WILEY-BLACKWELL

Cladistics (2013) 1–28

Cladistics

10.1111/cla.12042

Anatomy of a cladistic analysis

Nico M. Franz*

School of Life Sciences, Arizona State University, PO Box 874501, Tempe, AZ, 85287-4501, USA

Accepted 13 May 2013

Abstract

The sequential stages culminating in the publication of a morphological cladistic analysis of weevils in the *Exophthalmus* genus complex (Coleoptera: Curculionidae: Entiminae) are reviewed, with an emphasis on how early-stage homology assessments were gradually evaluated and refined in light of intermittent phylogenetic insights. In all, 60 incremental versions of the evolving character matrix were congealed and analysed, starting with an assembly of 52 taxa and ten traditionally deployed diagnostic characters, and ending with 90 taxa and 143 characters that reflect significantly more narrow assessments of phylogenetic similarity and scope. Standard matrix properties and analytical tree statistics were traced throughout the analytical process, and series of incongruence length indifference tests were used to identify critical points of topology change among succeeding matrix versions. This kind of parsimony-contingent rescoping is generally representative of the inferential process of character individuation within individual and across multiple cladistic analyses. The expected long-term outcome is a maturing observational terminology in which precise inferences of homology are parsimony-contingent, and the notions of homology and parsimony are inextricably linked. This contingent view of cladistic character individuation is contrasted with current approaches to developing phenotype ontologies based on homology-neutral structural equivalence expressions. Recommendations are made to transparently embrace the parsimony-contingent nature of cladistic homology.

Introduction

This study dissects the sequential steps of a morphological cladistic analysis of Neotropical broad-nosed weevils in the *Exophthalmus* genus complex (Coleoptera: Curculionidae: Entiminae *sensu* Alonso-Zarazaga and Lyal, 1999). The character matrix, now published (Franz, 2012), was the product of an iterative process aimed at identifying phylogenetic characters for tribes, genera, and species groups within the complex. In particular, I had created a sequence of 60 character matrices, each representing an incremental advancement from the first stages of analysis towards the published product. The decision to coalesce the intermediate stages of analysis was deliberate; the intent was to create an empirical study that would accompany a previous publication (Franz, 2005a). Therein it was argued that cladists can utilize the congruence test in a successive approximation approach that yields more accurately *scoped* character circumscriptions[1] and thereby increases the odds of reliable phylogenetic inference under the parsimony criterion. As this explanation for parsimony's success presupposes a procedural evolution from coarse to refined assessments of homology, it appeared worthwhile to document this trajectory in a deliberate, step-wise manner.

A number of clarifications are in place. First, the core sections of this paper frequently take on a first-person perspective. This is considered appropriate because I developed this study's content while chronicling my own actions. Second, my approach to cladis-

*Corresponding author.
E-mail address: nico.franz@asu.edu

[1]The terms "scope" and "scoping" are used herein in the sense of Rieppel (2007a). A properly scoped character or homology is formulated in such a way that it refers precisely to the intended set of taxa —no more (including unintended taxa within or outside of the analysis), and no less (failing to refer to intended taxa). Rieppel (2007a) discusses the effects of scope expansion and restriction (i.e. shifts in taxon sampling schemes) on character codings.

tics is naturalistic, open to likelihood considerations, and classification/language-centred (Franz, 2005a,b). This view implies that successful practice in science can have primacy over theory (Putnam, 1974). Accordingly, I place more emphasis on pragmatic insights than on longstanding cladistic/philosophical themes. I need not claim that my philosophical positions are exactly right, or even necessary to engage in the practice of cladistic character refinement (as long as this practice is *potentially* justifiable). Third, I assume that my taxonomic specialization plays a significant role in how I approach cladistics. Weevils in the superfamily Curculionoidea constitute a particularly challenging group for systematists, as aptly summarized by Thompson (1992, pp. 835–836):

> "Classification of weevils is like a mirage in that their wonderful variety of form and the apparent distinctness of many major groups lead one to suppose that classifying them will be fairly straightforward but, when examined closely, the distinctions disappear in a welter of exceptions and transformation series. As a result, a number of major groups are currently defined by single characters. This has produced a workable system but the groups so formed are inevitably artificial and the true relationships of their components are obscured."

The Curculionidae alone have close to 50 000 known species placed in some 225 tribes (Bouchard et al., 2011), many of which are considered unnatural (Thompson, 1992; Oberprieler et al., 2007; Franz and Engel, 2010). Contemporary systematists who aim to improve the mid-level classification of weevils are confronted with an intractable morphological complexity and taxonomic legacy, including many poor diagnostic features and few well-defined phylogenetic characters. One typically feels like the first person to examine these groups in a phylogenetic context, as the majority of characters must be revised or newly researched. While this condition is not unique to weevils, I cannot assume that every systematist will share my experiences in their entirety. Many other taxa are more (or even less) thoroughly explored than weevils, or present different inferential challenges (e.g. bacteria). On the other hand, *some* experiences should seem familiar to anyone who has published a morphology-based cladogram and classification. My point is this—it is difficult to fully understand cladistic practice without also invoking the specific evolutionary and classificatory properties of the taxa under study, as well as their effects on the practitioner. Moreover, if one holds that inference methods ought to adjust to the perceived properties of taxa, then the methods themselves have limited applicability. These limitations are hereby transparently acknowledged.

The structure of this paper is as follows. First, a brief introduction is offered summarizing the systematic challenges of the genus *Exophthalmus* Schoenherr and relatives. Then the study's trajectory over the course of 60 cladistic matrices is chronicled, starting with 52 taxa and ten characters and ending with 90 taxa and 143 characters. Line/scatter plots are provided to track the evolution of common matrix parameters: number of most parsimonious trees (MPTs), length, consistency index, etc. Incongruence length difference (ILD) tests are used to assess the significance of topological changes across succeeding matrices. As expected, the analytical trajectory included multiple stages of taxon/character assembly, evaluation, recoding, and reanalysis (perhaps approximating what Hennig, 1966; referred to as "reciprocal illumination"; Rieppel, 2003). Each stage is given an informal name for ease of communication.

In the final discussion, the insights of this study are connected (1) to a broader discourse about the procedural interdependency of cladistic character individuation and evolutionary plausibility considerations, as well as (2) to the implications of cladistic character scoping for the creation of structure-referencing expressions in phenotype ontologies (cf. Vogt et al., 2010; Deans et al., 2012a; Mungall et al., 2012). The motivation for shifting the focus to the latter topic is straightforward. We have entered an era in science where information about observations and theories is not just stored in computers but "explained" to them. Increasingly, the specific insights and terms used in scientific domains are translated into *ontologies*, i.e. controlled and structured vocabularies (classes, relationships) that allow integrated knowledge representation and reasoning (van Harmelen et al., 2008; Lord and Stevens, 2010; Franz and Goldstein, 2013). Several groups of authors promoting this approach for systematics have adopted a mostly homology-neutral vocabulary to construct phenotype ontologies. I argue here that this preference for homology neutrality is not well aligned with cladistic practice and will prohibit adequate ontological representation of the most refined linguistic products that cladists can bring into comparative biology.

Taxonomic antecedents

The systematic challenges of the *Exophthalmus* genus complex are described in detail in Franz (2012). When the study was initiated, *Exophthalmus* (Fig. 1) contained 86 species, roughly half of which occur on the Caribbean islands and in Mesoamerica, respectively. Since its first textual description in 1826, the genus has remained poorly delimited. *Exophthalmus* has long been regarded as paraphyletic in relation to several genera including *Compsus* Schoenherr (104 species), *Diaprepes* Schoenherr (16 species), *Eustylus* Schoenherr (26 species), *Exorides* Pascoe (29 species), *Lachnopus* Schoenherr (66 species), and other less speciose genera.

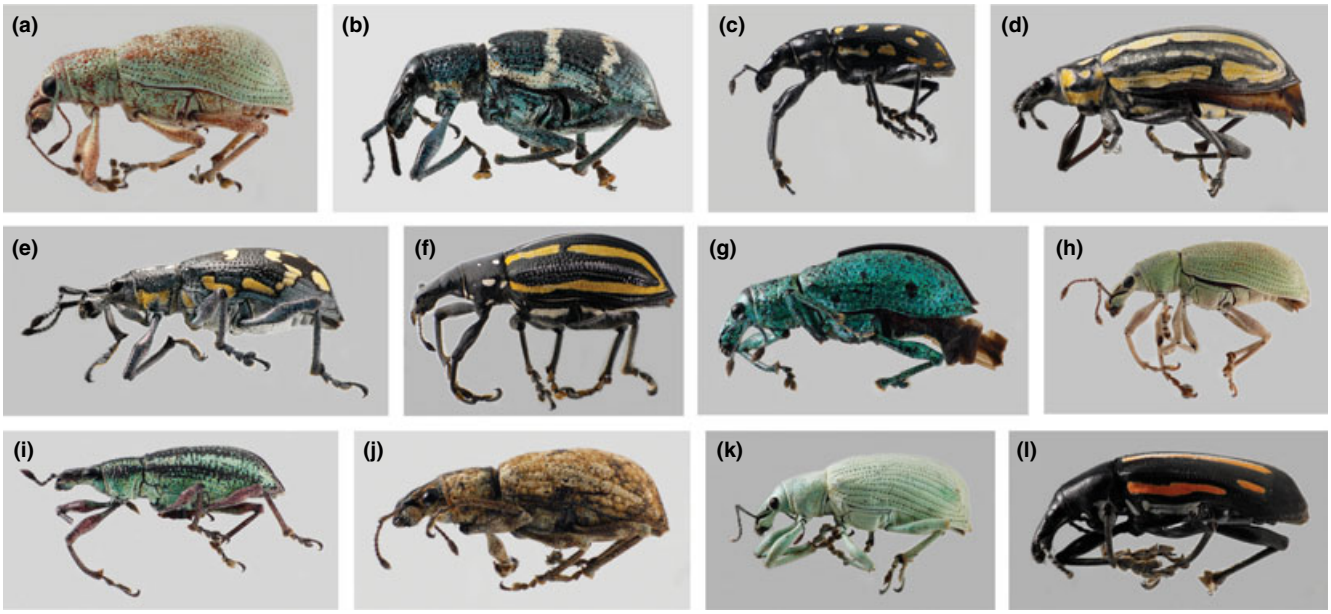

Fig. 1. Lateral habitus photographs of select species traditionally placed in the genus *Exophthalmus* (Fig. 9 for information on relative phylogenetic placement). (a) *E. agrestis* (Boheman); (b) *E. consobrinus* (Marshall); (c) *E. hieroglyphicus* Chevrolat; (d) *E. impressus* (Fabricius); (e) *E. nicaraguensis* Bovie; (f) *E. quadrivittatus* (Olivier); (g) *E. quinquedecimpunctatus* (Olivier); (h) *E. roseipes* (Chevrolat); (i) *E. sulcicrus* Champion; (j) *E. triangulifer* Champion; (k) *E. verecundus* (Chevrolat); (l) *E. vittatus* (Linnaeus).

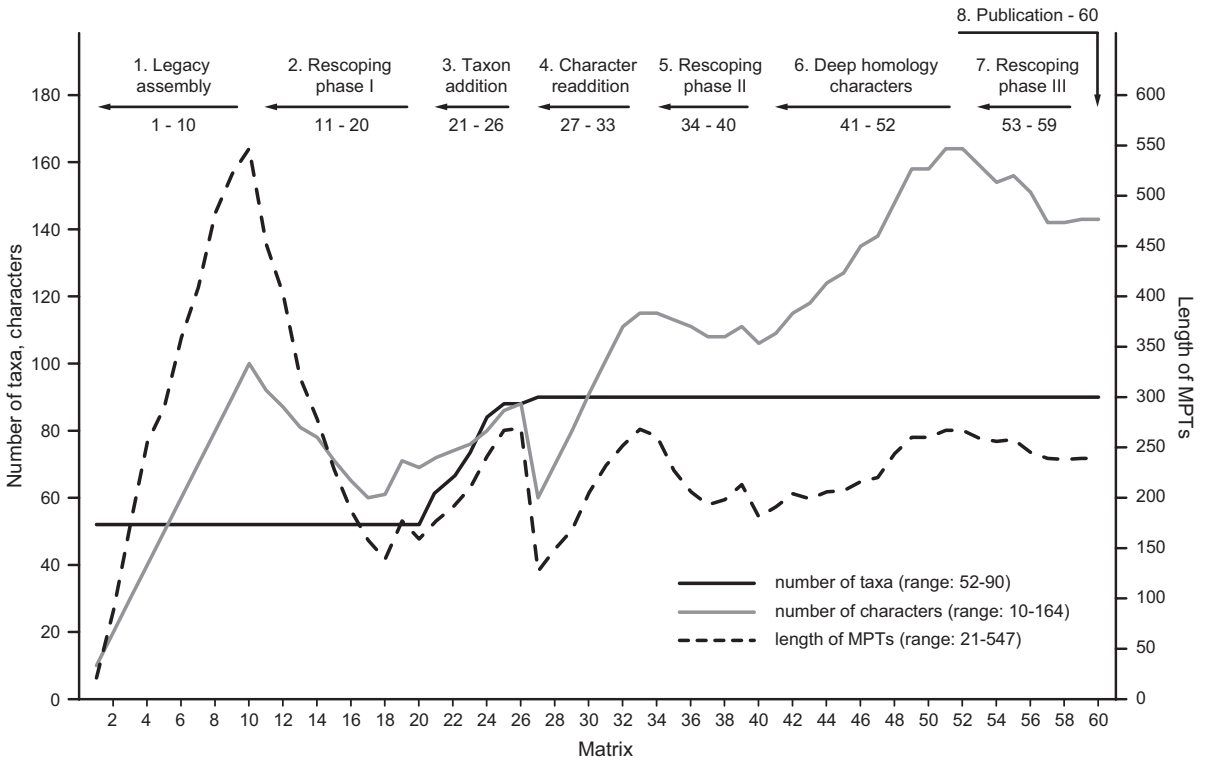

Fig. 2. Trajectory of the number of taxa, characters (scale on left axis), and length of the resulting most parsimonious trees (scale on right axis) throughout the 60 cladistic matrices and eight identified stages of analysis.

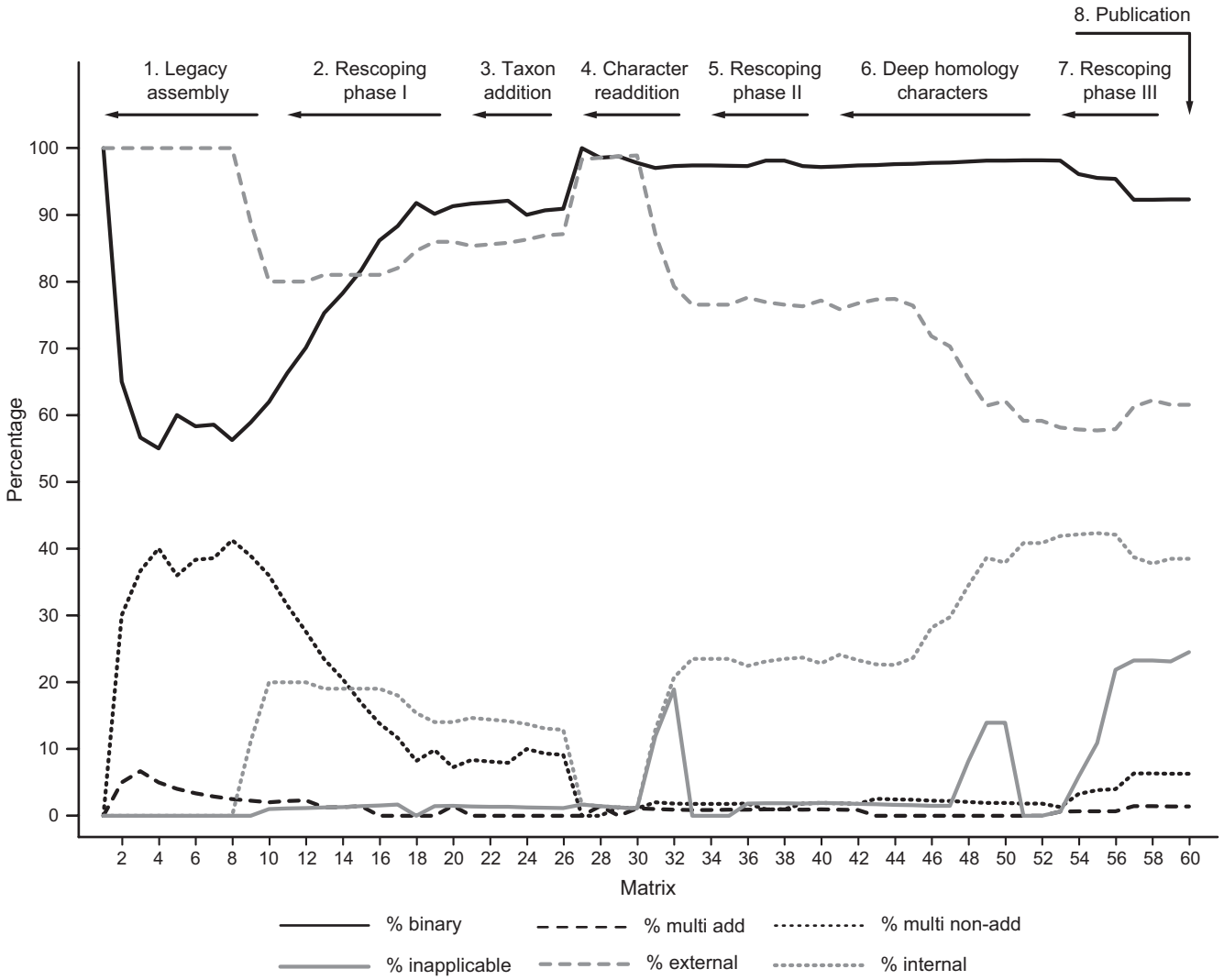

Fig. 3. Percentage contribution of binary, multistate additive, and multistate non-additive characters; characters with taxa coded as inapplicable (indicative of reductive coding); and characters corresponding to external versus internal (terminalia) structures throughout the stages of analysis.

