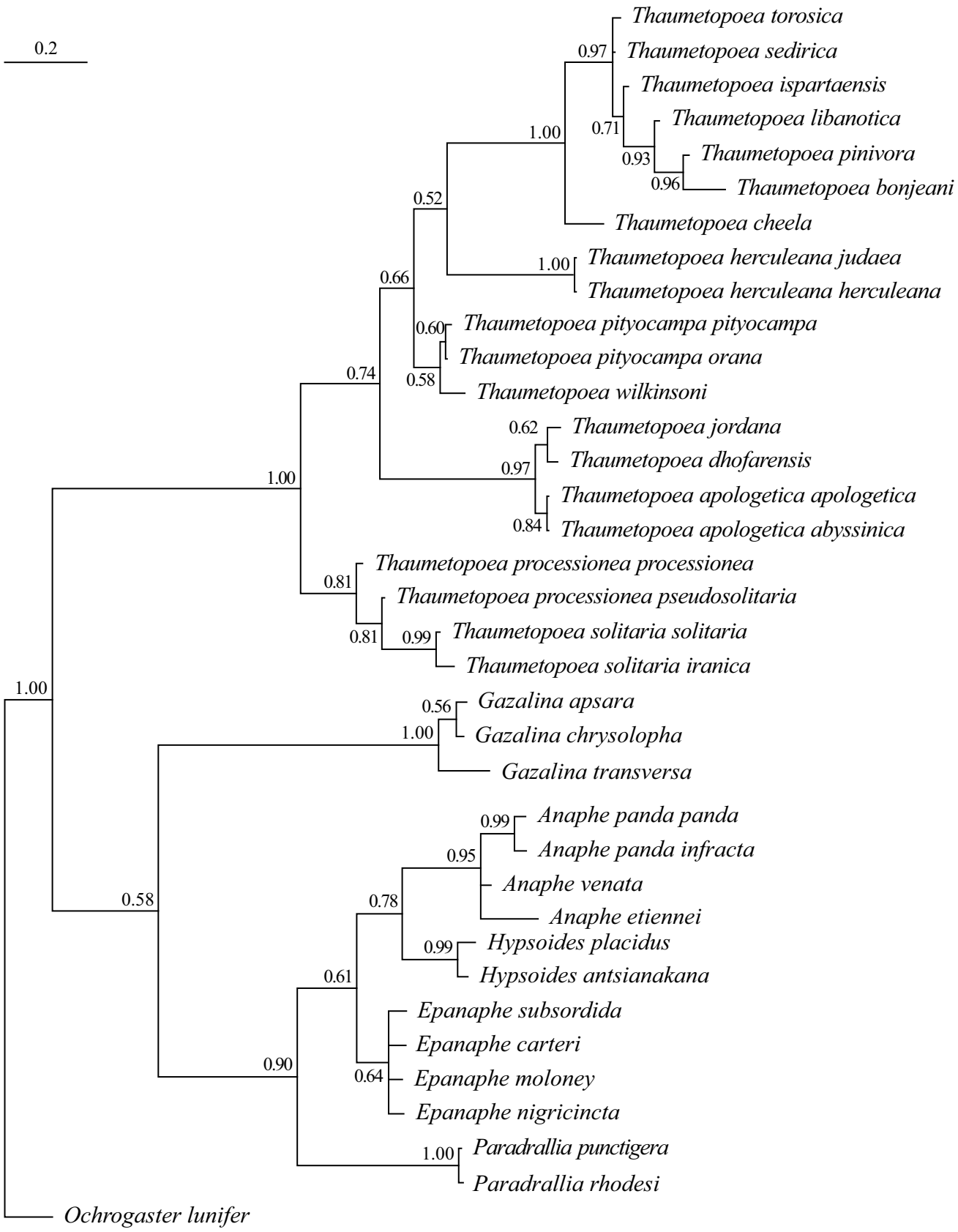


Figure S5. Bayesian tree inferred from Thau.morph dataset. Values at the base of each clade correspond to posterior probabilities.

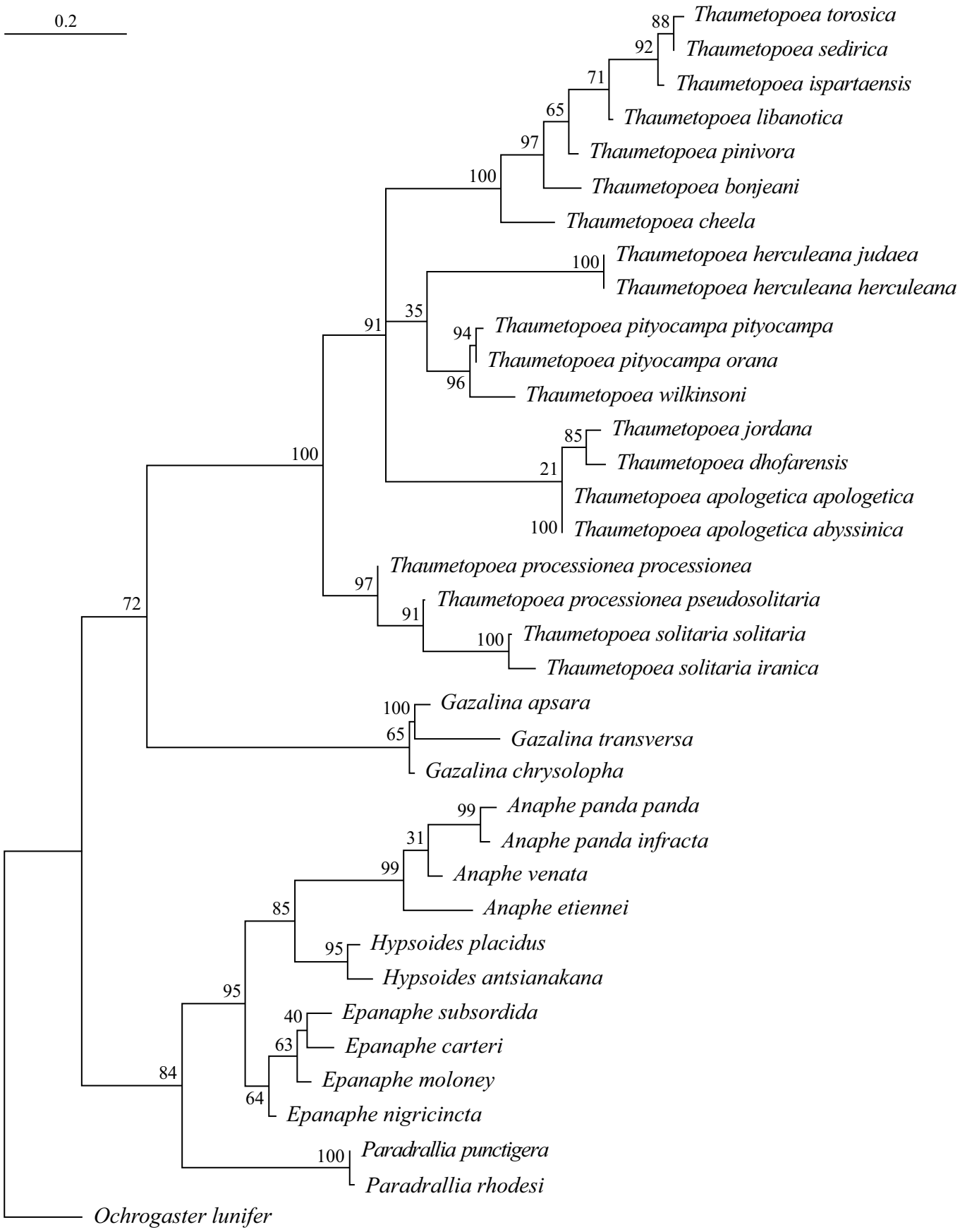


Figure S6. Maximum Likelihood tree (-ln=2971.8323) inferred from Thau.morph dataset performed using MK+FQ+ASC+G4 evolutionary model. Values at the base of each clade correspond to bootstrap support.