Three regional revisions of *Exophthalmus* are available (Champion, 1911; Hustache, 1929; Vaurie, 1961). Neither these nor other treatments have reconciled taxonomic inconsistencies borne out by many decades of piecemeal additions of new species to an originally ill-conceived taxon. All publications on *Exophthalmus* prior to the new analysis were conducted using a pre-Hennigian approach.

The classificatory challenges extend to the tribal level, as members of the *Exophthalmus* genus complex are variously assigned to the tribes Eustylini Lacordaire (18 genera) and Geonemini Gistel (40 genera). The 55 tribal concepts in the weevil subfamily Entiminae *sensu* Alonso-Zarazaga and Lyal (1999) remain similar to Lacordaire's (1863) classification and are considered "chaotic" (Oberprieler et al., 2007).

I initiated this analysis with limited experience in the systematics of *Exophthalmus* (Franz, 2010a), although I had previously published 12 cladistic analyses (totalling approximately 300 taxa and 475 characters) on other weevil groups. Moreover, the study was pragmatically designed as an intermediate step ("phylogenetic reassessment") towards a full-scale revision of *Exophthalmus* still in progress. Thus, an initial set of 52 species, approximately 12 outgroup and 40 ingroup (Table 1), was assembled using established criteria (Nixon and Carpenter, 1993). Several pertinent works were surveyed to extract characters of phylogenetic significance, in particular Champion's (1911) monograph of Mesoamerican weevils and van Emden's (1944) key to the entimine genera of the world. Many of these traditional diagnostic features were used to build up the early matrix versions.

Characterization and comparison of analytical stages

Matrix generation stages

The generation of 60 cladistic matrices took place over a period of 2 years. An attempt was made to capture with each matrix version a more or less comparable amount of advancement. For instance, the first ten versions were each saved after adding ten new characters. Thereafter, until matrix 26, each additional modification was recorded in a spreadsheet. A new version was named after combinations of ten of these actions had been performed: rerooting of the tree, new taxon addition, new character addition, character recoding (including resolution of initial missing entry codings such as "?"), character repolarization, and character deactivation. The modifications corresponding to matrices 27–60 were no longer recorded in detail, as the process of numbering "steps" of advancement became increasingly arbitrary and cumbersome to record (WinClada has no tracking option for such edits). Towards the end (matrices 38–60), much time was dedicated to re-evaluating previously established characters ("ok", "needs revision", "consider eliminating", etc.). A pragmatic way to coalesce changes during these final stages was to coin a new matrix version at the end of a work session. Thus, no claim is made here that each matrix version represents an evenly measured advancement over its predecessor. On the other hand, it is safe to assume that the most salient advancements from the first to the last version were captured in at least one of the 60 matrices.

Matrix analyses and comparisons

The approach used to analyse adult morphological characters of weevils in the *Exophthalmus* genus complex is straightforward, and is further detailed in Franz (2012). All analyses employ parsimony under equal weights, although naturally the reformulation or elimination of characters affects their weight in an analysis. The character matrices were assembled with ASADO (Nixon, 2008). They were analysed by spawning NONA (Goloboff, 1999), and using the parsimony ratchet search algorithm (Nixon, 1999). This software and analysis solution is well suited for smaller sized

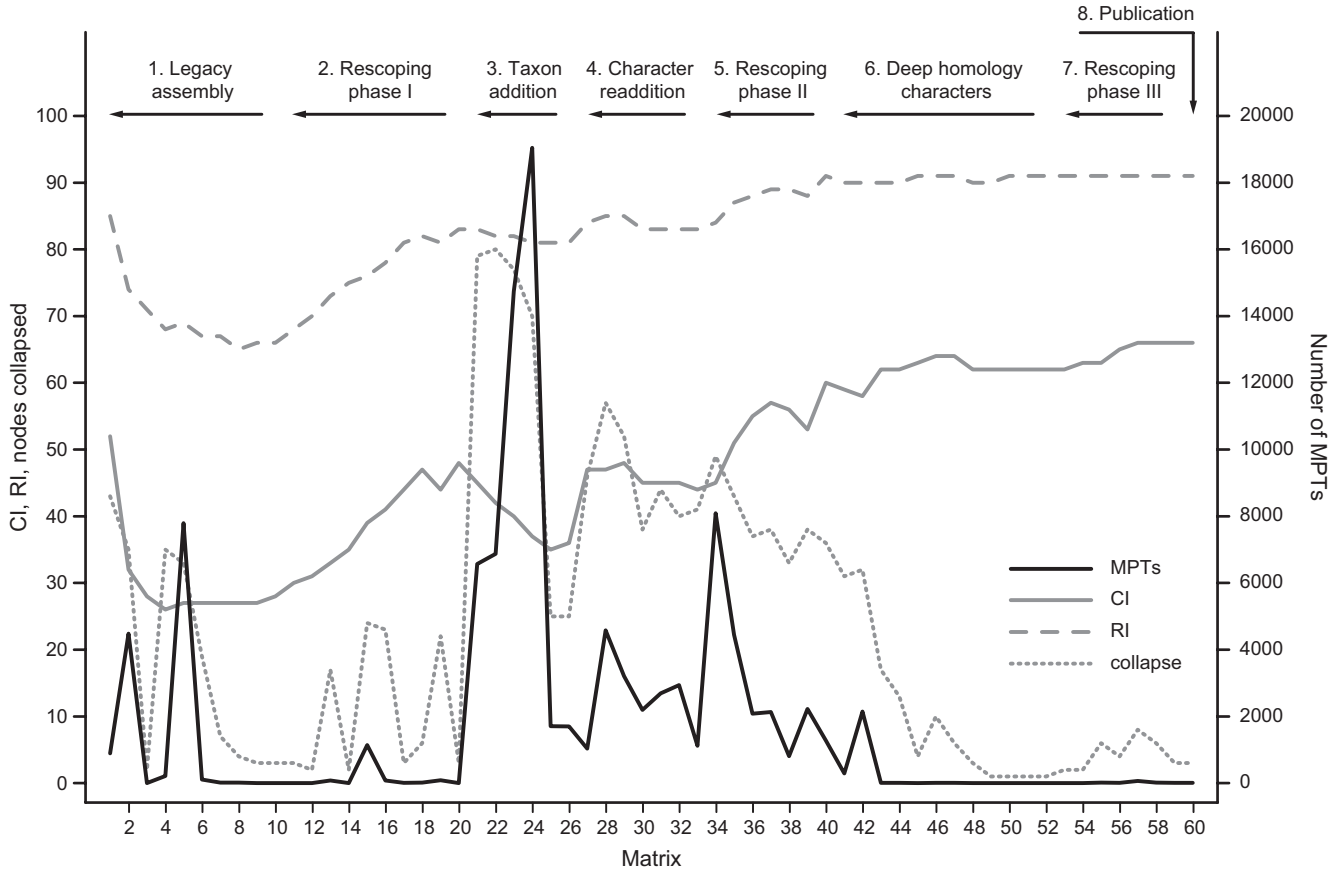

Fig. 4. Trajectory of common tree statistics, including consistency index (per cent), retention index (per cent), numbers of nodes collapsed in the strict consensus (all scale on left axis), and number of most parsimonious trees (scale on right axis) throughout the stages of analysis.

matrices. WinClada facilitates an instructive interplay between the WinDada (matrix view) and WinClados (tree view) interfaces where the effects of modifying a matrix are readily observable.

Each of the 60 matrices was analysed using the ratchet commands: 500 iterations per replication, 2 trees to hold per replication, 10% characters to sample, and 5 sequential ratchet runs. The results were resubmitted to NONA to find the best trees (hold 20000; amb=; poly-; max*; amb-: poly=; best;). The following information was documented after each analysis: number of taxa; number of characters (total, activated, deactivated); proportion (number) of characters coded as binary, multistate additive, and multistate non-additive; number of characters with inapplicables (Strong and Lipscomb, 1999); number of external versus internal (terminalia) characters (Song and Bucheli, 2010); number of MPTs, length (*L*), consistency index (CI), retention index (RI); and number of nodes collapsed in the strict consensus tree. The CI and RI values are reported as percentages (0–100), in keeping with the resulting publication (Franz, 2012). For select consensus trees (matrices 10, 20, 30 and 60; see Figs 6–9), Bremer branch support values (Bremer, 1994) were calculated in NONA with the commands: hold 25 000, suboptimal 15, and bsupport 15 (Franz, 2012).

Series of ILD tests (Farris et al., 1994) were run in WinClada to explore significant points of topology change (for discussions see Barker and Lutzoni, 2002; Hipp et al., 2004; Ramírez, 2006). The test conditions were set to 250 replications, 20 mult reps per replication, 20 trees to hold per replication, and 1000 trees to hold.

Sequence of analytical stages—from legacy character assembly to publication

Overview of analytical stages

In all, eight succeeding stages were identified in the analysis, as follows (Fig. 2): (1) legacy character assembly stage (matrices 1–10); (2) character rescoping phase I (matrices 11–20); (3) taxon addition phase (matrices 21–26); (4) character readdition phase (matrices 27–33); (5) character rescoping phase II (matrices 34–40); (6) deep homology character addition phase (matrices 41–52); (7) character rescoping phase III

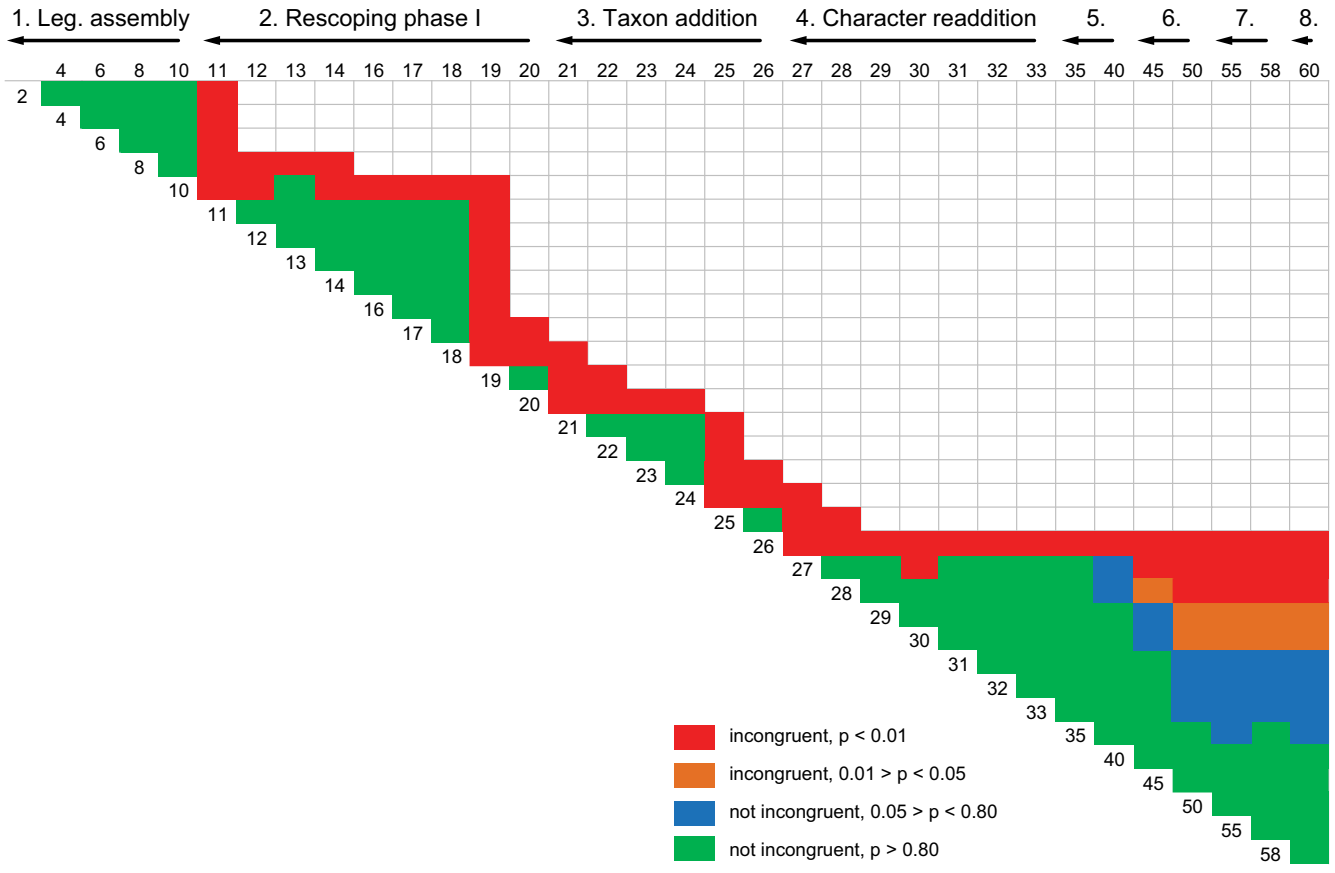

Fig. 5. Colour chart ("heat map") showing the results of reciprocal ILD tests among select pairs among the 60 matrices produced in the course of the analysis. Orange (and red) squares are indicative of significant changes in phylogenetic topology.

(matrices 53–59); and (8) publication stage (matrix 60). The arrangement of figures reflects the sequence of stages. Particular sets of figures display different aspects of each stage: i.e. Figs 2–4—core matrix properties and tree statistics; Fig. 5—ILD tests among select paired matrices; Figs 6–9—select examples of tree topologies; Figs 10 and 11—examples of character state optimizations; and Fig. 13—examples of publication stage characters. Tables 2 and 3 show examples of early stage (matrices 10–16) and publication stage (matrix 60) character circumscriptions, indicating their respective properties and actions taken towards their phylogenetic rescoping.

1—Legacy character assembly stage (matrices 1–10)

The initial series of matrices (1–10) were built up in a deliberately mechanical and unfiltered way, applying character circumscriptions from existing taxonomic keys and revisions such as Lacordaire (1863), LeConte and Horn (1876), Champion (1911), Pierce (1913), Hustache (1929), van Emden (1944), Kuschel (1955), Thompson (1992), Anderson (2002), and Franz (2010a). A new matrix version was coined once 10 such characters had been identified and coded for all 52 taxa. A large proportion of these characters, at times exceeding 40% of the total number, were multistate (Fig. 3). Several examples are provided in Table 2, for example the length relation of the funicular antennomeres I and II (van Emden, 1944), or the outline of the eye in lateral view (Anderson, 2002). Characters relating to the terminalia were first defined in matrices 9 and 10.

Admittedly, the initial characters and codings fell short of the (or specifically my) highest standards of homology (cf. Patterson and Johnson, 1997; Rieppel and Kearney, 2002; Wägele, 2005). Although these codings are observable and repeatable, they resulted from a quick and linear pass through the 52 taxa, with little consideration for detail homology or phylogenetic relevance. Even so, the cladistic outcome was surprisingly unwieldy, given that most characters had a traditional diagnostic value. In particular, the average character length rose steeply from 2.10 in matrix 1 to 6.38 in matrix 4, and lowered only slightly to 5.47 at the end of the assembly stage (matrix 10; Fig. 2). The resulting consensus topology of three MPTs was reasonably well resolved (only three nodes collapsed; Figs 4 and 6); however, unreversed synapomorphies were rare. Instead, the majority of nodes were supported by long series of homoplasious characters (Fig. 10). Several of these presumably informative characters had lengths of 14 steps and more (Table 2; Fig. 10).

The example of the length relation of the funicular antennomeres I and II illustrates a general trend for this stage. This character was numbered as 31 in matrix 10, and had been highlighted in O'Brien and Kovarik (2001) as one of the diagnostic traits separating *Diaprepes* from *Exophthalmus* and other genera. I approached the coding process with the expectation that a distinctly longer funicular antennomere II would reliably identify *Diaprepes* species. This expectation was altered in several ways by the experience of coding the character for the 52 taxa on hand. First, it became necessary to make a somewhat forced distinction between "II slightly longer than I" and "II conspicuously longer than I". Only the latter condition would apply to the six included species of *Diaprepes* plus one species of *Exophthalmus*, which was subsequently transferred (Franz, 2012). Second, the forced distinction also meant that an unrelated *Naupactus–Pantomorus* clade now shared the presumed apomorphic condition for *Diaprepes*. Third, coding the character for related genera further demonstrated its inadequacy: genera such as *Compsus*, *Exophthalmus*, and *Lachnopus* each seemed to have multiple independent origins of a particular state (Fig. 10). The result was an overall length of 17 steps occurring in clades that ranged from very inclusive to monotypic. The combination of poorly delimited states and poor phylogenetic performance led me to reassess the utility of the original for the present scope.

At the end of the legacy assembly stage, the cumulative consistency and retention indices were at or near their lowest values in the study (Fig. 4). Clade support was generally weak (Figs 6 and 10). While these outcomes may be acceptable in other contexts (Mindell and Thacker, 1996; Wenzel and Siddall, 1999; Song and Bucheli, 2010), they indicated to me that a great number of legacy characters were of questionable utility for identifying natural lineages in the *Exophthalmus* genus complex. I would not have attempted publishing this matrix at such an unrefined stage.

2—Character rescoping phase I (matrices 11–20)

A first reassessment phase began with a focus on the detail homology and phylogenetic informativeness of the coded characters.[2] This process led to the temporary deactivation of as many as 40 characters while generating matrices 11–18 (Fig. 2). Character length played an important yet not exclusive role in my decisions to deactivate characters (Table 2). Typically I retained characters that showed considerable levels of homoplasy but at the same time had (1) clearly circumscribed and assignable states and (2) identified at least one plausible multi-taxon clade. A suitable example is the presence of elytral humeri (Champion, 1911),

[2]I worked extensively with the WinClados interface and the Diagnoser Toggle function to review alternative tree topologies and character state optimizations during this and subsequent stages.

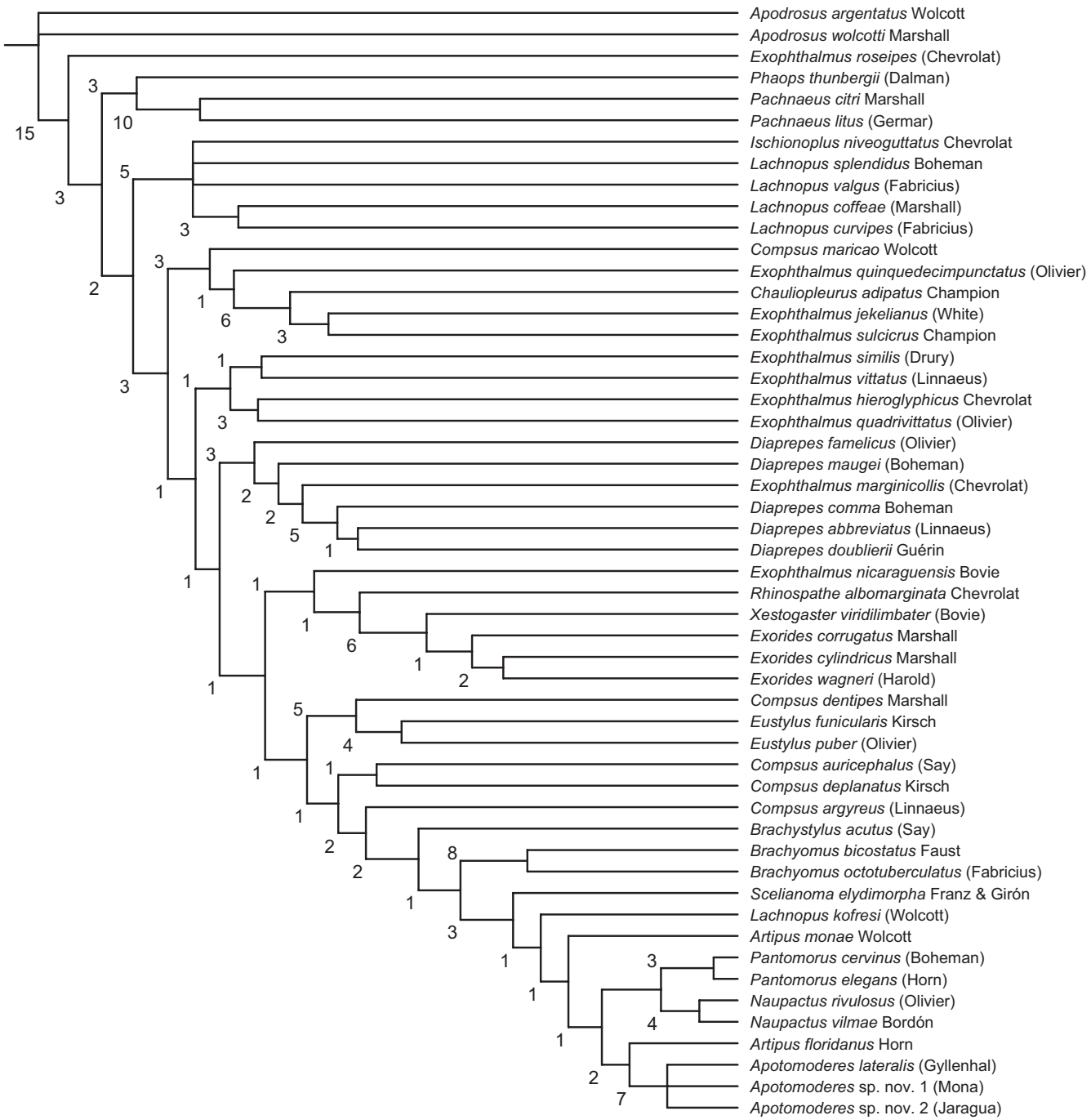

Fig. 6. Strict consensus of the four most parsimonious trees resulting from matrix 10, corresponding to the end of the legacy character assembly stage (stage 1; Fig. 2). Relevant analysis data: 52 taxa, 100 characters (61.0% binary), $L = 547$ steps, CI = 28, RI = 66, and three nodes collapsed in the strict consensus. Bremer branch support values are listed near the root of the corresponding clade. Average support per resolved node (adjusted for 45 nodes/52 taxa) = 2.62.

a character whose length *increased* from five steps (matrix 10, character 10; Fig. 10) to seven steps in the published analysis (matrix 60, character 62; Fig. 11). The character remained activated at all times.

The reassessment of poorly performing characters had very significant effects on the matrix properties (Figs 2–4). This process also produced the first significant difference in tree topology (Fig. 5). At first only characters found to be highly reliable were retained. Progressing accordingly from version 11 to 18, an average of 51 steps was saved with each new matrix. The relative contribution of binary characters

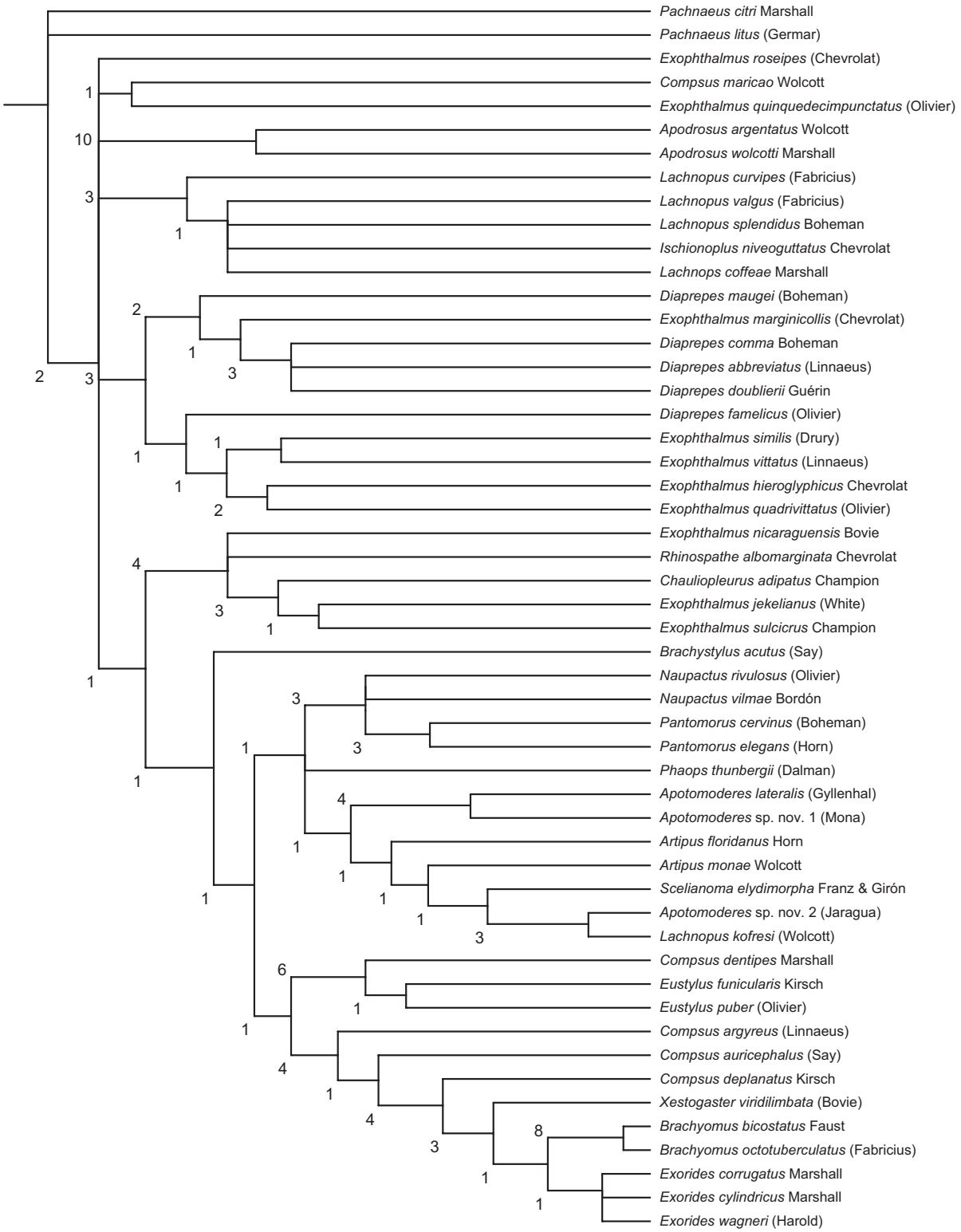

Fig. 7. Strict consensus of the three most parsimonious trees resulting from matrix 20, corresponding to the end of the character rescoping phase I (stage 2; Fig. 2). Relevant analysis data: 52 taxa, 69 characters (91.7% binary), L = 159 steps, CI = 48, RI = 83, and three nodes collapsed in the strict consensus. Bremer branch support values are listed near the root of the corresponding clade. Average support per resolved node (adjusted for 38 nodes/52 taxa) = 1.73.

Table 1
List of genera, species, outgroup/ingroup placement, and corresponding changes made in matrix 1 versus matrix 26 of the analysis

Number	Genus	Out/In matrix 1	Species (52 total)	Out/In matrix 26	Species (90 total)	Changes made—matrix 1 ⇒ matrix 26
1	*Achrastenus* Horn, 1876	In	0	Out	1*	In ⇒ Out; +1 species
2	*Apodrosus* Marshall, 1922	Out	2	Out	2	Unchanged
3	*Apotomoderes* Dejean, 1834	Out	3	Out	2	−1 species
4	*Artipus* Sahlberg, 1823	Out	2	Out	3	+1 species
5	*Brachyomus* Lacordaire, 1863	In	2	In	2	Unchanged
6	*Brachystylus* Schoenherr, 1845	In	1	Out	1	In ⇒ Out
7	*Chauliopleurus* Champion, 1911	Out	1	In	1	Out ⇒ In
8	*Cleistolophus* Sharp, 1891	Out	0	Out	1*	+1 species
9	*Compsus* Schoenherr, 1823	In	5	In	9	+4 species
10	*Diaprepes* Schoenherr, 1823	In	5	In	7	+2 species
11	*Epicaerus* Schoenherr, 1834	Out	0	Out	2*	+2 species
12	*Eustylus* Schoenherr, 1843	In	2	In	4	+2 species
13	*Exophthalmus* Schoenherr, 1823	In	10	In	24	+14 species
14	*Exorides* Pascoe, 1881	In	3	In	4	+1 species
15	*Ischionoplus* Chevrolat, 1878	Out	1	Out	1	Unchanged
16	*Lachnopus* Schoenherr, 1840	Out	5	Out	8	+3 species
17	*Melathra* Franz, 2011	Out	0	Out	1*	+1 species
18	*Naupactus* Dejean, 1821	Out	2	Out	2	Unchanged
19	*Otiorhynchus* Germar, 1822	Out	0	Out	2*	+2 species
20	*Pachnaeus* Schoenherr, 1826	Out	2	In	2	Out ⇒ In
21	*Pandeleteius* Schoenherr, 1834	Out	0	Out	1*	+1 species
22	*Pantomorus* Schoenherr, 1840	Out	2	Out	2	Unchanged
23	*Paululus* Howden, 1970	Out	0	Out	1*	+1 species
24	*Phaops* Sahlberg, 1823	In	1	In	1	Unchanged
25	*Phaopsis* Kuschel, 1955	In	0	In	1*	+1 species
26	*Rhinospathe* Chevrolat, 1878	Out	1	In	1	Out ⇒ In
27	*Scelianoma* Franz & Girón, 2009	In	1	Out	1	In ⇒ Out
28	*Tetrabothynus* Labram & Imhoff, 1852	Out	0	In	1*	Out ⇒ In
29	*Tropirhinus* Schoenherr, 1823	Out	0	In	1*	Out ⇒ In
30	*Xestogaster* Marshall, 1922	In	1	In	1	Unchanged

*Newly added genus.

increased from 62.0 to 91.8%, at the cost of eliminating or recoding many multi-state characters that originated from taxonomic keys (Table 2). For instance, the presence of irregular concavities or tumescences on the pronotum, initially encoded in a single character and applied to all taxa (matrix 10, character 45; Fig. 10), was subsequently split into four separate characters that narrowly identify apomorphic conditions of distinct clades (matrix 60, characters 42–45; Fig. 11). This type of *scope restriction* (Rieppel, 2007a) almost invariably produced more reductive and phylogenetically precise coding schemes. As a result, the average character length decreased from 4.91 (matrix 11) to 2.30 (matrix 20). The consistency and retention indices increased correspondingly (matrix 11: CI = 30 and RI = 68; matrix 20: CI = 48 and RI = 83; Fig. 4).

Towards the end the rescoping phase I (matrices 18–20), I started adding new (= non-legacy) characters derived both from parallel studies (Franz, 2010b; Girón and Franz, 2010, 2012) and from observations made during the first pass. At this point it became evident that the matrix was topologically unstable and clade support was generally weak (Fig. 7). For instance, the number of collapsed nodes progressed in a sequence of 23, 3, 6, 22, and 3 from matrices 16 to 20 (Fig. 4). Furthermore, the transition of matrices 18 and 19 produced the second significant change in tree topology (Fig. 5). This topology shift was facilitated in part by a decrease in the character/taxon ratio from 1.92 (matrix 10) to 1.17 (matrix 18), thus giving the new characters a high relative weight in shaping the outcome.