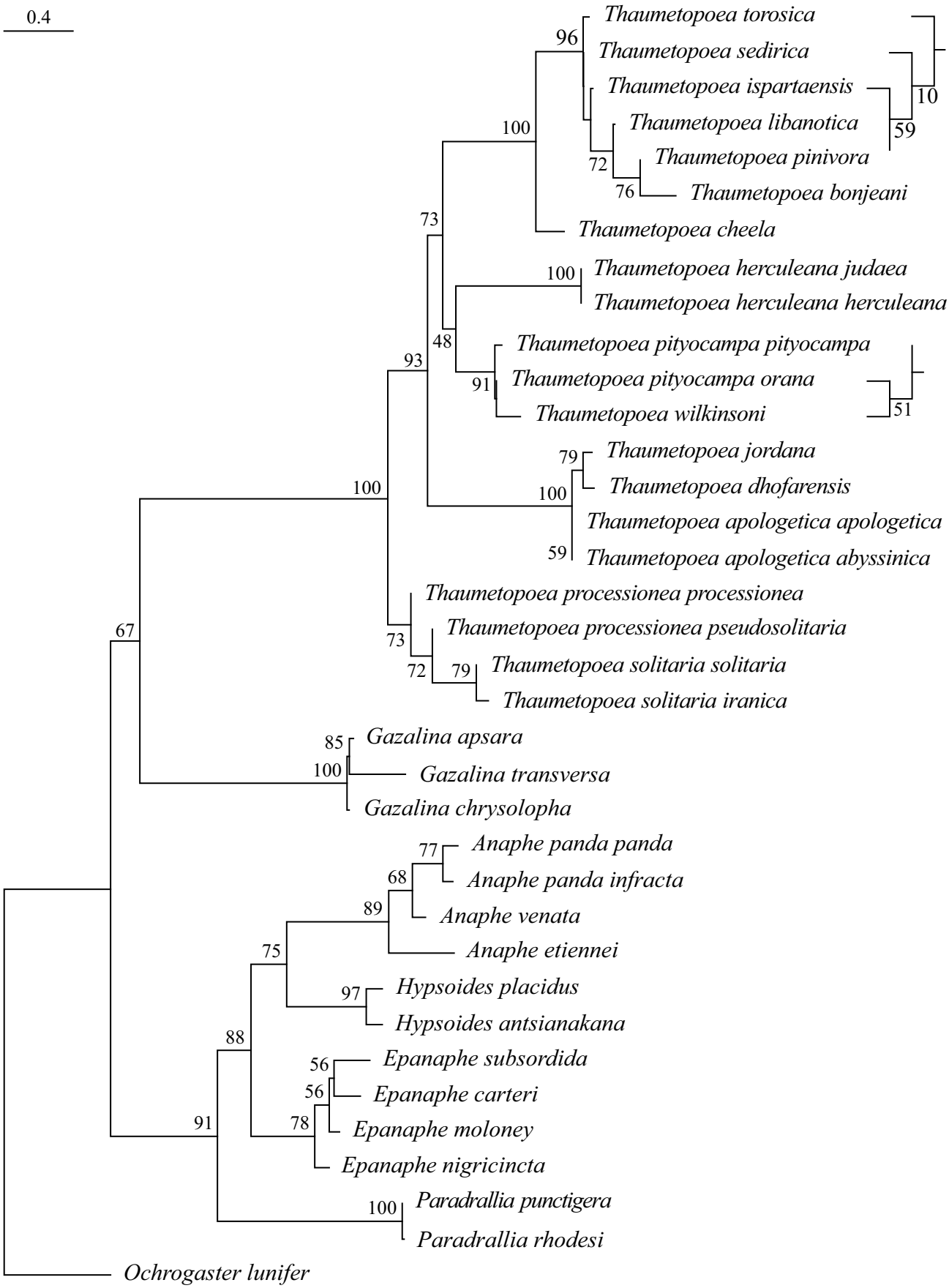


Figure S7. Maximum Likelihood tree (-ln=2536.9751) inferred from Thau.morph dataset performed using ORDERED+FQ+ASC+G4 evolutionary model. Values at the base of each clade correspond to bootstrap support. Partial topologies on right of the main tree to visualize better the relationships in those cases.

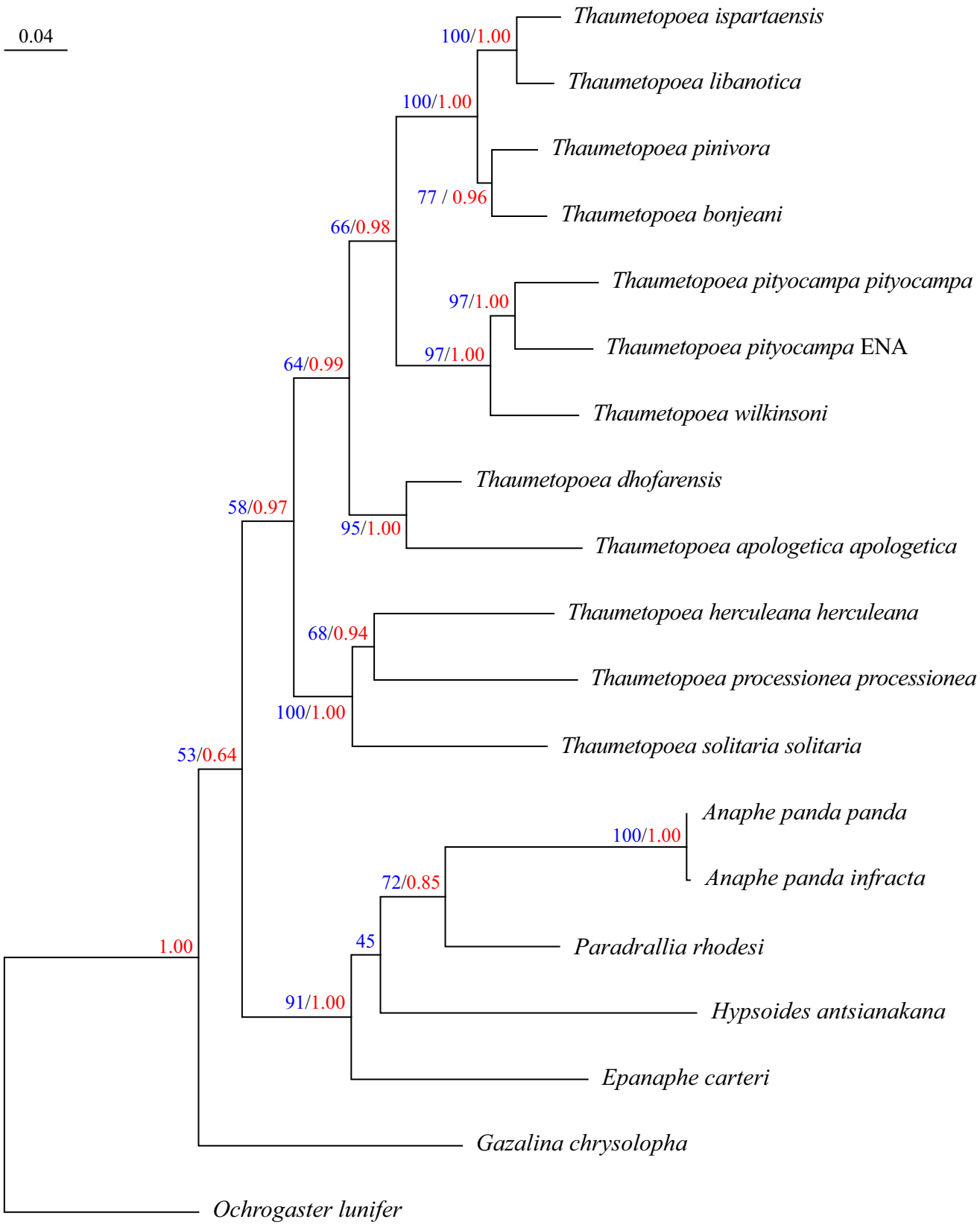


Figure S8. Maximum Likelihood tree (-ln=27774.0440) inferred from Thau.DNA dataset performed using GTR+I+G4 evolutionary model. Values at the base of each clade correspond to bootstrap support (in blue) and posterior probabilities (in red).