The apparent lack of robustness was in itself a cause for concern (Giribet, 2003). More importantly, I suspected that my codings of informative new character systems, such as surface sculpture traits of the head and terminalia, were compromised by insufficient taxon sampling. Inference problems related to sampling gaps were especially apparent in *Exophthalmus*, where many species-level configurations of, for example, the dorsal and ventral surface of the rostrum, the endophallic sclerites of the aedeagus, and the spermatheca appeared to represent "unique" conditions. I struggled to identify shared multi-species states for these character systems. Without sufficient taxon coverage, it would be impossible to propose precise detail homologies among taxa and thus infer informative character state transitions (Mickevich and Weller, 1990). I therefore opted to increase the taxon sampling from 52 species to

Table 2
Examples of traditional characters and states coded in the legacy assembly phase (matrices 1–10), their respective sources, and actions taken as part of the rescoping phase I (matrices 11–20)

Matrix/ Character	Description of character and states	Character source	Action taken	Motivation for action and subsequent trajectory
10/27	Rostrum, curvature of scrobe: (0) straight; (1) slightly curved; (2) strongly curved; and (3) angulate	van Emden (1944)	Deactivated ⇒ recoded	L = 14 steps, CI = 21, RI = 57; 4 ⇒ initially reduced to 3 states, then rescoped to emphasize scrobe orientation
10/53	Mandible, number of large setae: (0) 2; (1) 3–4; (2) 5–8; and (3) 9–15	Anderson (2002)	Deactivated ⇒ eliminated	L = 16 steps, CI = 18, RI = 38 ⇒ too variable, states poorly circumscribed
10/62	Maxilla, number of lacinial teeth: (0) 1; (1) 2–3; (2) 4–5; and (3) 6–8	Franz (2006)	Deactivated ⇒ eliminated	L = 11, CI = 18, RI = 62 ⇒ too variable, states poorly circumscribed
11/21	Rostrum, posterior extension of median sulcus: (0) anterior margin of eyes; (1) midpoint of eyes; (2) posterior margin of eyes; and (3) posterior region of head	Anderson (2002)	Deactivated ⇒ recoded	L = 13 steps, CI = 23, RI = 66 ⇒ "compounded character", states poorly circumscribed
11/38	Antenna, relation of funicular antennomeres I & II: (0) II slightly shorter than I; (1) II & I similar in length; (2) II slightly longer than I; and (3) II twice as long as I	van Emden (1944)	Deactivated ⇒ eliminated	L = 17 steps, CI = 17, RI = 53 ⇒ too variable, states poorly circumscribed
11/42	Head, outline of eye in lateral view: (0) elliptical, horizontally longer; (1) circular; (2) semi-circular, anterior margin subrectate; (3) vertically pear-shaped, ventrally tapered; (4) semi-circular, posterior margin subrectate; (5) semi-circular, anterior & posterior margins subrectate; and (6) elliptical, vertically longer	Anderson (2002)	Deactivated ⇒ eliminated	L = 14 steps, CI = 42, RI = 72 ⇒ too variable, states poorly circumscribed
14/16	Rostrum, elevation of epistoma: (0) equate with adjacent regions; (1) slightly convex; (2) slightly concave; and (3) strongly concave	Franz (2010a)	Recoded: 4 ⇒ 2 states	Matrix 13 (original): L = 11 steps, CI = 27, RI = 73 Matrix 14 (recoded): L = 5 steps, CI = 20, RI = 77 ⇒ Recoded to two states in light of topology: (0) slightly to strongly concave; and (1) equate to slightly convex
15/34	Pronotum, dorsal sculpture: (0) punctate; (1) with small, irregular concavities; (2) with large, irregular concavities; and (3) with irregular tumescences	LeConte and Horn (1876)	Recoded: 4 ⇒ 3 states, multiple chars	Matrix 15 (original): L = 4 steps, CI = 75, RI = 90 Matrix 16 (recoded): L = 5 steps, CI = 40, RI = 81 ⇒ Recoded into 5 characters (42–45); and adjustment of states
15/39	Metendosternite, position of anterior tendons: (0) mesal; (1) mesolateral; and (2) lateral	Franz (2010a)	Deactivated ⇒ eliminated	L = 8 steps, CI = 25, RI = 70 ⇒ too variable, states poorly circumscribed
16/57	Male terminalia, configuration of central endophallic tube: (0) posterior & anterior sclerites subequal in length; (1) anterior sclerite significantly longer; (2) anterior sclerite shorter; and (3) sclerites small, widely separated	Franz (2010a)	Recoded: 4 ⇒ 2 states, multiple chars	L = 5 steps, CI = 06, RI = 71 ⇒ Concept of "endophallic sclerite" significantly narrowed and contextualized; many species recoded

90 species, including ten additional genera and 14 additional species of *Exophthalmus* (Fig. 2; Table 1). My expectation was that this increase in taxon sampling would provide critical information needed to understand shared apomorphic conditions and character state transformations in the ingroup taxa.

3—Taxon addition phase (matrices 21–26)

The incorporation of 38 new taxa spanned across matrices 21–26 (Fig. 2). The process is magnified here because the addition of a single new taxon, plus the coding of all existing characters and states for that taxon, were jointly counted as a token step in a sequence of ten steps leading to an increment of the matrix version. In addition, some 20 characters were added in this stage. Typically these were characters whose phylogenetic significance was apparent at the moment of adding them. For example, *E. nicaraguensis* Bovie and *E. consobrinus* (Marshall) share a unique pattern of obliquely orientated stripes of scales on the elytra (Figs 1 and 11; matrix 60, character 78). This

character was added and coded for all taxa in matrix 23, at the time of including *E. consobrinus*. The majority of characters created in the taxon addition phase made contributions *only* to resolving apical clades. Complex multi-state or terminalia-related characters were not yet coded at this juncture (Fig. 3).

The gradual inclusion of the 38 new taxa provoked the third and fourth significant topology changes at the transitions of matrices 20–21 and 24–25, respectively (Fig. 5). There was also a decrease in the overall consistency index, from 48 (matrix 21) to 36 (matrix 26; Fig. 4). The likely reason for this decrease lay in my failure (at this specific stage) to rescope characters coded up until matrix 20 in such a way that no "erroneous homoplasy" would be introduced when applying them to the new taxa (Farris, 1983; Rieppel, 1988, 2007a; Wenzel and Siddall, 1999; Nixon and Carpenter, 2012).

Towards the middle of the taxon addition stage there was a very significant increase in the number of MPTs (matrix 20: three MPTs; matrix 24: 19 045 MPTs), and consequently in the nodes collapsed in the strict consensus (matrix 21: three nodes collapsed; matrix 23: 80 nodes collapsed). This effect is exaggerated (Fig. 4) because for practical reasons I had delayed coding certain sets of characters for the newly added taxa, thereby creating a group of "floating taxa" with many missing characters (Adams, 1972; Maddison, 1993; Kearney, 2002; Pol and Escapa, 2009). The situation normalized towards matrix 26, which yielded 1699 MPTs and 25 nodes collapsed in the strict consensus.

4—Character readdition phase (matrices 27–33)

At the end of the taxon addition phase (matrix 26), I had (1) obtained a character/taxon ratio of 1.00 (88/88), (2) dealt with all phylogenetically inadequate legacy characters through the process of rescoping, and (3) closed some of the apparent sampling gaps in the matrix. Nevertheless, phylogenetic resolution remained poor and clade support weak, with an average adjusted Bremer branch value of 1.18 per node (Fig. 8). The main reasons for this were threefold. First, as many as 44 characters left over from phases 1 and 2 had remained deactivated and thus required reexamination and possible reactivation or elimination. Second, the matrix now consisted of two separate generations of characters: one that was developed during the first rescoping phase (phase 2, 68 characters), and another that was supplemented during the taxon addition phase (phase 3, 20 characters). Although each generation of characters had been coded for all taxa, neither set had originated with all taxa in mind. Instead, they were coded as "local optima" connected to one or the other set of taxa. Thus, in an analogy to arguments for global parsimony (Farris, 1982; Rieppel, 2007a), it remained critical to reassess and reformulate the collective set of characters once more and adjust the scope of each description to the full set of taxa sampled. Third, I still had to incorporate new characters from underexamined regions such as the head and terminalia. The three shortcomings were addressed in phases 4, 5, and 6, respectively.

The transition from matrix 26 to matrix 27 was abrupt, resulting in the fifth and last significant change in tree topology (Fig. 5). I intermittently deactivated 28 characters that I considered somehow unsatisfactory, with an emphasis on multi-state non-additive characters (8 characters) and internal traits (16 characters; Fig. 3). The average character length decreased from 3.06 to 2.12. I then gradually re-/integrated 55 characters over the course of six matrix versions (Fig. 2). Only three of these were coded as multi-state, and 27 referred to properties of the terminalia—the highest proportion yet with 23.5% (Fig. 3). In retrospect, the transition to matrix 31 produced the earliest topologies that are not significantly incongruent with the published results (Figs 5 and 8).

This phase also marked the first deliberate attempt to utilize reductive coding with inapplicables (Strong and Lipscomb, 1999). My motivation for this was as follows. Character systems such as the endophallic sclerites of the male terminalia were observed to be highly variable in presence, number, size, shape, and topographic arrangement (cf. Andersson, 1994). When examining the intermediate trees and optimizations, I noticed the reoccurrence of anatomically similar structures (e.g. a central elongate sclerite) in multiple phylogenetically distant subclades (Fig. 13e and f and Table 3; characters 110, 115, and 116). These structures were present in more than one species per subclade, and furthermore showed specific transformations *within* a confined lineage. I decided to abandon the approach of coding the "same" states across all taxa for these characters. Instead, I utilized reductive coding as a simple (and arguably suboptimal) solution to create multiple characters whose phylogenetic context is "dynamic" and explicitly constrained (for discussion see Maddison, 1993; Strong and Lipscomb, 1999; Wheeler, 2001; Ramírez, 2007; Rieppel, 2007a).

In spite of the aforementioned efforts, none of the matrices generated in the character readdition phase I yielded fewer than 1000 MPTs or fewer than 40 nodes collapsed in the strict consensus (Figs 4 and 8). The lack of phylogenetic resolution probably contributed to the fact that the topologies of late stage 4 matrices are not significantly different from the final results (Fig. 5). There were no remarkable changes in the consistency index (range: 44–48) or retention index (range: 83–85). The relationships among most of the high-level focal lineages remained obscure (Fig. 8).

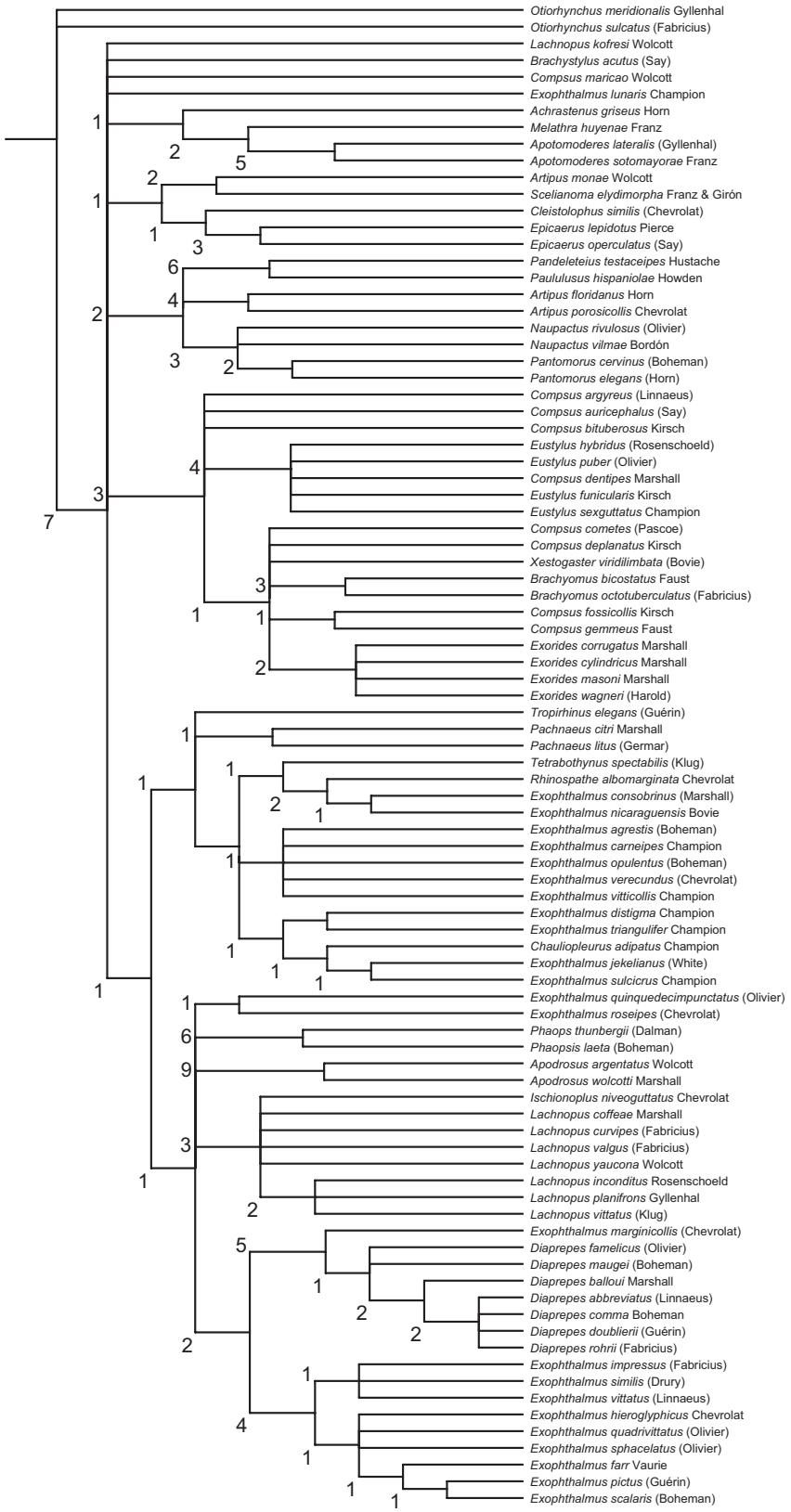

Fig. 8. Strict consensus of the 2192 most parsimonious trees resulting from matrix 30, corresponding to the middle of the character addition phase (stage 4; Fig. 2). Relevant analysis data: 90 taxa, 91 characters (97.8% binary), L = 205 steps, CI = 45, RI = 83, and 38 nodes collapsed in the strict consensus. All other matrices in this stage have even higher numbers of collapsed nodes in the strict consensus (Fig. 4). Bremer branch support values are listed near the root of the corresponding clade. Average support per resolved node (adjusted for 45 nodes/90 taxa) = 1.18.

5—Character rescoping phase II (matrices 34–40)

Another full pass through the characters was performed during this phase to reassess their phylogenetic utility and extensional scope. Most adjustments were minor and dealt with exploring combinations of variant codings for characters and states that appeared to show erroneous homoplasy. The concomitant changes in the recorded statistics—i.e. the number of characters, tree length, consistency, and retention indices, number of MPTs, nodes collapsed in the strict consensus, and tree topology—were largely insignificant (Figs 2–5).

The inability to obtain a more conclusive result at this stage gradually led me to explore additional character systems suited to resolve the many polytomies obtained in the strict consensus (see stage 6). I utilized two strategies to gain insight into the sources of ambiguity. These were exploratory in nature (cf. Grant and Kluge, 2003), intended to yield more accurate homology assessments, and not particularly successful. First, I replaced all states coded as inapplicable ("–") for 21 characters with zeros ("0"). While this was not an acceptable approach to encoding homology, it was useful for understanding—in a mechanistic sense—whether the reductively coded characters of stage 4 were (in part) responsible for the lack of resolution (Maddison, 1993; Wilkinson, 1995; Kearney, 2002; Wiens, 2003). They were not. Nevertheless, I did not reinstate the inapplicables until matrix 48. Second, I started examining character state optimizations on interim majority rule consensus trees. Naturally these were well resolved even if support clade support was weak. My motivation was to see if certain characters contributed more or less to a majority rule topology, which would in turn give more weight to certain ways of representing them in the matrix. Again, my epistemology was questionable given the nature of majority rule consensus trees (cf. Nixon and Carpenter, 1996; Sharkey and Leathers, 2001). The approach yielded few new insights.