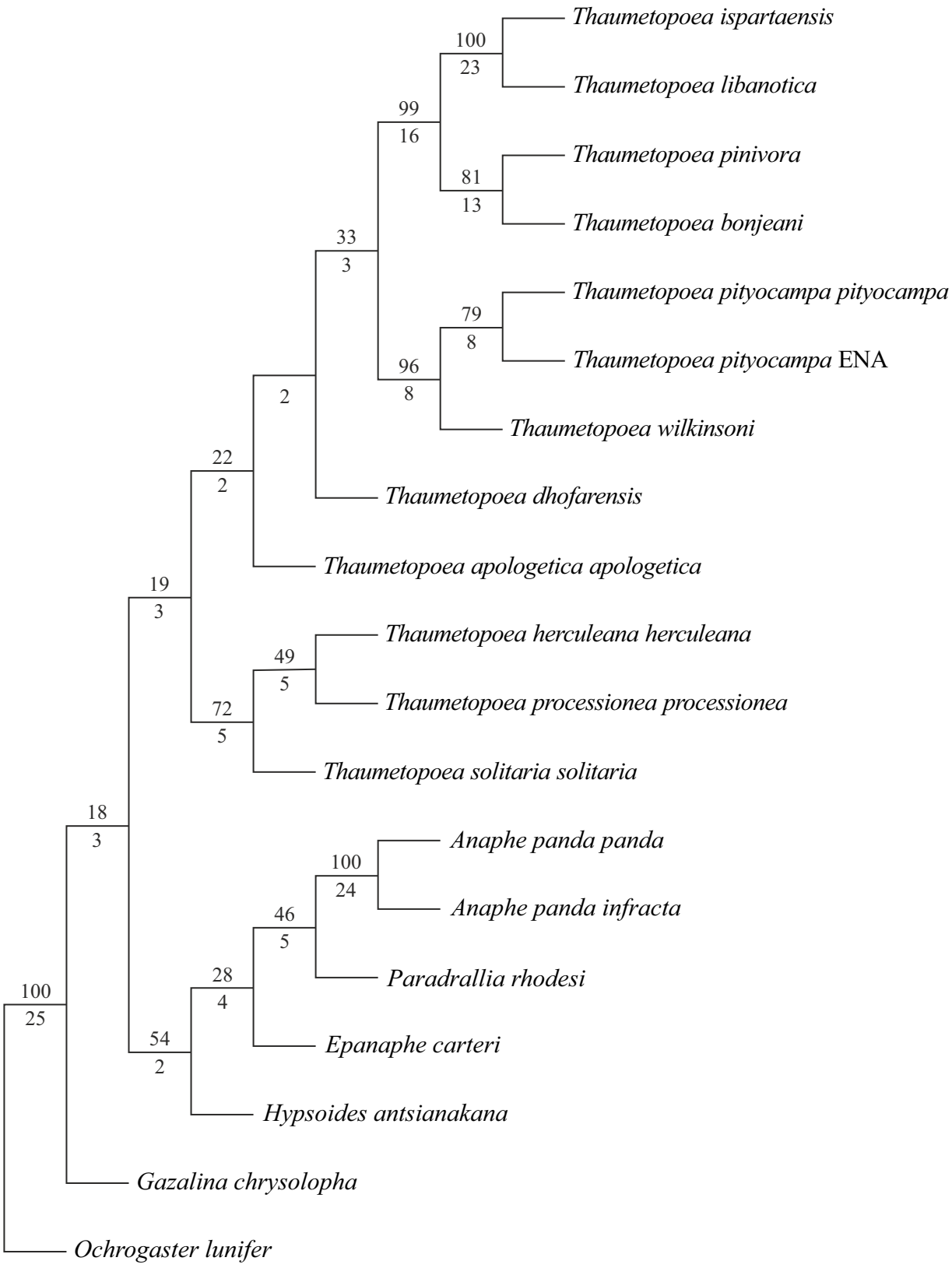


Figure S9. Most parsimonious trees inferred from Thau.DNA dataset (TBR) (Length= 4484, CI= 0.623, RI= 0.494). Values at the base of each clade correspond to bootstrap support (above) and Bremer support (below).

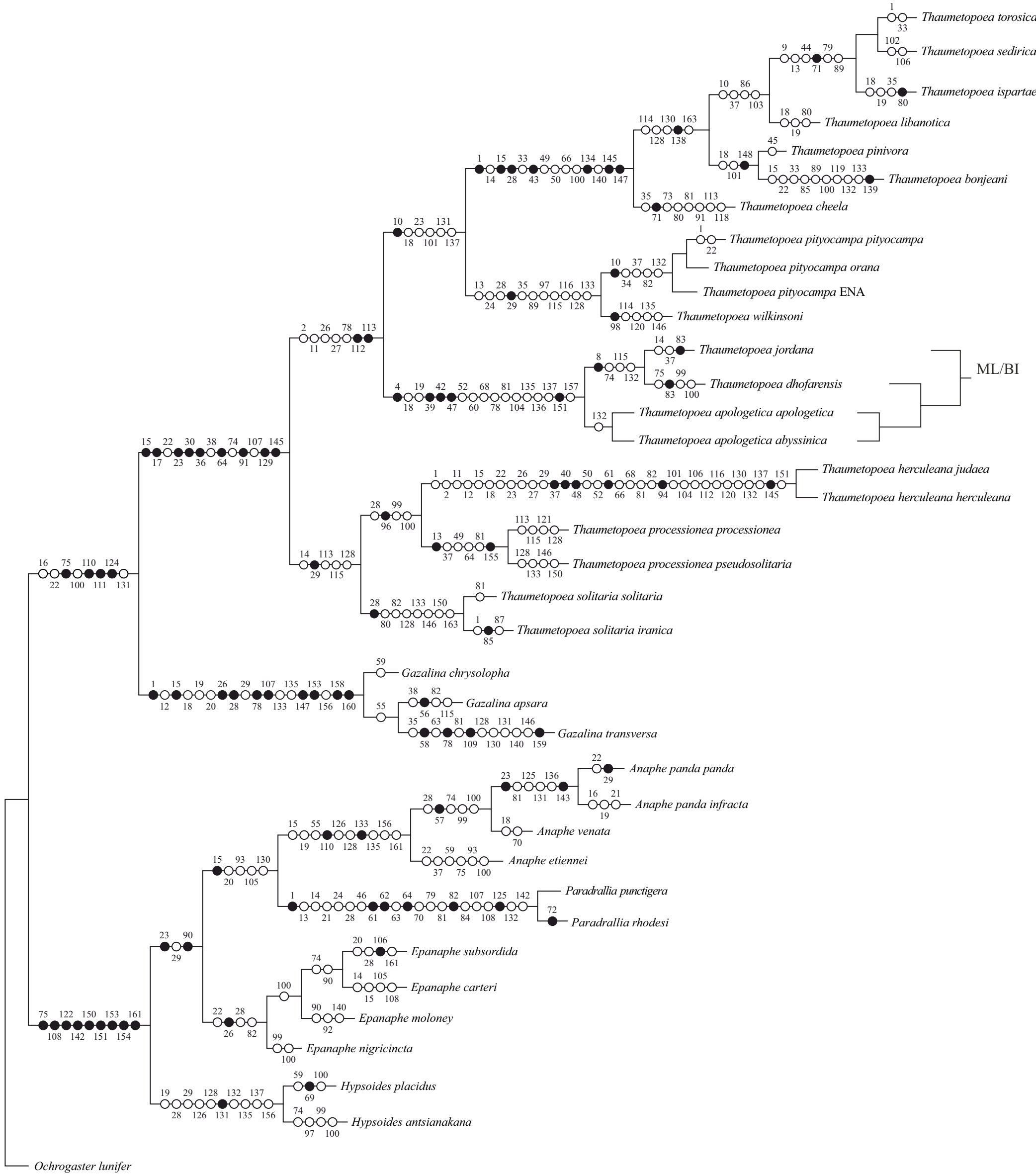


Figure S10 The most parsimonious tree (Length= 5046 , CI= 0.611, RI= 0.569) obtained from Thau.tot.comp dataset. Optimization: only unambiguous changes are mapped. Black circles, synapomorphic characters; white circles, homoplasious changes. Partial topology shows the difference with ML and BI trees. MP-Bremer, Bremer support to the node; MP-bt, bootstrap support to the node; ML-ord-UFB, ultrafast bootstrap support to the node, obtained from ordered model; ML-mk-UFB, ultrafast bootstrap support to the node, obtained from unordered model; BI-pp, posterior probability support to the node.

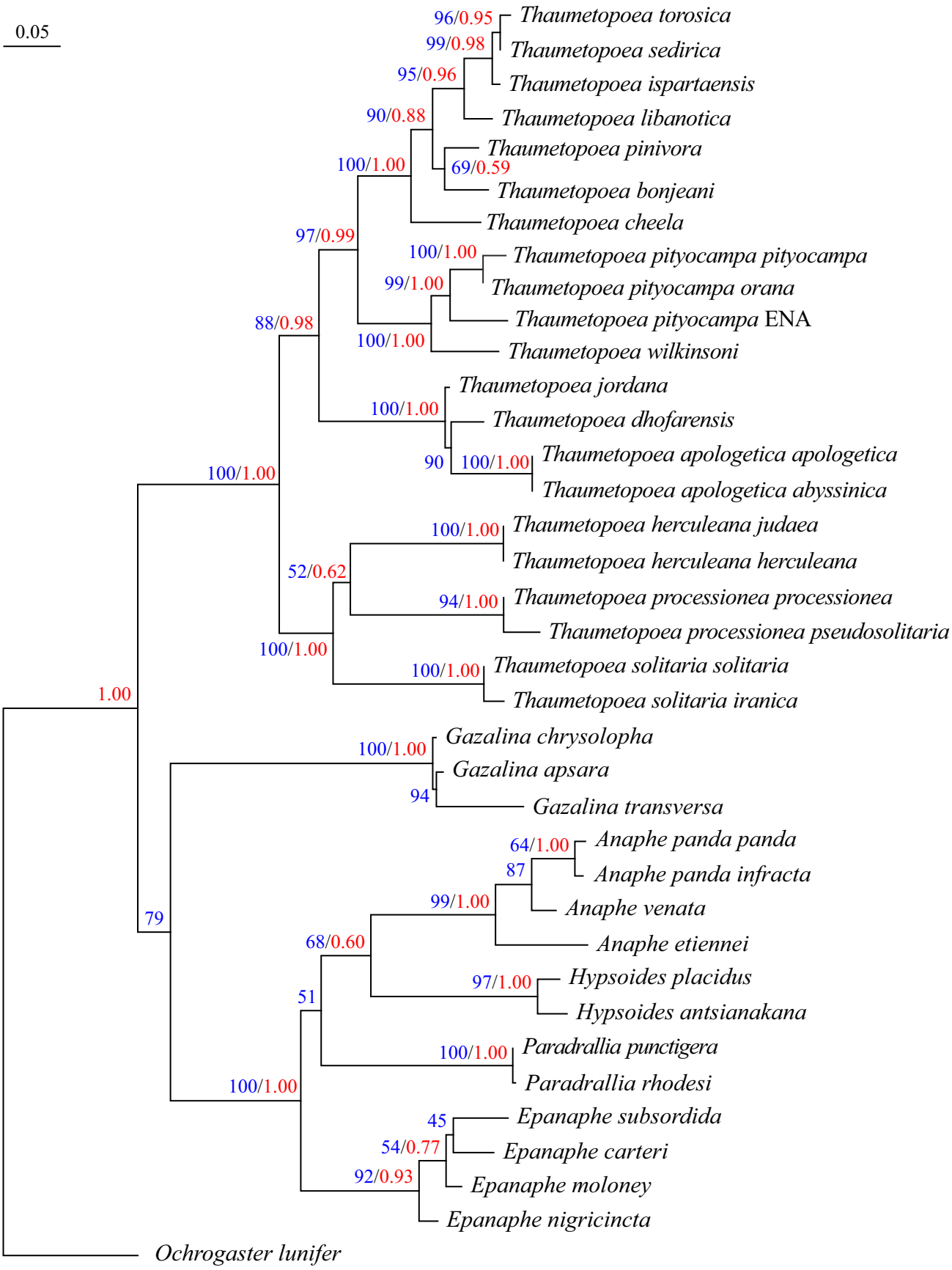


Figure S11. Maximum Likelihood tree (-ln=30433.391857) inferred from Thau.tot.comp dataset, performed using GTR+I+G4 evolutionary model for molecular partition and ORDERED+FQ+ASC+G4 evolutionary model for morphological partition. Values at the base of each clade correspond to bootstrap support (in blue) and posterior probabilities (in red).

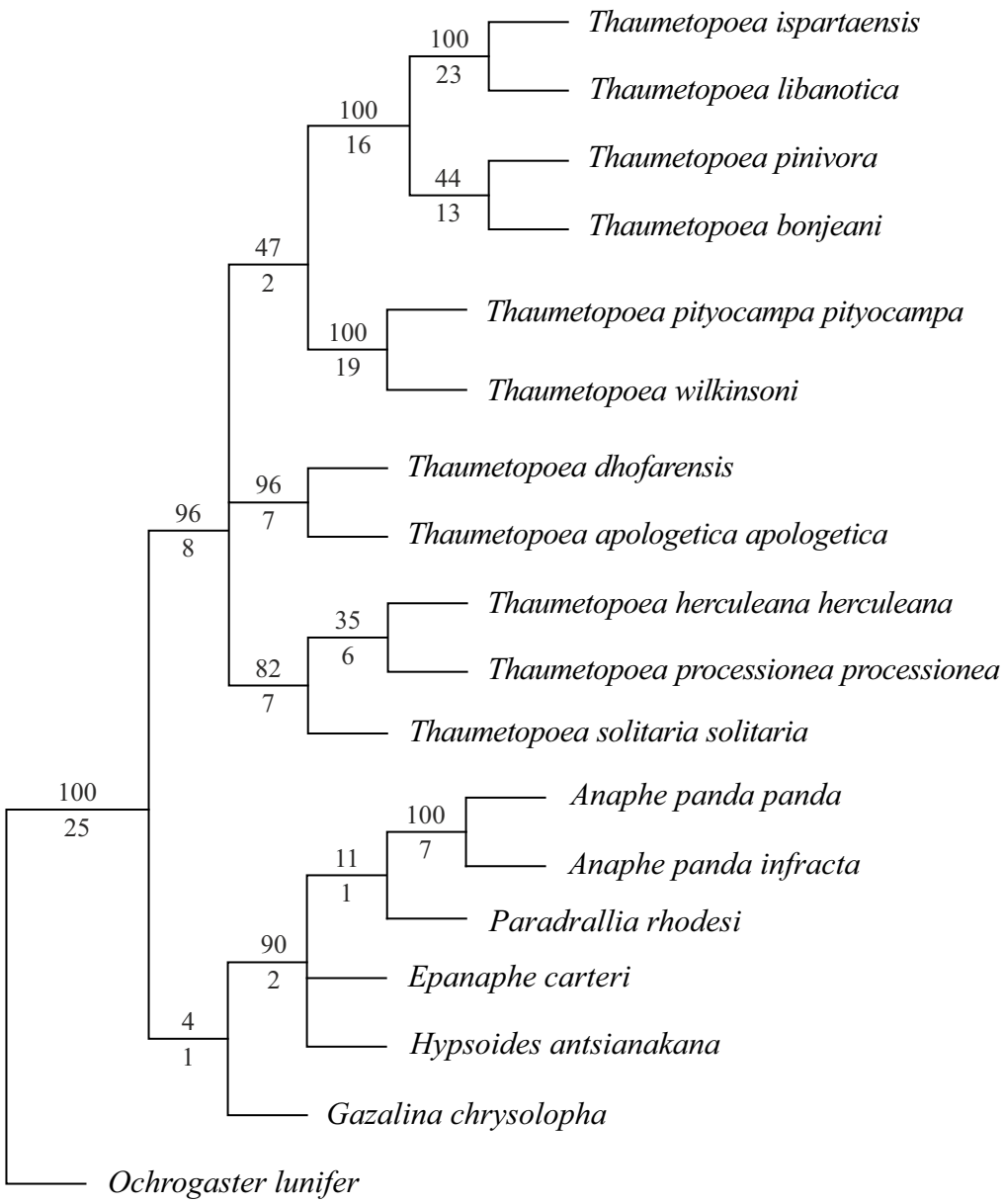


Figure S12. Strict consensus of three most parsimonious trees inferred from Thau.tot.red dataset (TBR) (Length= 4658, CI= 0.641 , RI= 0.511). Values at the base of each clade correspond to bootstrap support (above) and Bremer support (below).