6—Deep homology character addition phase (matrices 41–52)[3]

Two seemingly overdue adjustments occurred at this stage. On one hand, I re-examined key character systems—primarily on the head (rostrum) and in the male and female terminalia—where I thought that phylogenetic information was available with more accurate study and coding. On the other hand, I took an even more aggressive approach to scoping characters and states narrowly to match the set of taxa on hand, as opposed to adhering to more traditional and globally applicable circumscriptions. Many authors have discussed the ins and outs of such an approach (Richards, 2003; Jenner, 2004; Mishler, 2005; Ramírez, 2007; Winther, 2009). According to Rieppel (2007a, pp. 305–306):

> […] if character statements come with a variable scope, then the identification of the valid scope for character statements cannot be a matter of mere ostension, or rigid designation, but must be a matter of scientific theory construction. A character statement initially introduced with a restricted scope may certainly be amenable to scope expansion, but if so and how requires substantial knowledge to be brought to bear on the issue. Scope expansion of character statements can result in a situation where purportedly similar structures, apparently denoted by the same name (proper name or kind name), are in fact not the same. The nonhomology of such characters may be revealed through morphological complexity at the comparative level, by tree topology at the analytical level, or both. The logical consequence is the subdivision of the original character such that different homologues are denoted by separate anatomical terms. […] the decision to render a character statement incomplete in the context of an expanded domain of discourse, or else to expand the scope of a character statement while allowing for ambiguity of reference, is one that ultimately rests with the investigator […] Disambiguating reference of anatomical terms by scrutinizing morphological detail at the comparative level would appear to result in stronger hypotheses of primary homology, but at a more restricted scope.

I make no claim here that this approach is new or different from methods promoted, for example, in systematics textbooks. Nevertheless, the analysis was now moving from "structural kind characters" accumulated in the legacy stage to a full-blown, scope-contingent approach to character state individuation (cf. Mahner and Bunge, 1997; Rieppel and Kearney, 2002, 2007). It took time, empirically gained insight, and resolution of inner conflicts to reach this stage. The latter perhaps also because as a post-evolutionary taxonomy, total evidence systematist (cf. Mayr and Bock, 2002; Williams and Ebach, 2008; Rieppel, 2009), one might be disinclined to pursue or defend this approach (i.e. it may have a "just so" aspect). In my personal experience, implementing Rieppel's (2007a) approach can lead to counter-intuitive coding schemes, especially if one were to argue that character state propositions are merely conjectures that require a limited amount of theorizing (for discussion see de Pinna, 1991; Brady, 1994; Vergara-Silva, 2009).

An example of coding the presence/absence of longitudinal keel-shaped elevations—typically called "carinae"—on the dorsal surface of the rostrum illustrates the above (Fig. 12). Rostral carinae of *some* type are widespread throughout the weevils (Anderson, 2002). Coding their presence or absence in the widest sense

[3]"Deep" in the present context means non-superficial, or not readily recognized, requiring more in-depth analysis. This usage is more colloquial than, and differs from, the genetic/regulatory context in which the term is also used (Scotland, 2010).

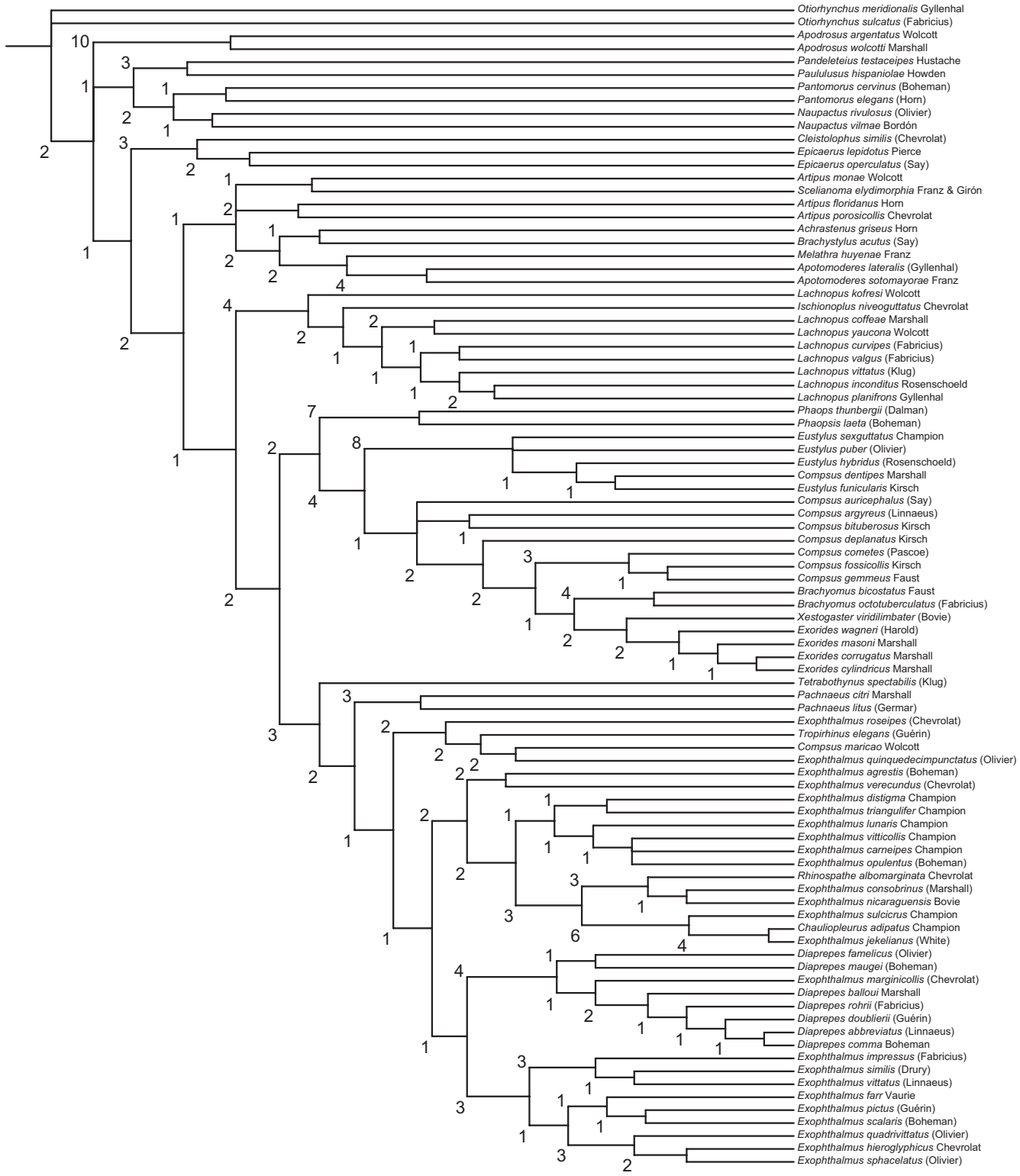

Fig. 9. Strict consensus of the eight parsimonious trees resulting from matrix 60, corresponding to the end stage of the analysis used for publication (stage 8; Fig. 2). Relevant analysis data: 90 taxa, 143 characters (92.3% binary), L = 239 steps, CI = 66, RI = 91, and three nodes collapsed in the strict consensus. Bremer branch support values are listed near the root of the corresponding clade. Average support per resolved node (adjusted for 81 nodes/90 taxa) = 1.89.

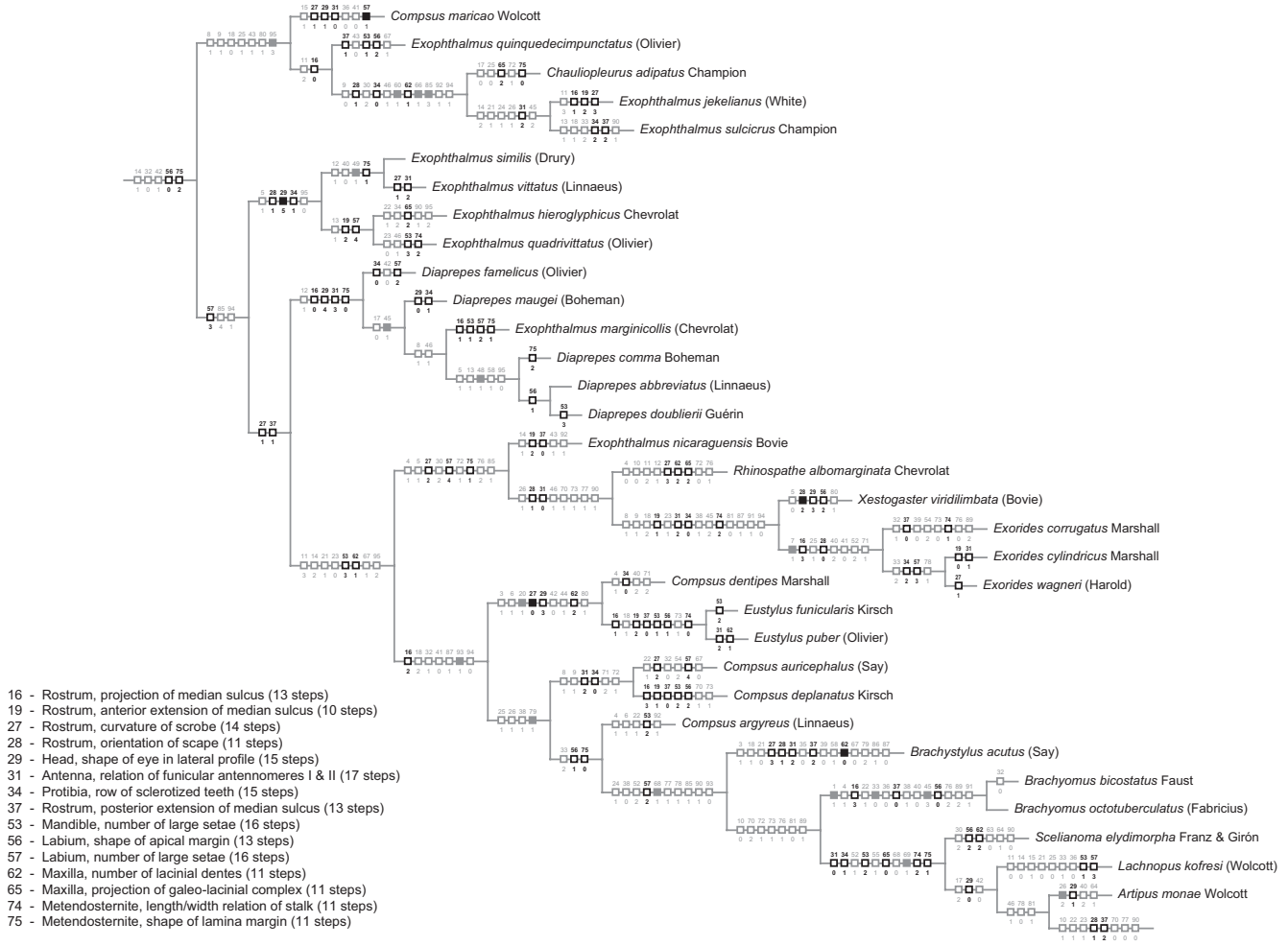

Fig. 10. Example of a section of most parsimonious tree 1 (of four trees in total) at the end of the legacy character assembly stage (matrix 10; L = 547 steps, CI = 28, RI = 66), showing characters and states according to ACCTRAN optimization. Solid rectangles indicate (single) non-homoplasious character state transformations, whereas empty white rectangles indicate (multiple) homoplasious character state transformations. The numbers above and below each rectangle correspond to the coded characters and states, respectively. Fifteen focal characters exceeding a length of 10 steps are highlighted in bold. Bremer branch support values are provided in Fig. 6.

would lead to a character with hundreds of homoplasious steps within the superfamily Curculionoidea. In the case of *this* restricted analysis, several species show "carinae" on the rostrum. But eventually it became apparent that members of the genus *Diaprepes* have a special configuration of rostral carinae (Fig. 12a). I described this as follows (Franz, 2012: 520; character 17): "tricarinate, with a characteristic combination of one median carina and two (dorso-) lateral, apically slightly diverging carinae, each carina narrow, moderately sharp." Taxa *not* assigned to this apomorphic condition include: (1) species of *Exophthalmus* that are primarily monocarinate although they may have lateral rounded elevations (Fig. 12b); (2) species of the outgroup genus *Otiorhynchus* Germar which have three more sharply elevated, mesally clustered, and apically triangularly configured carinae (Fig. 12c); (3) species of *Pachnaeus* Schoenherr that display three blunt carinae (Fig. 12d); (4) species of *Phaops* Sahlberg that are members of a phylogenetically separate clade (Figs 9 and 11), thus making it unlikely that their tricarinate rostrum (Fig. 12e) is "truly" homologous (Brigandt, 2009) to that of *Diaprepes* from which it otherwise differs only subtly (i.e. the carinae are slightly wider and parallel throughout); and (5) species of *Rhinospathe* Chevrolat that are superficially tricarinate but whose median carina is apparent in part because the adjacent lateral regions are concave and inflected (Fig. 12f). In this particular example, then, species of (1) *Exophthalmus* and (3) *Pachnaeus* were coded as "0" (= special condition absent); and the remaining taxa (groups 2, 4, and 5) were coded as "–" (= inapplicable). This dichotomy was established in the preceding character (character 16 in Franz, 2012) where, following concurrent phylogenetic evidence, the homology of the median carina of a grade including

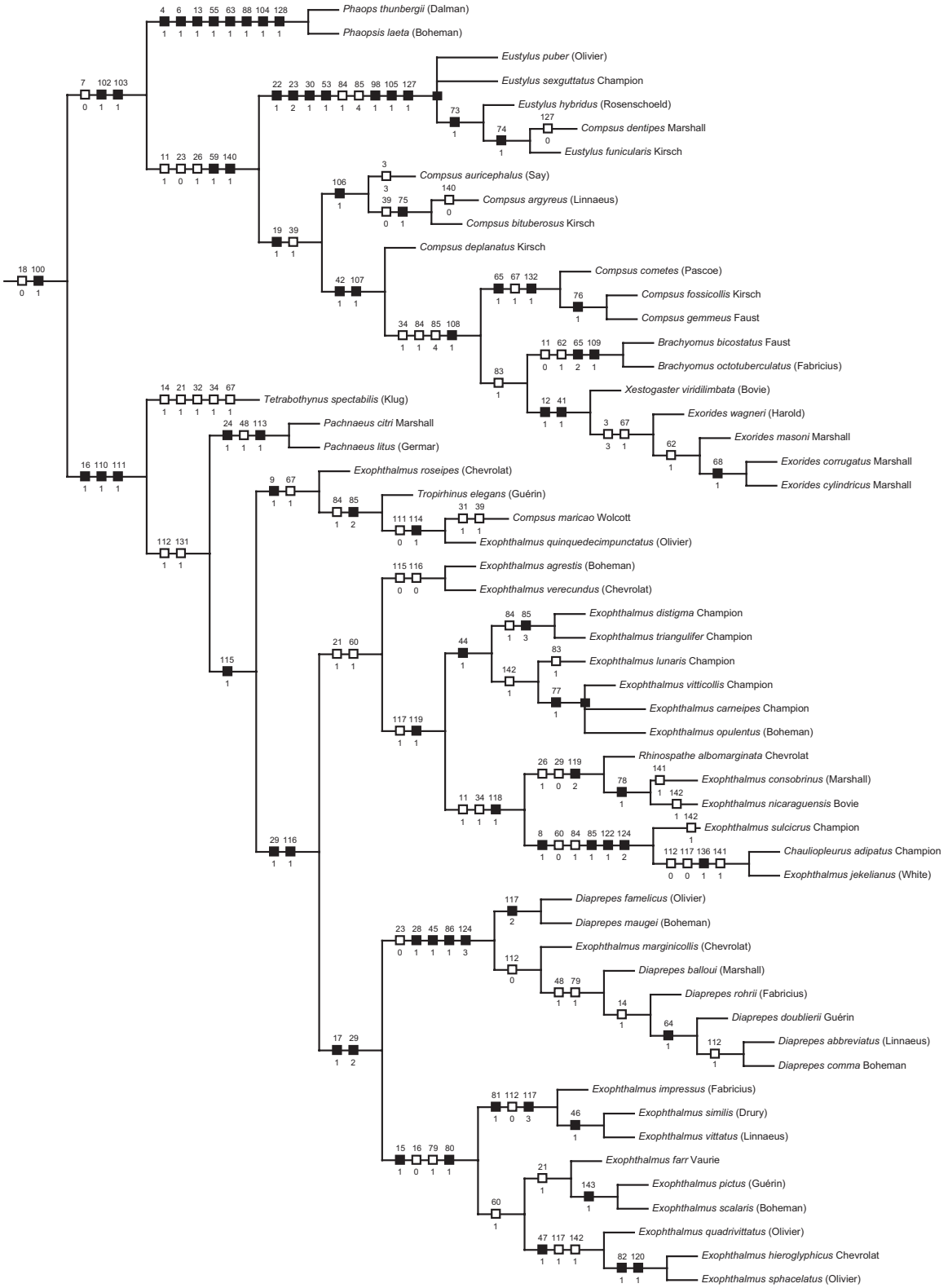

Fig. 11. Example of a section of most parsimonious tree 1 (of eight trees in total) at the publication stage (matrix 60; L = 239 steps, CI = 66, RI = 91), showing characters and states according to ACCTRAN optimization, with unsupported nodes collapsed. Bremer branch support values are provided in Fig. 9 (see also Fig. 10).