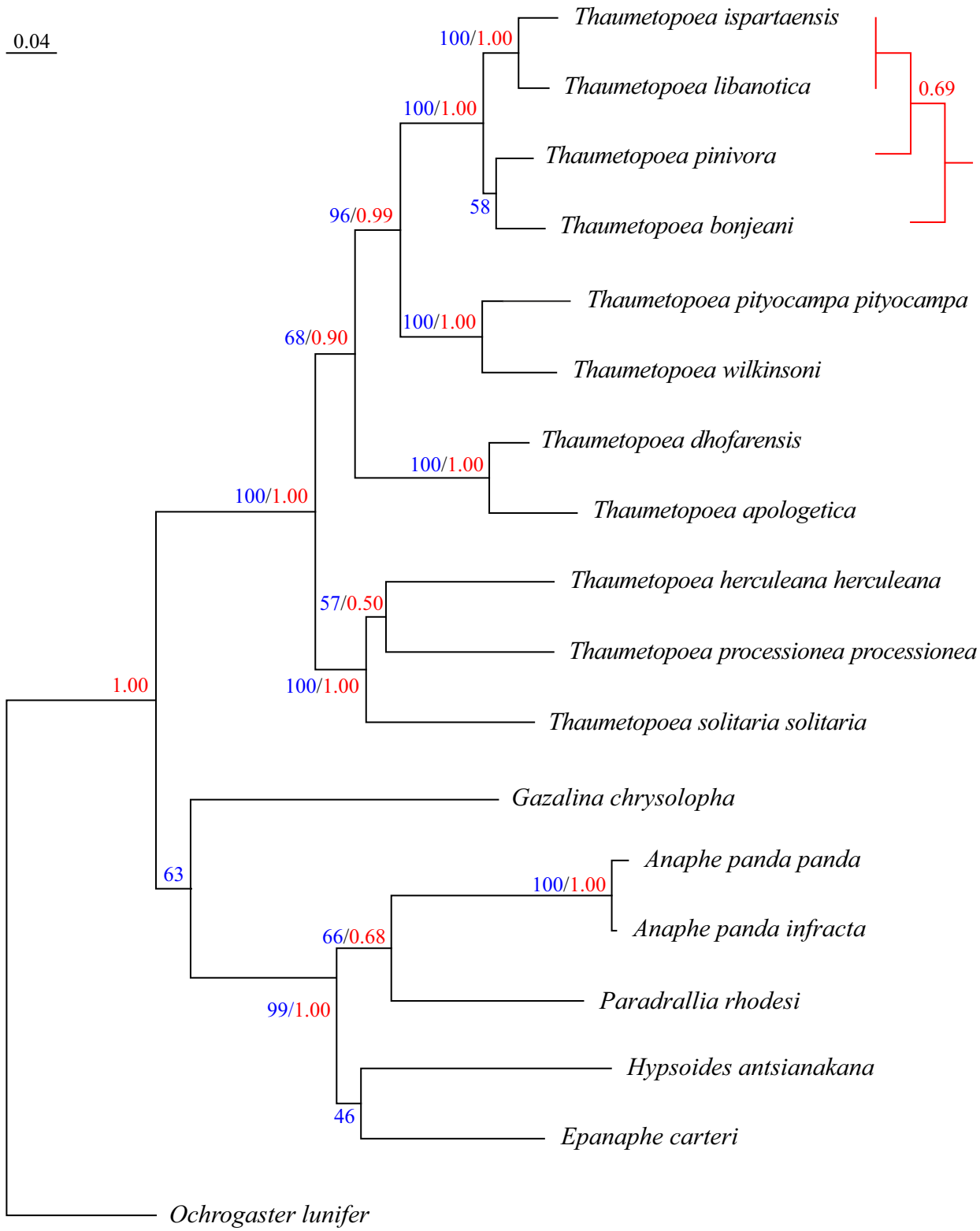


Figure S13. Maximum Likelihood tree (-ln=29835.0572) inferred from Thau.tot.red dataset, performed using GTR+I+G4 evolutionary model with molecular partition and ORDERED+FQ+G4 evolutionary model with morphological partition. Values at the base of each clade correspond to bootstrap support (in blue) and posterior probabilities (in red). Partial topology in red-colour shows the difference with BI tree.

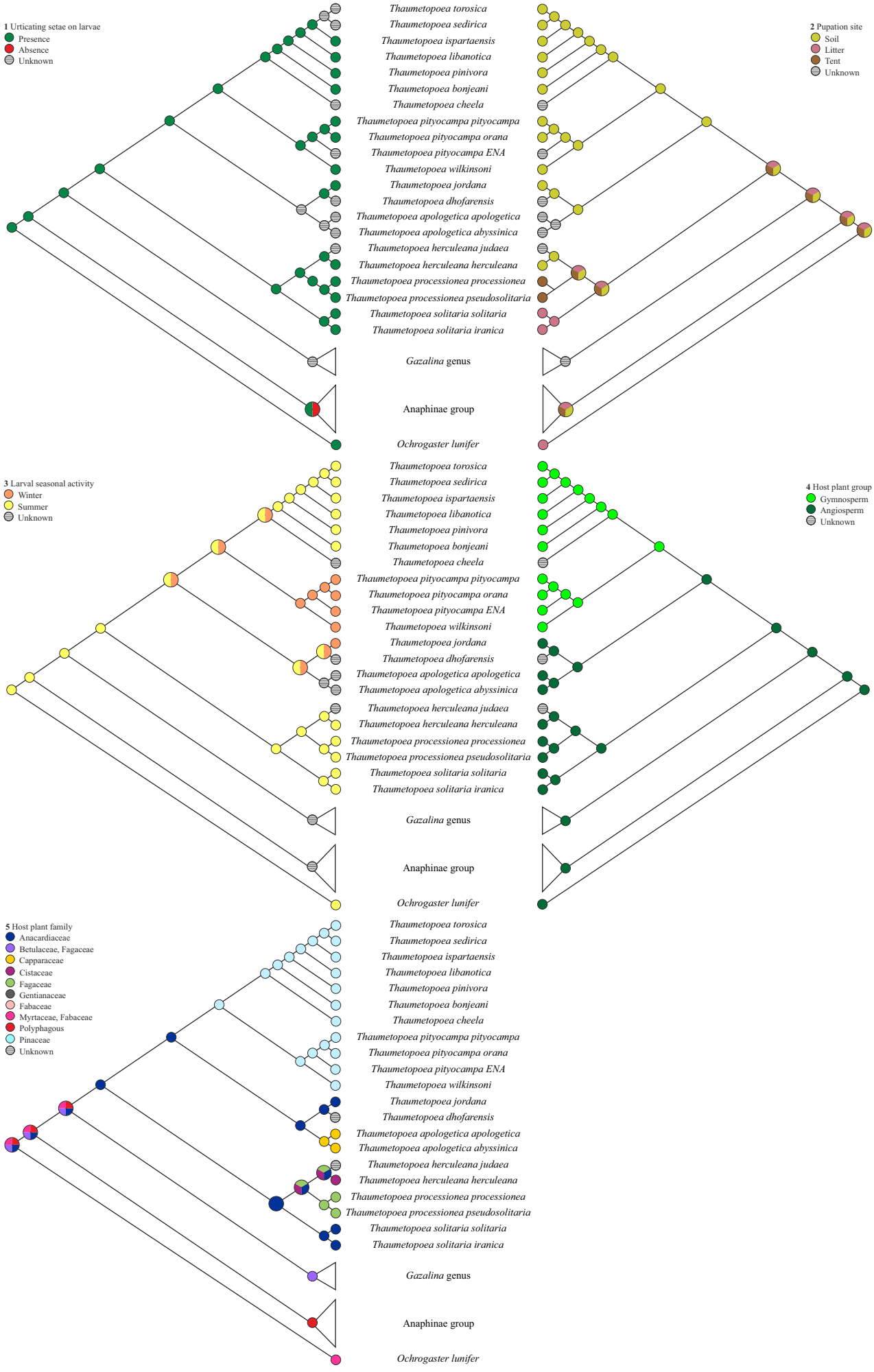


Figure S14. Mapping of biological and ecological traits for each species, reconstructed on reference tree applying maximum parsimony algorithm implemented in Mesquite software. Pie charts are enlarged on ambiguous ancestral nodes to increase clarity.

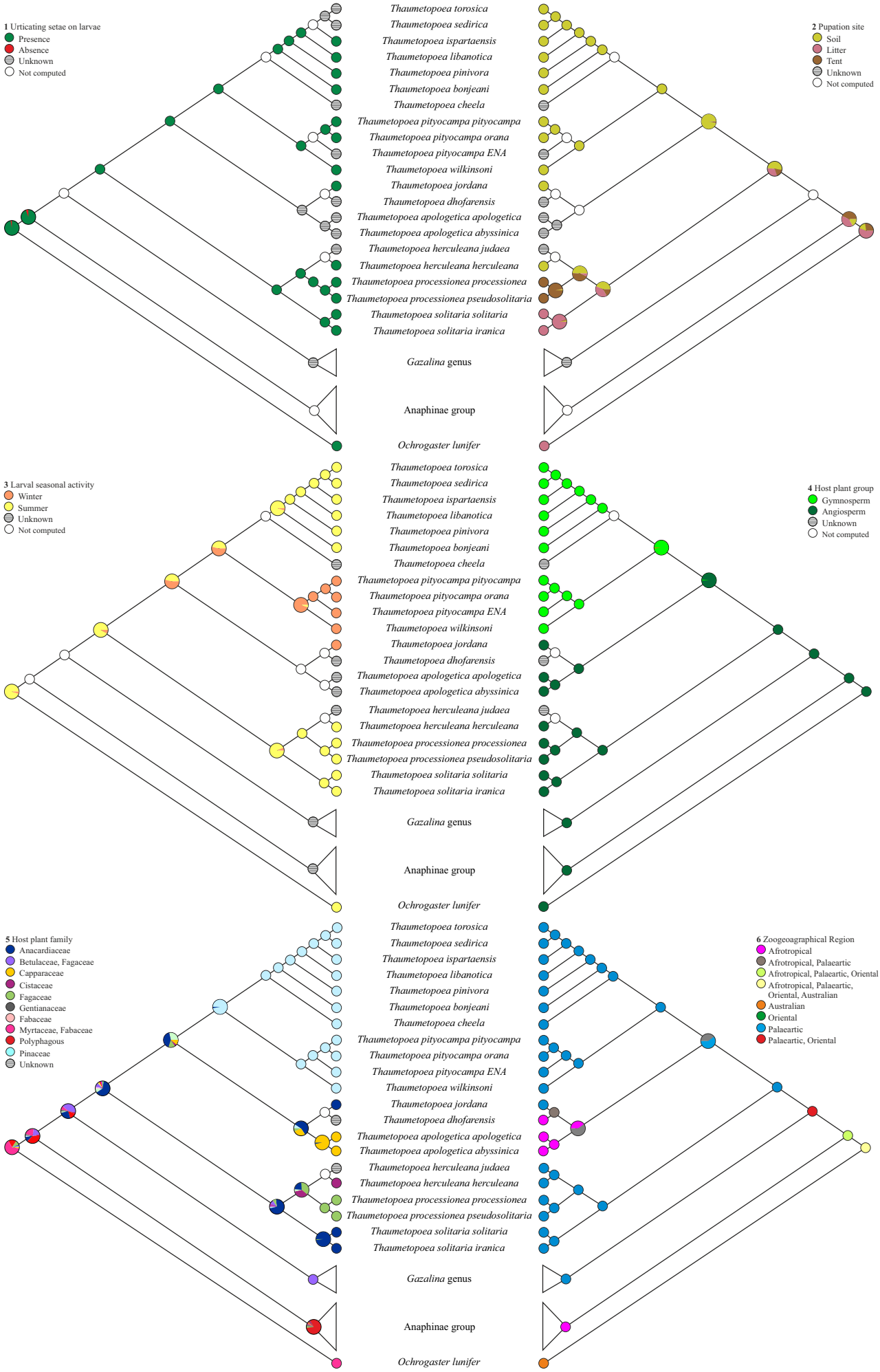


Figure S15. Mapping of biological and ecological traits for each species, reconstructed on reference tree applying maximum likelyhood algorithm implemented in Mesquite software. Pie charts sections represent the proportional likelihood of each character state at the node and are enlarged on ambiguous ancestral nodes to increase clarity. Trait number 6 was mapped with S-DIVA software

Tab.S1A. Specimens examined in this work. List of collections where specimens are preserved: BMNH: Natural History Museum (formerly British Museum of Natural History), London UK. DAFNAE: Department of Agronomy, Food, Natural resources, Animals and Environment - University of Padua, Padua Italy. Private coll.: Private collection of A. Schintlmeister, Dresden Germany. RBINS: Royal Belgian Institute of Natural Sciences, Brussels Belgium. RMCA: Royal Museum of Central Africa, Tervuren Belgium. MCZR: Museo Civico di Zoologia, Rome Italy. MZUR: Museo di Zoologia, 'Sapienza' University of Rome, Rome Italy. WITT: Witt Museum, Munich Germany. ZSM: The Bavarian State Collection of Zoology, Munich Germany. HT= holotype, PT= paratype, ST= syntype.

N°	Ref. slides	Taxon	Locality	Ref. Specimens
1	Noto_2098	*Thaumetopoea apologetica abyssinica*	Ethiopia	BMNH_1378622
2	Noto_2099	*Thaumetopoea apologetica abyssinica*	Ethiopia	BMNH_1378623
3	Noto_2100	*Thaumetopoea apologetica abyssinica*	Ethiopia	BMNH_1378624
4	Noto_2101	*Thaumetopoea apologetica abyssinica*	Ethiopia	BMNH_1378625
5	Noto_298	*Thaumetopoea apologetica abyssinica*	Ethiopia	BMNH
6	Noto_2095	*Thaumetopoea apologetica apologetica*	Arabia	BMNH_1378605
7	Noto_2096	*Thaumetopoea apologetica apologetica*	Arabia	BMNH_1378607
8	Noto_2097	*Thaumetopoea apologetica apologetica*	Arabia	BMNH_1378603
9	Noto_297	*Thaumetopoea apologetica apologetica*	Uganda	BMNH
10	TH 4	*Thaumetopoea apologetica apologetica*	Zimbabwe	Private coll._#29.538
11	TH 5	*Thaumetopoea apologetica apologetica*	Zimbabwe	Private coll._#29.540
12	TH 56	*Thaumetopoea apologetica apologetica*	Zimbabwe	Private coll._#29.537
13	TH 57	*Thaumetopoea apologetica apologetica*	Zimbabwe	Private coll._#29.539
14	TH 58	*Thaumetopoea apologetica apologetica*	Tzaneen	Private coll._#29.541
15	TH 59	*Thaumetopoea apologetica apologetica*	South Africa	Private coll._#29.542
16	TH 6	*Thaumetopoea apologetica apologetica*	Namibia	Private coll._#29.543
17	TH 60	*Thaumetopoea apologetica apologetica*	Zimbabwe	Private coll._#29.544
18	Noto_2083	*Thaumetopoea bonjeani*	Morocco	BMNH_1378573
19	TH 18	*Thaumetopoea bonjeani*	Algeria - Tala Guilef	DAFNAE_TH18
20	TH 19	*Thaumetopoea bonjeani*	Algeria - Tala Guilef	DAFNAE_TH19
21	TH 20	*Thaumetopoea bonjeani*	Algeria - Tala Guilef	DAFNAE_TH20
22	TH 21	*Thaumetopoea bonjeani*	Algeria - Tizi Oujavoub	DAFNAE_TH21
23	TH 22	*Thaumetopoea bonjeani*	Algeria - Tizi Oujavoub	DAFNAE_TH22
24	TH 23	*Thaumetopoea bonjeani*	Algeria - Tizi Oujavoub	DAFNAE_TH23
25	Noto_2111	*Thaumetopoea cheela* ST	India	BMNH_1378635
26	TH 73	*Thaumetopoea dhofarensis*	Oman	WITT_TH73
27	TH 74	*Thaumetopoea dhofarensis*	Oman	WITT_TH74
28	Whiltshire_1911	*Thaumetopoea dhofarensis*	Oman	BMNH
29	Noto_2084	*Thaumetopoea herculeana*	Algeria	BMNH_1378597
30	Noto_2085	*Thaumetopoea herculeana*	Algeria	BMNH_1378596
31	Noto_2086	*Thaumetopoea herculeana*	Morocco	BMNH_1378599
32	Noto_2087	*Thaumetopoea herculeana*	Morocco	BMNH_1378600
33	Noto_2088	*Thaumetopoea herculeana*	Morocco	BMNH_1378601
34	TH 26	*Thaumetopoea herculeana*	Spain	DAFNAE_TH26
35	TH 33	*Thaumetopoea herculeana*	Spain	MZUR_TH33
36	TH 34	*Thaumetopoea herculeana*	Algeria	MZUR_TH34
37	Noto_2089	*Thaumetopoea herculeana judaea*	Palestine	BMNH_1378602
38	Noto_2090	*Thaumetopoea herculeana judaea*	Palestine	BMNH_1378638
39	TH 30	*Thaumetopoea ispartaensis*	Turkey - Isparta	DAFNAE_TH30