Table 3
Examples of five characters present in final matrix 60 (publication stage), with circumscriptions of coded states, scoping refinements, phylogenetic significance, and character statistics (Fig. 13)

Number and character	Coded states	Scoping refinements	Phylogenetic significance	Statistics
3. Labium, shape of prementum in ventral profile	(0) rhomboidal, slightly longer than wide, lateral margins angled (Fig. 13a); (1) subquadrate, all sides similar in length; (2) cordate, lateral margins apically diverging, apicolateral edges variously angled or rounded (Fig. 13d); and (3) subrectangular, wider than long	DELTRAN optimization preferred, as ACCTRAN optimization posits a reversal from rhomboidal (0) to cordate (2) to rhomboidal (0) in the out-group clades *Otiorhynchus* and *Naupactus –Pantomorus*	State (0) is present in *Otiorhynchus* and in the *Naupactus–Pantomorus* clade; state (1) is synapomorphic for *Apodrosus;* state (2) is convergently present in the *Pandeleteius–Paululus* clade and in the Geonemini–Eustylini complex; and state (3) is nested within that complex, and is convergently present in *C. auricephalus* and *Exorides*	$L = 5$, CI = 60, RI = 80
29. Rostrum, ventral side, presence and shape of triangular impression	(0) rostrum ventrally without a large, triangularly shaped impression (Fig. 13b); (1) rostrum ventrally with a short, moderately deep, evenly triangularly shaped impression flanked by hypostomal–labial sutures (Fig. 13c); and (2) impression longer, narrowly triangular, and deeper (Fig. 13d)	Coded as additive, reflecting a presumed evolutionary transition from state (0) to state (1), and to state (2)	State (1) is synapomorphic for the *E. agrestis–E. sphacelatus* clade (although present only in the *E. agrestis –E. jekelianus* clade), with a reversal to state (0) in the *R. albomarginata– E. nicaraguensis* clade; whereas state (2) is synapomorphic for the *D. famelicus– E. sphacelatus* clade	$L = 3$, CI = 66, RI = 97
110. Aedeagus, endophallus, presence of elongate tubular sclerite	(0) endophallus without an elongate tubular sclerite in mid region and (1) endophallus with a variously configured, elongate tubular sclerite in mid region (Fig. 13e, f)	Coded as inapplicable in taxa that lack large endophallic sclerites (see char. 100).	Synapomorphy for the *T. spectabilis– E. sphacelatus* clade	$L = 2$, CI = 100, RI = 100
115. Aedeagus, endophallus, presence of two-part laminate anterior sclerite	(0) anterior endophallic sclerite not separated into two lateral laminate parts and (1) anterior endophallic sclerite separated into two lateral, opposed subrectangular, laminar parts that are connected via membranes (Fig. 13e, f)	Coded as inapplicable in taxa that lack a separate anterior endophallic sclerite (see char. 114); otherwise applicability as in character 112	Synapomorphy for the *E. roseipes– E. sphacelatus* clade, with a reversal in the *E. agrestis–E. verecundus* clade	$L = 2$, CI = 50, RI = 75
116. Aedeagus, endophallus, presence of arched intersclerite connection	0) anterior and posterior sclerites either slightly separate or contiguous, although lacking an arched intersclerite connection and (1) lateral laminae of anterior sclerite with shorter, variously ampullate sclerite that extends into two arched (in lateral profile), posterolateral arms that connect to the posterior sclerite (Fig. 13e, f)	Applicability as in char. 115. ACCTRAN optimization preferred	Accordingly, state (1) is synapomorphic for the *A. agrestis–Exophthalmus sphacelatus* clade, with a reversal in the *E. agrestis–E. verecundus* clade	$L = 2$, CI = 50, RI = 83

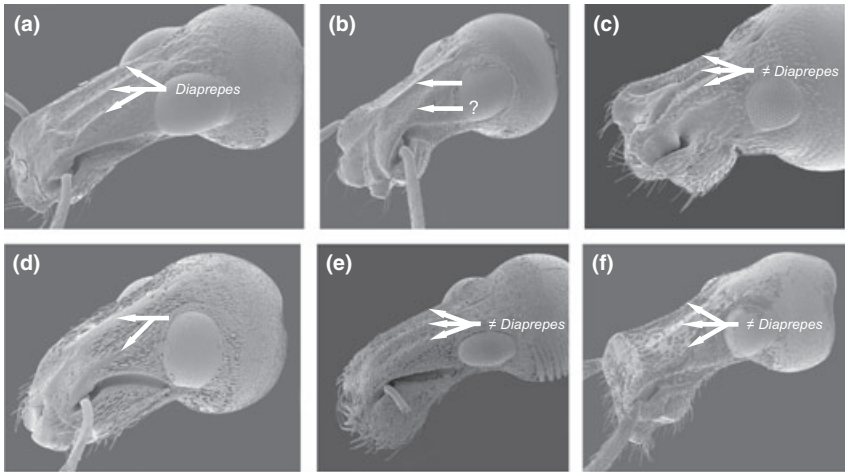

Fig. 12. Dorsolateral aspect of the variously carinate rostra of select species of entimine weevils (see also text). (a) *Diaprepes abbreviatus* (Linnaeus); (b) *Exophthalmus sulcicrus*; (c) *Otiorhynchus meridionalis* Gyllenhal; (d) *Pachnaeus litus* (Germar); (e) *Phaops thunbergii* (Dalman); (f) *Rhinospathe albomarginata* Chevrolat.

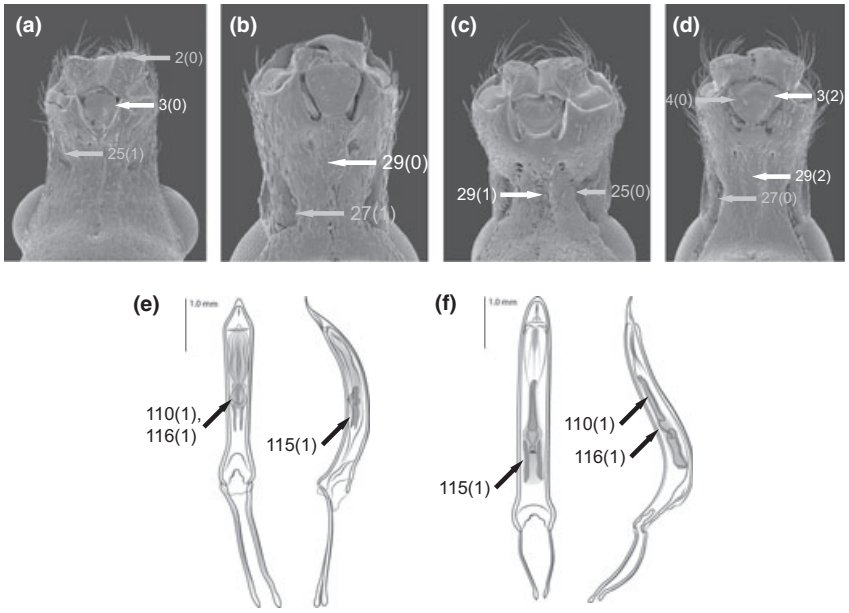

Fig. 13. Illustrations of exemplary characters and states used in matrix 60 (cf. Franz, 2012; Table 3). (a–d) Ventral aspect of rostrum and head: (a) *Naupactus rivulosus* (Olivier); (b) *Lachnopus curvipes* (Fabricius); (c) *Exophthalmus vitticollis* Champion; (d) *Diaprepes abbreviatus*. (e, f) Ventral and lateral aspect of aedeagus: (e) *Diaprepes abbreviatus*; (f) *Exophthalmus quadrivittatus*.

Diaprepes, *Exophthalmus*, and *Pachnaeus* was asserted. Adopting the narrowly scoped coding scheme meant (1) initially observing apparent tricarinate rostra in a wide range of species but then (2) rejecting that description in the cladistic sense for all except those with the special *Diaprepes* configuration. The more specialized anatomical language was created as the predicates "carinate" or "tricarinate" in isolation are not scoped *just* to refer to the homologous *Diaprepes* rostra.

One might view this approach as an attempt to "aim at the historical molecular level" (cf. Rieppel, 2005; Rieppel and Kearney, 2007; Brigandt, 2009). Characters and states are reformulated such that they are most likely to refer to clade-specific historical transformations in the genotype. In the above example, emerging phylogenetic evidence suggested that the tricarinate rostrum of *Diaprepes* species is unique to them in a historical or taxic homology sense, even though in a more shallow phenotypic or ahistorical molecular or ontogenetic sense it is perhaps not unique. The inference of historical/taxic uniqueness is *explicit* in the refined observational language and cladistic coding scheme. Table 3 and the accompanying

Fig. 13 further illustrate how the "deep homology" approach can affect the formulation of homology propositions.

Using this approach, the number of characters increased from 112 in matrix 41 to 164 in matrix 52 (Fig. 2). None of the 52 newly added characters was multi-state, and 41 (78.8%) referred to structures of the terminalia (Fig. 3). The average character length decreased from 1.75 (matrix 41) to 1.63 (matrix 52), along with slight increases in the consistency index (from 59 to 62) and retention index (from 90 to 91; Fig. 4). ILD tests detected no significant difference in topology throughout this phase (Fig. 5). However, very significant changes occurred in the number of MPTs and nodes collapsed in the strict consensus: the former fell from 1277 trees to two trees, and the latter from 36 nodes to one node (Fig. 4). Towards the end of this phase (matrices 48–50; Fig. 4), I experimented with relaxing the practice of coding zeros instead of the more appropriate inapplicables (see phase 5). This had no analytical effect, thus giving me more confidence in the robustness of the outcome. Although clade support was not strong in many instances, I was now approaching a specific hypothesis of phylogeny for the ingroup taxa (compare Figs 8 and 9).

7—Character rescoping phase III (matrices 53–59)

This final rescoping phase had little phylogenetic impact. The topologies of, for example, matrices 53 and 60 are almost identical. The numbers of MPTs and clades resolved varied minimally (Fig. 4). I concentrated on reviewing the language for characters and states so that it would now reflect the theory-laden homology assessments coming out of the previous analysis stages. I also recoded characters with an excessively reductive coding scheme, and thereby recreated several transformationally coherent multi-state characters (cf. Sereno, 2007). The number of binary characters fell from 156 to 132, and the number of multi-state characters rose from three to 11 (Fig. 3). Inapplicables were reintroduced and properly applied to 35 characters, most of them referring to the terminalia (Table 3). Redundant formulations of the same terminalia-referencing characters and states were discarded. For conventional purposes, character states polarities were adjusted to match the overall orientation of cladogram (Nixon and Carpenter, 1993).

Clade support remained moderate, with an average adjusted Bremer branch support of 1.89 per resolved node (Figs 9 and 11). I had kept at least five deactivated characters from previous stages in the matrix to assess whether they would be more informative now than then. This was not the case. In light of the relative topological robustness and my inability to add significant new characters, the analysis was nearing the stage of submission.

A select number of submission-ready characters are shown in Table 3 and Fig. 13 (Franz, 2012). They illustrate the mix of deliberate referential imprecision, precision, and scope restriction that I had decided on at this stage. For instance, character 3 has a traditional multi-state circumscription (shape of labial prementum; Fig. 13a, d), with the only refinement being my preference for a DELTRAN optimization for reasons of evolutionary plausibility (cf. Agnarsson and Miller, 2008). Character 29 was coded as multi-state additive (presence of triangular rostral impression; Fig. 13b, d), a preference that was informed by the emerging phylogenetic results (Fig. 11). Characters 110, 115, and 116 are described with specific domains of (frequently nested) applicability, thus reflecting the presumed localized appearances and subsequent modifications of endophallic structures within Caribbean and Central American species of *Exophthalmus* and close relatives (Fig. 13e, f). In each case there is detailed topographic information, and the inferred plesiomorphic and apomorphic conditions are either (1) very specific (e.g. "endophallus *without* an elongate tubular sclerite"—character 110, state 0; "anterior endophallic sclerite separated into two lateral, opposed subrectangular, laminar parts…"—character 115, state 1) or (2) purposefully ambiguous (e.g. "anterior and posterior sclerites *either* slightly separate *or* contiguous, although…"—character 116, state 0; "endophallus with a *variously configured*, elongate tubular sclerite in mid region"—character 110, state 1). Such context-specific qualifications are expected under an "aim for the molecular level" approach.

8—Publication stage

I had no problems in publishing this paper with minor revisions.[4] One referee (all were anonymous) called the manuscript "excellent", whereas the other two referees were more restrained. Both discussed the merits of my aggressive scoping of characters and states. One referee stated, perceptively:

> I note that most characters have a CI = 1, something unusual in taxonomically challenging and difficult weevil groups like the one herein studied. […] Many character state changes presented as unique, non-homoplasious (black rectangles) in

[4]This is not meant to imply much about the study's quality. Time will hopefully tell; the reliability of scientific inferences is an *a posteriori* phenomenon (Boyd, 1983). For related reasons, I also need not elaborate here on independently generated and still unpublished molecular evidence in support of this morphological analysis (Mazo-Vargas, A. et al., in preparation).

> Fig. 2, could have been shown as independently evolved in different clades if they were not reformulated. […] Even with this reformulation of characters, the support of the clades, particularly the "tribes" is not high. However, the level of resolution is good, something not easy to achieve when only adult morphology is used to resolve species or genus level relationships.

This statement presents an interesting contrast with Thompson's (1992) quote in the introduction. The trade-off is, either using analytically contingent inferences with gains in resolution, or adhering to more "objective" conceptions of similarity at the cost of ambiguous results. I responded with yet another quote, taken from Davis's (2011) cladistic analysis of baridine weevils. In that analysis of 301 taxa and 113 characters, the author used a more conventional approach to character individuation, which yielded an average character length of 39.9 steps per character on the 33 MPTs ($L = 4509$ steps; CI = 5; RI = 51). The visual impression of mapped character state optimizations in that publication (the author's figure 117) resembles that of Fig. 10. In a section denominated "final thoughts and future directions", the author writes (Davis, 2011, p. 138):

> […] homoplasies not only give structure to trees as synapomorphies do, but they also delineate which characters do not have the same qualities. […] separating the amount of homoplasy that is the result of noise and that which is the result of evolutionary history is integral in examining and improving large phylogenetic studies such as this one.

More data will probably help sort out these questions (Rieppel, 2007b). However, differences in homology formulations are significant as well. Had the reviewers challenged me strongly on this issue, I would have had the ability to illustrate the options on hand.

Review of stages 1–8—topology changes and implications for clade inference

In review, the trajectory of the analysis produced at least five significant changes in tree topology among succeeding matrix versions (Fig. 5). All of these occurred during the first three stages of analysis, with the transition from the taxon addition phase to the character readdition phase being the last major shift (matrices 26 and 27). Many clades inferred during the final stages were not fully formed at earlier stages (cf. Figs 6–9). Several species or species groups were particularly difficult to accommodate, as is reflected in their vastly different placements throughout the analysis. Examples include *Brachystylus* Schoenherr, *Pachnaeus*, *Phaops*, *Scelianoma* Franz & Girón, and all members of the *E. roseipes–E. quinquedecimpunctatus* clade (as identified in Fig. 9). The position of these "floating taxa" began stabilizing once the phylogenetic sampling gaps of the early stages were reduced.

The later stages saw the emergence of several geographically well-circumscribed lineages (Fig. 9); i.e. (1) a largely Caribbean geonemine grade (*Cleistolophus–Lachnopus*), (2) a South American eustyline clade (*Phaops–Exorides*), and (3) a Caribbean/Central American clade which includes all examined species of the highly paraphyletic *Exophthalmus*, and nested within them the exclusively Caribbean *Diaprepes*. Such geographical consilience is relevant and encouraging. Within the inclusive *Exophthalmus* clade, as many as 25 subclades have at least one unreversed synapomorphy (Fig. 11), resulting in relatively short, unqualified, and testable phylogenetic diagnoses (for details see Franz, 2012, pp. 546–548). Only two comparable clades were recovered at an earlier stage of analysis (matrix 10; Fig. 10).

Discussion

On the analytical process-dependency of cladistic character individuation

The actions taken in my analysis could be measured by revisiting long-standing arguments about the relationship of "proper" cladistic practice and "proper" philosophy of science (cf. Franz, 2005a; Fitzhugh, 2006; Rieppel, 2009). But perhaps this would not constitute the most fruitful reading. Hence my focus for discussion is not whether I conducted "good", or well justified, science. By not putting forward the best possible homology statements in matrices 1–10, I have introduced bias towards preconceived notions (Franz, 2005a). In arriving at the late-stage homology assessments, I may have utilized methods that are not well justified, or failed to utilize better suited methods (Pogue and Mickevich, 1990; Hawkins, 2000; Grant and Kluge, 2003; Goloboff et al., 2006, 2008; Ramírez, 2007; Catalano et al., 2010). And in the end, my inferences may turn out partially or entirely unreliable; worse still, history was not sufficiently information-preserving to allow for a reliable and precise cladistic outcome (Felsenstein, 1978; Farris, 1983; Sober, 1988). None of these otherwise reasonable caveats is particularly relevant to the following discussion.

Instead, I meant to demonstrate *how* some aspects of the actual *process* of matrix production manifest themselves in a specific case study under the cladistic parsimony paradigm. In doing so, I showed that the process of positing and refining homology assessments *may* be driven by the idiosyncratic dynamics of the taxa and evolutionary phenomena under study, and by likelihood considerations and projections of the reliability of early- versus late-stage encodings of phyloge-

netic informations (Wägele, 2005; Sereno, 2009; Vogt et al., 2010). While the specifics of my analysis are what they are, the need to empirically weigh and scope evidence so as to reflect homology at the targeted depth of inference is shared among most analyses including molecular studies (Wenzel and Siddall, 1999; Sanderson and Shaffer, 2002; Kearney and Rieppel, 2006; Lienau et al., 2006; Rodríguez-Ezpeleta et al., 2007; Philippe et al., 2011). Granted, other analyses may start off with a more scrutinized legacy of characters than I had access to (Table 2). But many systematists will share the experience that *some* traditionally diagnostic and informative traits—once submitted to the congruence test and mapped onto a cladogram (Fig. 10)—must be reconfigured to better reflect phylogenetic similarity.

Thus I suggest that analytical phylogenetic methods not only *organize* character information, but furthermore have the purpose of *shaping character individuation* (Franz, 2005a; Rieppel, 2007a; Sereno, 2009; Vogt et al., 2010).[5] While I could have done better than just code poorly performing legacy characters in my initial matrices (Fig. 10, Table 2), I should have been hard pressed to arrive at the final descriptions of characters and states without benefitting from intermittent parsimony-driven inferences that led to the reweighing and rescoping of earlier homology assessments (Figs 11–13, Table 3). Under the cladistic paradigm the most precise inferences of homology are parsimony-influenced and parsimony-contingent, and the two notions are inextricably linked and *entrenched* in our maturing observational terminology.

If this practice is applied over the course of a single study, or across multiple succeeding studies, then the metaphor of a circular or spiralling successive approximations approach towards positing homology is apt. Cladistic practice rarely if ever adheres to linear epistemological trajectory (Wägele, 2005; Sereno, 2009), i.e. singular observation (primary stage) → singular congruence test (secondary stage) → acceptance of outcomes for publication (tertiary stage). While some authors may regard an iterative trial-and-error approach as "tinkering" (or worse),[6] I suggest that using algorithmic tools to improve character individuation is more ubiquitous than many published works let on. It is not always conducive to a researcher's reputation to expose these practices. They can seem less than "rigorous", but they do and must occur frequently for cladistics to succeed.

Implications of cladistic character individuation for phenotype ontology creation

The use of algorithmic projections to improve character individuation is not restricted to phenotypic traits (Sanderson and Shaffer, 2002), although it is that latter realm where they have historically made the most genuine and valuable contributions to systematics (Wheeler, 2004; Franz, 2005a; Assis, 2009; Assis and de Carvalho, 2010; Mooi and Gill, 2010). This is so because systematists whose research addresses the projectibility of phenotypic characters are more inclined to make explicit *linguistic* contributions to the field (Franz, 2005b; though see Scotland et al., 2003; for an alternative view). Their contributions may take the form of phylogenetically scoped characters and character state circumscriptions, phylogenetic diagnoses of clades that reference inferred synapomorphies and relevant homoplasious traits, and phylogenetically revised classifications. This leads to the predictive, causally grounded reference system that Hennig (1966) advocated and which has empowered phenotype-centred systematic research with a new epistemological quality.

As phenotype-centred research continues to position itself in the age of phylogenomics (Philippe et al., 2005, 2011; Bybee et al., 2010), there are opportunities to reappraise and reinforce the unique linguistic contributions that systematists make to a field now flooded with complementary molecular data and myriad tools for statistical analysis and visual data representation (Franz, 2005a,b; Rieppel, 2007b; Assis, 2009; Assis and de Carvalho, 2010). One can assume that phenotypic traits will remain central to narratives about evolutionary phenomena, although at increasingly large scales of analysis (Ramírez et al., 2007; Balhoff et al., 2010; Dahdul et al., 2010a; Franz and Thau, 2010; Deans et al., 2012a).

The advent of ontology-based representations of phenotypic data promises a pathway for large-scale *syntheses* of genomic and phenotypic information. Phenotype ontologies are controlled, structured vocabularies for phenotypic traits—i.e. hierarchies or networks of entity-quality expressions—amenable to computerized representation and reasoning (for review see Deans et al., 2012a,b). They are in widespread use in the model organism research communities and are becoming more prevalent in comparative biology (Franz and Thau, 2010; Vogt et al., 2010). However, the relationship between the sort of theory-laden homology statements advocated in this study and computable entity-quality expressions is not straightforward. Indeed, phenotype ontology design has so far largely bypassed the notion of cladistic homology. For

[5]Hennig's (1966) notion of reciprocal illumination might be interpreted in the same general direction, but I shall not presume that he would have endorsed the views and practices presented in this study.

[6]This view is not incompatible with the notion that one can *misuse* algorithmic methods to produce a preconceived result. Nor would it justify the view that algorithms *in themselves* make reliable science. The verdict of *a posteriori* reliability affects all systematic methods.

instance, Seltmann et al. (2012, pp. 79) write (with specific reference to the Hymenoptera Anatomy Ontology, HAO; Yoder et al., 2010):

> Fundamentally, the HAO project rests on recognizing different instances of a topographically-defined concept as "the same" (e.g., the fore wing of taxon A is the same structure as the fore wing of taxon B) […] The HAO employs the principle of "structural equivalence" to discuss topographical sameness. In biology, however, homology is often more explicit, referring to a more profound "sameness," because it expresses a theory about structures sharing a common evolutionary origin even if they appear structurally dissimilar […] The dynamic nature of homology hypotheses conflicts with the HAO's goal of unambiguous circumscription of anatomical concepts, and, as such, overt references to homology hypotheses are avoided in constructing HAO definitions.

Similarly, Mungall et al. (2012, pp. 12–13) clarify, with regard to the construction of Uberon, an overarching, cross-species phenotype ontology:

> […] Uberon contains grouping classes "eye" and "wing," despite the fact that neither of these are homophyletic—they evolved multiple times. The inclusion of a class in the ontology should not be taken as an indication of shared evolutionary descent (homology), merely that classes have some property or properties in common. We have taken an integrative approach in the building of Uberon, and in doing so embrace multiple axes of classification. […] This homology-neutrality of Uberon is a deliberate design feature of the ontology.

Historically versed readers may infer that the phrase "multiple axes of classification" can be code for: *not* phylogenetic systematics (Windsor, 2000). Lastly, in a more sophisticated argument for homology neutrality in phenotype ontologies, Vogt et al. (2010, p. 306; see also Vogt, 2009) state:

> Explanatory homology hypotheses should not be mistaken and blended with morphological descriptions, which in their turn are by nature descriptive and not explanatory. […] we differentiate phylogenetic investigations into the step of producing data and the step of phylogenetic reasoning. Producing data includes conducting morphological studies and generating well-documented descriptions of organismic traits, but excludes primary homology assessment and matrix generation. Phylogenetic reasoning, by contrast, is the step of establishing a phylogenetic argumentation by identifying, delimiting, and evaluating evidence units, which includes primary homology assessment, character coding, and alignment or matrix generation, and their evaluation (i.e. weighting) during numerical tree inference.

One can sympathize with the objective of having formalized and well-defined structural phenotypic information available for large-scale analyses, although the motivations for this objective may lie outside of the realm of systematics proper (cf. Dahdul et al., 2010b).[7] Homology expressions would then constitute an additional, optional layer of annotations made "on top of" a phenotype-referencing ontology. The latter, in turn, is tied to "immediate observations" of structures that permits reasoning in various homology-neutral contexts.

Though one might ask, looking at my phylogenetic analysis of the *Exophthalmus* genus complex as reconstructed above, where *exactly* should I have stopped? Was the transition from matrix 10 to matrix 11 the turning point from "data production" to "phylogenetic reasoning"? Is it most appropriate for ontology creation and reasoning to work with an average character length of nearly 5.5 steps (Figs 2 and 10)? Or should we consider the six weevil rostra illustrated in Fig. 12 just as "tricarinate", in spite of phylogenetic inferences suggesting more precise labels? Or perhaps I should have been more "descriptive" still, thus somehow setting aside the homology implied by using such terms as "lacinial teeth" and "galeo-lacinial complex" (Ting, 1936) when counting modified setae on the maxillae of weevil specimens (Table 2). After all, these descriptive terms *do* carry theoretical, parsimony-employing implications of how weevil mouthparts evolved and are homologous to those of (e.g.) grasshoppers (Grimaldi and Engel, 2005). In short, where does one transparently and consistently draw the line between "mostly context neutral" and "fully context dependent" in a continuous spiralling process of cladistic character refinement within and across analyses?

Polemics aside, it is safe to say that in actual practice the expressions used in phenotype ontologies are *not* homology neutral or homology agnostic. They cannot and should not have these attributes if they grew out of the linguistic legacy of phenotype-centred systematic research. Instead, it is more accurate to say that structural equivalence expressions occurring in such ontologies are *phylogenetically underdetermined* (Franz and Thau, 2010). It is precisely the lack of full phylogenetic determinacy that would allow structural equivalence expressions to participate in reasoning *beyond* the scope delimited by refined taxic homology and non-erroneous (correctly inferred and narrowly scoped) homoplasy. Clearly, a context-relaxing approach towards homology can reveal a wealth of important phenotypic patterns at large phylogenetic scales (cf. Dahdul et al., 2010b). However, systematists might argue that relaxing homology criteria lies beyond *their* best interests.

Neutralizing the phylogenetic context-contingency runs counter to the objective of having a maximally precise and predictive terminology (Rieppel, 2007a). In particular, by failing to deploy the most refined terms as the basis for phenotypic similarity, we inflate

[7] At root, this approach is probably traceable to an acceptance of "universals", i.e. "an invariant pattern in reality which is multiply exemplified in an indefinitely extendable range of different instances" (Smith, 2006, p. 292). For further discussion see Merrill (2010a,b) and Smith and Ceusters (2010); reviewed in Franz and Goldstein (2013).

the population of positive instances for targeted phenotypes. As Proctor (1996, p. 144) argues in a separate yet related context (Wenzel and Carpenter, 1994):

> Ecological researchers are likely to define their characters very broadly for two reasons: first, because the phenomenon of interest may occur in a wide range of taxa and thus is unlikely to be similar in the narrow, phylogenetically homologous sense; second, because statistical power increases with the sample size (number of independent evolutions). Thus ecologists define their characters of interest very broadly in order to *maximize the probability of homoplasy*. Rarely are the behaviors investigated in comparative studies thought to be true homologs, rather they are suites of behaviors serving similar functions.

Similarly, one can gain broader inference powers by constructing structural equivalence expression in ontologies, at the great cost of accepting phylogenetic underdeterminacy. The gains will be most beneficial to disciplines other than systematics. If there is indeed a trade-off between breadth and depth of inference (Brachman and Levesque, 2004), then it is prudent for phenotype ontology design to pursue both directions. Recent attempts to clearly express and embed in phenotype ontologies the specific nature and phylogenetic scope of homology-referencing expressions should be developed further (cf. Parmentier et al., 2010; Balhoff et al., 2011; Travillian et al., 2011; Mungall et al., 2012; Niknejad et al., 2012). Research in this direction is needed to fully deploy the insights of systematic reasoning in an ontology-based framework.

Conclusions and recommendations

Using a microscopic, temporally sliced dissection of my cladistic analysis of the *Exophthalmus* genus complex as a test case, I have shown that the process of individuating homologous characters and states under the cladistic paradigm is strongly influenced by parsimony considerations and related assessments of evolutionary plausibility. It is instructive—to say the least—to observe how an initial set of proposed homologies is optimized for the first time on a parsimony-based cladogram (Fig. 10). Parsimony-informed scope adjustment is most relevant if characters under new cladistic scrutiny stem from a systematic legacy that has lacked precise phylogenetic scoping.

I have suggested that this *kind* of parsimony-influenced rescoping of character information is in principle justifiable (Franz, 2005a) and in practice necessary and widespread. The *extent* to which rescoping occurs in an analysis may vary greatly, depending (*inter alia*) on a researcher's expertise and ability to utilize and improve upon the quality of the existing terminology. Over time and across multiple cladistic studies, this terminology will become increasingly precise and erroneous statements of homology and homoplasy will become less frequent.

To the extent that narrowly scoped homology assessments and terms hold up as synapomorphies and relevant homoplasious features of their corresponding clades, these terms become entrenched in our phylogenetic vocabulary and adopted by other biological disciplines. One would expect a strong correlation between the degree of entrenchment of a specific homology-referencing expression and its performance in parsimony-based studies. In the long term, repeated testing for congruence under parsimony may contribute to the proper phylogenetic scoping of virtually every character and state represented in a cladistic matrix (cf., Tables 2 and 3 and Figs 10 and 11). In that sense, the notions of parsimony, congruence, and homology are inextricably linked in cladistics, and embedded in its emerging observational language.

The ability to refine traditional phenotype-referencing terms through cladistic treatment puts systematists in a privileged position. Phenotypic traits are most immediately connected to our sensory capabilities. These verbalized traits remain most central to creating explanatory and predictive classifications and evolutionary narratives. In addition to evaluating the phylogenetic adequacy of previously proposed traits, phenotype-centred research can (1) significantly *reconfigure* how these traits are individuated, and (2) *propose* new homologous traits to be integrated with the existing information. More annotational transparency and cross-analytical compatibility are urgently needed (Sereno, 2009), and would enhance the creation and curation of a precise language for phylogenetic phenomena.

The advent of phenotype ontologies offers much promise for adopting formalized definitions and relationships for phenotypic structures (cf. Deans et al., 2012a). However, by integrating expressions of structural equivalence at increasingly greater scales, these ontologies also run the risk of "dialling down" the most precise and phylogenetically scoped assessments of homology that systematists can produce. There is room for exploring ontologies that make fuller use of the context- and parsimony-interdependencies that characterize a truly cladistic language for phenotypic traits.

Reviewers of an earlier version of this paper suggested that I end with a set of clear and practical recommendations. I do this here, with the following caveats: (1) I cannot regard the recommendations as new; and (2) as someone who favours a naturalistic, explanatory perspective towards scientific methods I believe some vagueness (as opposed to strict normativity) is necessary.

1 Systematists who work directly with phenotypic features should *not* take the existing observational terminology for putatively homologous traits as a given.

A suitable starting assumption is that this terminology will be phylogenetically *underdetermined* to an unknown degree and needs refinement.

2 The referential *refinement* of this terminology—i.e. a more precise language that more narrowly maps onto synapomorphic and properly inferred homoplasious traits—should be an objective and expectation for new cladistic analyses. Systematics is a language that aims to map precise terms onto inferred historical biological phenomena. New systematic insights should have tangible and clearly marked linguistic consequences.