Tab.S1B. Continue from previous page.

N°	Ref. slides	Taxon	Locality	Ref. Specimens
40	TH 61	*Thaumetopoea ispartaensis*	Turkey - Isparta	DAFNAE_TH61
41	Noto_2080	*Thaumetopoea jordana*	Arabia	BMNH_1378574
42	Noto_2082	*Thaumetopoea jordana*	Jordan	BMNH_1378576
43	TH 72	*Thaumetopoea jordana*	Palestine	ZSM_TH72
44	Noto_2081	*Thaumetopoea jordana* *	Jordan	BMNH_1378575
45	Noto_2078	*Thaumetopoea pinivora*	France	BMNH_1378579
46	Noto_2079	*Thaumetopoea pinivora*	Poland	BMNH_1378578
47	TH 24	*Thaumetopoea pinivora*	Germany	WITT_TH24
48	TH 35	*Thaumetopoea pinivora*	Germany	MZUR_TH35
49	Noto_2073	*Thaumetopoea pityocampa orana*	Morocco	BMNH_1378569
50	Noto_2074	*Thaumetopoea pityocampa orana*	Tunisia	BMNH_1378567
51	Noto_2075	*Thaumetopoea pityocampa orana*	Algeria	BMNH_1378565
52	Noto_2076	*Thaumetopoea pityocampa orana*	Algeria	BMNH_1378566
53	Noto_2077	*Thaumetopoea pityocampa orana*	Algeria	BMNH_1378568
54	Noto_2064	*Thaumetopoea pityocampa pityocampa*	Portugal	BMNH_1378589
55	Noto_2065	*Thaumetopoea pityocampa pityocampa*	Portugal	BMNH_1378588
56	Noto_2066	*Thaumetopoea pityocampa pityocampa*	Algeria	BMNH_1378587
57	Noto_2067	*Thaumetopoea pityocampa pityocampa*	Majorca	BMNH_1378585
58	Noto_2068	*Thaumetopoea pityocampa pityocampa*	Republic of Macedonia	BMNH_1378584
59	Noto_2071	*Thaumetopoea pityocampa pityocampa*	Greece	BMNH_1378583
60	Noto_2072	*Thaumetopoea pityocampa pityocampa*	Greece	BMNH_1378582
61	Noto_2112	*Thaumetopoea pityocampa pityocampa*	France	BMNH_1378637
62	Noto_2113	*Thaumetopoea pityocampa pityocampa*	-	BMNH_1378563
63	TH 10	*Thaumetopoea pityocampa pityocampa*	Spain	DAFNAE_TH10
64	TH 12	*Thaumetopoea pityocampa pityocampa*	Italy	DAFNAE_TH12
65	TH 13	*Thaumetopoea pityocampa pityocampa*	Italy	DAFNAE_TH13
66	TH 17	*Thaumetopoea pityocampa pityocampa*	Italy	DAFNAE_TH17
67	TH 32	*Thaumetopoea pityocampa pityocampa*	Croatia	MZUR_TH32
68	TH 46	*Thaumetopoea pityocampa pityocampa*	Italy - Calbarina	DAFNAE_TH46
69	TH 47	*Thaumetopoea pityocampa pityocampa*	Italy - Calbarina	DAFNAE_TH47
70	TH 9	*Thaumetopoea pityocampa pityocampa*	Spain	DAFNAE_TH9
71	TH 27	*Thaumetopoea processionea*	Austria	WITT_TH27
72	Noto_1634	*Thaumetopoea processionea processionea*	France	BMNH
73	Noto_2091	*Thaumetopoea processionea processionea*	Hungary	BMNH_1378610
74	Noto_2092	*Thaumetopoea processionea processionea*	Switzerland	BMNH_1378611
75	Noto_2093	*Thaumetopoea processionea processionea*	France	BMNH_1378613
76	TH 36	*Thaumetopoea processionea processionea*	-	MZUR_TH36
77	TH 37	*Thaumetopoea processionea processionea*	Italy	MCZR_TH37
78	TH 63	*Thaumetopoea processionea processionea*	France	DAFNAE_TH63
79	Noto_2094	*Thaumetopoea processionea pseudosolitaria*	Macedonia	BMNH_1378612
80	TH 65	*Thaumetopoea processionea pseudosolitaria*	Republic of Macedonia	DAFNAE_TH65
81	TH 66	*Thaumetopoea processionea pseudosolitaria*	Italy - Sicilia	ZSM_TH66
82	TH 29	*Thaumetopoea sedirica* PT	Turkey - Sedir	DAFNAE_TH29
83	TH 31	*Thaumetopoea sedirica* PT *	Turkey - Sedir	DAFNAE_TH31
84	Noto_2105	*Thaumetopoea solitaria*	Cyprus	BMNH_1378618

*, specimens without genitalia

Tab.S1C. Continue from previous page.

N°	Ref. slides	Taxon	Locality	Ref. Specimens
85	Noto_2106	*Thaumetopoea solitaria*	Cyprus	BMNH_1378617
86	Noto_2107	*Thaumetopoea solitaria*	Iraq	BMNH_1378616
87	Noto_2108	*Thaumetopoea solitaria iranica*	Iran	BMNH_1378621
88	Noto_2109	*Thaumetopoea solitaria iranica*	Iran	BMNH_1378620
89	Noto_2110	*Thaumetopoea solitaria iranica*	Iran	BMNH_1378619
90	TH 3	*Thaumetopoea solitaria solitaria*	Pakistan	DAFNAE_TH3
91	TH 38	*Thaumetopoea solitaria solitaria*	Bulgaria	MCZR_TH38
92	TH 39	*Thaumetopoea solitaria solitaria*	Republic of Macedonia	MCZR_TH39
93	TH 64	*Thaumetopoea torosica*	Turkey - Adana	DAFNAE_TH64
94	TH 28	*Thaumetopoea torosica* *	Turkey - Isparta	DAFNAE_TH28
95	Noto_2104	*Thaumetopoea wilkinsoni*	Cyprus	BMNH_1378571
96	TH 11	*Thaumetopoea wilkinsoni*	Dishon - Israel	DAFNAE_TH11
97	TH 48	*Thaumetopoea wilkinsoni*	Israel - Dishon	DAFNAE_TH48
98	TH 49	*Thaumetopoea wilkinsoni*	Israel - Dishon	DAFNAE_TH49
99	TH 50	*Thaumetopoea wilkinsoni*	Israel - Western Negev	DAFNAE_TH50
100	TH 51	*Thaumetopoea wilkinsoni*	Israel - Western Negev	DAFNAE_TH51
101	TH 52	*Thaumetopoea wilkinsoni*	Israel - Judean Foothills	DAFNAE_TH52
102	TH 53	*Thaumetopoea wilkinsoni*	Israel - Judean Foothills	DAFNAE_TH53
103	TH 54	*Thaumetopoea wilkinsoni*	Israel - Southern Judean	DAFNAE_TH54
104	TH 55	*Thaumetopoea wilkinsoni*	Israel - Southern Judean	DAFNAE_TH55
105	Noto_2062	*Thaumetopoea wilkinsoni* [1]	Lebanon	BMNH_1378592
106	Noto_2063	*Thaumetopoea wilkinsoni* [1]	Lebanon	BMNH_1378593
107	Noto_2069	*Thaumetopoea wilkinsoni* [1]	Rhodes	BMNH_1378580
108	Noto_2070	*Thaumetopoea wilkinsoni* [1]	Rhodes	BMNH_1378581
109	Whiltshire_281	*Thaumetopoea wilkinsoni* HT	Cyprus	BMHN(E)_1378629
110	Noto_2102	*Thaumetopoea wilkinsoni* PT	Cyprus	BMNH_1378628
111	Noto_2103	*Thaumetopoea wilkinsoni* PT	Cyprus	BMNH_1378627
112	TH 75	*Anaphe etiennei*	DR of Congo	RMCA_TH75
113	Noto_271	*Anaphe panda*	South Africa	BMNH
114	TH 78	*Anaphe panda*	Tanzania	ZSM_TH78
115	TH 15	*Anaphe panda infracta*	Zambia	DAFNAE_TH15
116	TH 7	*Anaphe panda infracta*	DR of Congo	DAFNAE_TH7
117	TH 14	*Anaphe reticulata*	South Africa	DAFNAE_TH14
118	TH 76	*Anaphe venata*	DR of Congo	RMCA_TH76
119	TH 79	*Anaphe venata*	Congo da Lemba	RMCA_TH79
120	Noto_276	*Epanaphe* (*Anaphe*) *carteri*	Ghana	BMNH
121	Noto_275	*Epanaphe* (*Anaphe*) *moloney*	Zambia	BMNH
122	Noto_302	*Epanaphe* (*Anaphe*) *subsordida*	Nigeria	BMNH
123	TH 70	*Epanaphe carteri*	Angola	ZSM_TH70
124	TH 81	*Epanaphe maynei*	DR of Congo	RMCA_TH81
125	TH 69	*Epanaphe moloney*	Nigeria	ZSM_TH69
126	TH 80	*Epanaphe nigricincta*	Uganda	ZSM_TH80
127	Noto_278	*Epanaphe subsordida*	Cameroon	BMNH
128	TH 82	*Epanaphe subsordida*	DR of Congo	RMCA_TH82
129	TH 16	*Gazalina apsara*	Myanmar	WITT_TH16

[1], labelled as *T. pityocampa pityocampa* in BMNH; *, specimens without genitalia

Tab.S1D. Continue from previous page.

N°	Ref. slides	Taxon	Locality	Ref. Specimens
130	TH 67	*Gazalina apsara*	Myanmar	WITT_TH67
131	TH 77	*Gazalina apsara*	Pakistan	DAFNAE_TH77
132	TH 2	*Gazalina chrysolopha*	Pakistan	DAFNAE_TH2
133	TH 68	*Gazalina transversa*	Nepal	ZSM_TH68
134	TH 84	*Hypsoides antsianakana*	Madagascar	ZSM_TH84
135	TH 71	*Hypsoides placidus*	-	ZSM_TH71
136	TH 8	*Hypspoides antsianakana* *	Madagascar	DAFNAE_TH8
137	TH 85	*Ochrogaster lunifer*	Australia	DAFNAE_TH85
138	Noto_269	*Ochrogaster lunifer* (*Marane rubricorpus*)	Australia	BMNH
139	Noto_303	*Ochrogaster lunifer* (*ruptimacula*)	-	BMNH
140	Noto_270	*Ochrogaster lunifer* (*Teara contraria*)	Australia	BMNH
141	TH 83	*Paradrallia punctigera*	DR of Congo	RBINS_TH83
142	Noto_299	*Paradrallia rhodesi*	Zambia	BMNH
143	TH 1	*Paradrallia rhodesi*	Zambia	DAFNAE_TH1
144	TH 25	*Paradrallia rhodesi*	Zambia	DAFNAE_TH25

*, specimens without genitalia

Tab.S2. Mitochondrial genes used.

	cox1-cox2-atp6-atp8-cox3	partial *cob-nad1*	*nad4-nad5*
Thaumetopoea processionea processionea	HE963114	HE956699	HE864329
Thaumetopoea solitaria solitaria	HE963115	HE956700	HE864330
Thaumetopoea herculeana herculeana	HE963108	HE956693	HE864323
Thaumetopoea apologetica apologetica *	LT614824	—	—
Thaumetopoea dhofarensis *	LT614825	—	—
Thaumetopoea bonjeani	HE963107	HE956692	HE864322
Thaumetopoea pinivora	HE963111	HE956696	HE864326
Thaumetopoea libanotica	HE963110	HE956695	HE864325
Thaumetopoea ispartaensis	HE963109	HE956694	HE864324
Thaumetopoea pityocampa pityocampa	HE963112	HE956697	HE864327
Thaumetopoea pityocampa ENA	HE963113	HE956698	HE864328
Thaumetopoea wilkinsoni	HE963116	HE956701	HE864331
Anaphe panda panda *	HM892083	—	—
Anaphe panda infracta *	LT614827	—	—
Epanaphe carteri *	HM892014	—	—
Gazalina chrysolopha *	LT614826	—	—
Hypsoides antsianakana *	LT614829	—	—
Ochrogaster lunifer	AM946601	AM946601	AM946601
Paradrallia rhodesi *	LT614828	—	—

*, sequenced only the barcoding portion of *cox1*; —, not available; underline, sequences produced for the present work.

Tab.S3. Dataset content.

	morphology	*cox1*	*cox2*	*cox3*	*atp6*	*atp8*	*cob*	*nad1*	*nad4*	*nad5*
Thaumetopoea processionea processionea	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Thaumetopoea processionea pseudosolitaria	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea solitaria solitaria	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Thaumetopoea solitaria iranica	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea herculeana herculeana	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Thaumetopoea herculeana judaea	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea apologetica apologetica	🟢	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea apologetica abyssinica	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea dhofarensis	🟢	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea jordana	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea cheela	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea bonjeani	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Thaumetopoea pinivora	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Thaumetopoea libanotica	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Thaumetopoea ispartaensis	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Thaumetopoea sedirica	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea torosica	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea pityocampa pityocampa	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Thaumetopoea pityocampa orana	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Thaumetopoea pityocampa ENA	🔴	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Thaumetopoea wilkinsoni	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Anaphe etiennei	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Anaphe panda panda	🟢	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Anaphe panda infracta	🟢	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Anaphe venata	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Epanaphe carteri	🟢	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Epanaphe moloney	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Epanaphe nigricincta	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Epanaphe subsordida	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Gazalina apsara	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Gazalina chrysolopha	🟢	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Gazalina trasversa	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Hypsoides antsianakana	🟢	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Hypsoides placidus	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Ochrogaster lunifer	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Paradrallia punctigera	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Paradrallia rhodesi	🟢	🟢	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴

🟢 Complete data, 🟢 Barcoding sequence (about 650 bases), 🔴 Missing data.

Length of genes (bases). *cox1*: 1518 ca.; *cox2*: 679 ca.; partial *cox3*: 126 ca.; *atp6*: 165 ca.; *atp8*: 678 ca.; *cob*: 1062 ca.; *nad1*: 939 ca.; partial *nad4*: 490 ca.; *nad5*: 1509 ca.

Note. **Molecular markers coverage for the ingroup**. Ten of the 16 *Thaumetopoea* species possess the whole set of markers (8) with the total alignment spanning 6358 positions. Only the barcode-portion of *cox1* was sequenced for *Thaumetopoea apologetica* and *Thaumetopoea dhofarensis*. No molecular data available for *Thaumetopoea jordana*, *Thaumetopoea cheela*, *Thaumetopoea sedirica*, and *Thaumetopoea torosica*.