3 The congruence test, along with an evolving set of additional parsimony-implementing inference tools, may be used *flexibly* and at virtually *any stage* of analysis to reassess the performance and plausibility of intermittent homology propositions. Often this can lead to down-weighing or discarding legacy assessments, or to subdividing broad homology propositions into sets of separate and phylogenetically more localized characters and states, coding each (essentially) as if the other were non-existent. In some relevant sense, history never repeats and "acceptable" homoplasy may simply reflect one's inability to recognize and formulate differences among traits after every attempt was made to do so. However, it is acceptable to posit (e.g.) "tricarinate as in this clade" (character 1: present/absent) versus "tricarinate as in that clade" (character 2: present/absent) if no better formulation is possible.

4 Systematic epistemology should *embrace* the contingent, theory-laden, and intersubjective nature of cladistic homology propositions. Good, rigorous practice means annotating the underlying considerations for a particular solution extensively, laying out plausible alternatives and their effects of cladistic outcomes, signalling uncertainty and vagueness in one's assessments, and making explicit use of the contingencies that reflect the specific scope of an analysis.

5 One should be *bold* in the practice of seeking the best-fitting scope to codify homology; however, a good working criterion for stopping is when one senses that certain codings are no longer defensible in discussions with one's most highly regarded peers.

6 Systematists should understand that sophisticated, parsimony-contingent formulations of homology pose a challenge for ontology-based representation and reasoning, particularly if the latter adheres to the tenet of homology neutrality. However, alternative schools of ontology design hold that ontologies should follow the representation and inference needs of particular scientific disciplines, and not vice versa (Merrill, 2010a,b). Until such ontologies are developed, it is prudent not to water down the special linguistic contributions that systematists bring to biology by virtue of their parsimony-contingent homology propositions and scope refinements.

Acknowledgements

I thank Juliana Cardona-Duque, Jennifer Girón Duque, Anyimilehidi Mazo-Vargas, and three anonymous referees for helpful feedback on issues regarding the homology of entimine weevil structures in the process of developing the original analysis upon which this contribution is based. My research on Neotropical entimine weevils was funded in part by the National Science Foundation, awards DEB-641231 and DEB-1155984.

References

Adams, E.N., 1972. Consensus techniques and the comparison of taxonomic trees. Syst. Zool. 21, 390–397.

Agnarsson, I., Miller, J.A., 2008. Is ACCTRAN better than DELTRAN? Cladistics 24, 1–7.

Alonso-Zarazaga, M.A., Lyal, C.H.C., 1999. A World Catalogue of Families and Genera of Curculionoidea (Insecta: Coleoptera) (Excepting Scolytidae and Platypodidae). Entomopraxis, Barcelona, Spain.

Anderson, R.S., 2002. Family 131. Curculionidae. In: Arnett., R.H. Jr, Thomas, M.C., Skelley, P.E., Frank, J.H. (Eds.), American Beetles, Polyphaga: Scarabaeoidea to Curculionoidea. CRC Press, Boca Raton, FL, Vol. 2, pp. 722–815.

Andersson, M., 1994. Sexual Selection. Princeton University Press, Princeton, NJ.

Assis, L.C.S., 2009. Coherence, correspondence, and the renaissance of morphology in phylogenetic systematics. Cladistics 25, 528–544.

Assis, L.C.S., de Carvalho, M.R., 2010. Key innovations: further remarks on the importance of morphology in elucidating systematic relationships and adaptive radiations. Evol. Biol. 37, 247–254.

Balhoff, J.P., Dahdul, W.M., Kothari, C.R., Lapp, H., Lundberg, J.G., Mabee, P., Midford, P.E., Westerfield, M., Vision1, T.J., 2010. Phenex: ontological annotation of phenotypic diversity. PLoS One 5, e10500.

Balhoff, J.P., Midford, P.E., Lapp, H., 2011. Integrating anatomy and phenotype ontologies with taxonomic hierarchies. In: Proceedings of the Second International Conference on Biomedical Ontology, July 26–30, 2011. Buffalo, NY, pp. 426–427. Available at http://icbo.buffalo.edu/ICBO-2011_Proceedings.pdf (accessed 5 June 2012).

Barker, F.K., Lutzoni, F., 2002. The utility of the incongruence length difference test. Syst. Biol. 51, 625–637.

Bouchard, P., Bousquet, Y., Davies, A.E., Alonso-Zarazaga, M.A., Lawrence, J.F., Lyal, C.H.C., Newton, A.F., Reid, C.A.M., Schmitt, M., Ślipiński, S.A., Smith, A.B.T., 2011. Family-group names in Coleoptera (Insecta). ZooKeys 88, 1–972.

Boyd, R.N., 1983. On the current status of the issue of scientific realism. Erkenntnis 19, 45–90.

Brachman, R.J., Levesque, H.J., 2004. Knowledge Representation and Reasoning. Morgan Kaufmann Publishers Inc., San Francisco, CA.

Brady, R.H., 1994. Pattern description, process explanation, and the history of morphological sciences. In: Grande, L., Rieppel, O. (Eds.), Interpreting the Hierarchy of Nature. From Systematic Patterns to Evolutionary Process Theories. Academic Press, San Diego, CA, pp. 7–31.

Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.

Brigandt, I., 2009. Natural kinds in evolution and systematics: metaphysical and epistemological considerations. Acta Biotheor. 57, 77–97.

Bybee, S.M., Zaspel, J.M., Beucke, K.A., Scott, C.H., Smith, B.W., Branham, M.A., 2010. Are molecular data supplanting

morphological data in modern phylogenetic studies? Syst. Entomol. 35, 2–5.
Catalano, S.A., Goloboff, P.A., Giannini, N.P., 2010. Phylogenetic morphometrics (I): the use of landmark data in a phylogenetic framework. Cladistics 26, 539–549.
Champion, G.C., 1911. Otiorhynchinae Alatae. In: Biologia Centrali-Americana. British Museum (Natural History), London, Vol. 4, Part 3, pp. 178–317.
Dahdul, W.M., Balhoff, J.P., Engeman, J., Grande, T., Hilton, E.J., Kothari, C., Lapp, H., Lundberg, J.G., Midford, P.E., Vision, T.J., Westerfield, M., Mabee, P.M., 2010a. Evolutionary characters, phenotypes and ontologies: curating data from the systematic biology literature. PLoS One 5, e10708.
Dahdul, W.M., Lundberg, J.G., Midford, P.E., Balhoff, J.P., Lapp, H., Vision, T.J., Haendel, M.A., Westerfield, M., Mabee, P.M., 2010b. The Teleost Anatomy Ontology: anatomical representation for the genomics age. Syst. Biol. 59, 369–383.
Davis, S.R., 2011. Delimiting baridine weevil evolution (Coleoptera: Curculionidae: Baridinae). Zool. J. Linn. Soc. 161, 88–156.
Deans, A.R., Yoder, M.J., Balhoff, J.P., 2012a. Time to change how we describe biodiversity. Trends Ecol. Evol. 27, 78–84.
Deans, A.R., Mikó, I., Wipfler, B., Friedrich, F., 2012b. Evolutionary phenomics and the emerging enlightenment of arthropod systematics. Invertebr. Syst. 26, 323–330.
van Emden, F.I., 1944. A key to the genera of Brachyderinae of the world. Ann. Mag. Nat. Hist., 11, 503–532, 559–586.
Farris, J.S., 1982. Outgroups and parsimony. Syst. Zool. 31, 328–334.
Farris, J.S., 1983. The logical basis of phylogenetic analysis. In: Platnick, N.I., Funk, V.A. (Eds.), Advances in Cladistics. Columbia University Press, New York, NY, Vol. 2, pp. 7–36.
Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of congruence. Cladistics 10, 315–319.
Felsenstein, J., 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27, 401–410.
Fitzhugh, K., 2006. The abduction of phylogenetic hypotheses. Zootaxa 1145, 1–110.
Franz, N.M., 2005a. Outline of an explanatory account of cladistic practice. Biol. Philos. 20, 489–515.
Franz, N.M., 2005b. On the lack of good scientific reasons for the growing phylogeny/classification gap. Cladistics 21, 495–500.
Franz, N.M., 2006. Towards a phylogenetic system of derelomine flower weevils (Coleoptera: Curculionidae). Syst. Entomol. 31, 220–287.
Franz, N.M., 2010a. Redescriptions of critical type species in the Eustylini Lacordaire (Coleoptera: Curculionidae: Entiminae). J. Nat. Hist. 44, 41–80.
Franz, N.M., 2010b. Revision and phylogeny of the genus *Apotomoderes* Dejean (Coleoptera: Curculionidae: Entiminae). ZooKeys 49, 33–75.
Franz, N.M., 2012. Phylogenetic reassessment of the *Exophthalmus* genus complex (Curculionidae: Entiminae: Eustylini, Geonemini). Zool. J. Linn. Soc. 164, 510–557.
Franz, N.M., Engel, M.S., 2010. Can higher-level phylogenies of weevils explain their evolutionary success? A critical review. Syst. Entomol. 35, 597–606.
Franz, N.M., Goldstein, A.M., 2013. Phenotype ontologies: are homology relations central enough? A reply to Deans et al.. Trends Ecol. Evol. 28, 131–132.
Franz, N.M., Thau, D., 2010. Biological taxonomy and ontology development: scope and limitations. Biodiv. Inform. 7, 45–66.
Giribet, G., 2003. Stability in phylogenetic formulations and its relationship to nodal support. Syst. Biol. 52, 554–564.
Girón, J.C., Franz, N.M., 2010. Revision, phylogeny, and historical biogeography of the genus *Apodrosus* Marshall, 1922 (Coleoptera: Curculionidae: Entiminae). Insect Syst. Evol. 41, 339–414.
Girón, J.C., Franz, N.M., 2012. Phylogenetic assessment of the Caribbean weevil genus *Lachnopus* Schoenherr, 1840 (Coleoptera: Curculionidae: Entiminae). Invertebr. Syst. 25, 67–82.
Goloboff, P.A., 1999. NONA, Version 2.0 (for Windows). Published by the author, Tucumán, Argentina. Available at http://www.cladistics.com.
Goloboff, P.A., Mattoni, C.I., Quinteros, A.S., 2006. Continuous characters analyzed as such. Cladistics 22, 589–601.
Goloboff, P.A., Carpenter, J.M., Arias, J.S., Miranda-Esquivel, D.R., 2008. Weighting against homoplasy improves phylogenetic analysis of morphological data sets. Cladistics 24, 758–773.
Grant, G., Kluge, A.G., 2003. Data exploration in phylogenetic inference: scientific, heuristic, or neither. Cladistics 19, 379–418.
Grimaldi, D., Engel, M.S., 2005. Evolution of the Insects. Cambridge University Press, New York, NY.
van Harmelen, F., Livschitz, V., Porter, D. (Eds.), 2008. The Handbook of Knowledge Representation. Elsevier, Oxford, UK.
Hawkins, J.A., 2000. A survey of primary homology assessment: different botanists perceive and define characters in different ways. In: Scotland, R., Pennington, R.T. (Eds.), Homology and Systematics. Coding Characters for Phylogenetic Analysis. Taylor & Francis, London, UK, pp. 22–53.
Hennig, W., 1966. Phylogenetic Systematics. University of Illinois Press, Urbana, IL.
Hipp, A.L., Hall, J.C., Sytsma, K.J., 2004. Congruence versus phylogenetic accuracy: revisiting the incongruence length difference (ILD) test. Syst. Biol. 53, 81–89.
Hustache, A., 1929. Curculionides de la Guadeloupe, Première Partie. In: Gruvel, A. (Ed.), Faune des Colonies Françaises. Société d'Éditions Géographiques, Maritimes et Coloniales, Paris, Vol. 3, pp. 165–267.
Jenner, R.A., 2004. The scientific status of metazoan cladistics: why current research practice must change. Zool. Scrip. 33, 293–310.
Kearney, M., 2002. Fragmentary taxa, missing data, and ambiguity: mistaken assumptions and conclusions. Syst. Biol. 51, 369–381.
Kearney, M., Rieppel, O., 2006. Rejecting "the given" in systematics. Cladistics 22, 369–377.
Kuschel, G., 1955. Nuevas sinonimias y anotaciones sobre Curculionoidea. Rev. Chil. Entomol. 4, 261–312.
Lacordaire, T., 1863. Histoire Naturelle des Insectes. Genera des Coléoptères ou Exposé Méthodique et Critique de Tous les Genres Proposés jusqu"ici dans cet Ordre d"Insectes. Roret, Paris, Vol. 6.
LeConte, J.L., Horn, G.H., 1876. The Rhynchophora of America north of Mexico. Proc. Am. Philos. Soc. 15, 1–455.
Lienau, E.K., de Salle, R., Rosenfeld, J.A., Planet, P.J., 2006. Reciprocal illumination in the gene content tree of life. Syst. Biol. 55, 441–453.
Lord, P., Stevens, R., 2010. Adding a little reality to building ontologies for biology. PLoS One 5, e12258.
Maddison, W.P., 1993. Missing data versus missing characters in phylogenetic analysis. Syst. Biol. 42, 576–581.
Mahner, M., Bunge, M., 1997. Foundations of Biophilosophy. Springer, Berlin, Germany.
Mayr, E., Bock, W.J., 2002. Classifications and other ordering systems. J. Zoolog. Syst. Evol. Res. 40, 169–194.
Merrill, G.H., 2010a. Ontological realism: methodology or misdirection? Appl. Ontol. 5, 79–108.
Merrill, G.H., 2010b. Realism and reference ontologies: considerations, reflections and problems. Appl. Ontol. 5, 189–221.
Mickevich, M.F., Weller, S.J., 1990. Evolutionary character analysis: tracing character change on a cladogram. Cladistics 6, 137–170.
Mindell, D.P., Thacker, C.E., 1996. Rates of molecular evolution: phylogenetic issues and applications. Annu. Rev. Ecol. Syst. 27, 279–303.
Mishler, B.D., 2005. The logic of the data matrix in phylogenetic analysis. In: Albert, V.A. (Ed.), Parsimony, Phylogeny, and Genomics. Oxford University Press, Oxford, UK, pp. 57–80.
Mooi, R.D., Gill, A.C., 2010. Phylogenies without synapomorphies —a crisis in fish systematics: time to show some character. Zootaxa 2450, 26–40.
Mungall, C.J., Torniai, C., Gkoutos, G.V., Lewis, S.E., Haendel, M.A., 2012. Uberon, an integrative multi-species anatomy ontology. Genome Biol. 13, R5.

Niknejad, A., Comte, A., Parmentier, G., Roux, J., Bastian, F.B., Robinson-Rechavi, M., 2012. vHOG, a multispecies vertebrate ontology of homologous organs groups. Bioinformatics 28, 1017–1020.

Nixon, K.C., 1999. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15, 407–414.

Nixon, K.C., 2008. ASADO, Version 1.85 TNT-MrBayes Slaver Version 2; mxram 200 (vl 5.30). Made available through the author, Cornell University, Ithaca, NY.

Nixon, K.C., Carpenter, J.M., 1993. On outgroups. Cladistics 9, 413–426.

Nixon, K.C., Carpenter, J.M., 1996. On consensus, collapsibility, and clade concordance. Cladistics 12, 305–321.

Nixon, K.C., Carpenter, J.M., 2012. On homology. Cladistics 28, 160–169.

Oberprieler, R.G., Marvaldi, A.E., Anderson, R.S., 2007. Weevils, weevils, weevils everywhere. Zootaxa 1668, 491–520.

O'Brien, C.W., Kovarik, P.W., 2001. The genus *Diaprepes:* its origin and geographical distribution in the Caribbean region. In: Futch, S.H. (Ed.), Diaprepes Short Course. Citrus Research and Education Center, University of Florida Institute of Food and Agricultural Sciences, Lake Alfred, FL, pp. 1–7.

Parmentier, G., Bastian, F.B., Robinson-Rechavi, M., 2010. Homolonto: generating homology relationships by pairwise alignment of ontologies and application to vertebrate anatomy. Bioinformatics 26, 1766–1771.

Patterson, C., Johnson, C.D., 1997. The data, the matrix, and the message: comments on Begle's "Relationships of the Osmeroid Fishes". Syst. Biol. 46, 358–365.

Philippe, H., Delsuc, F., Brinkmann, H., Lartillot, N., 2005. Phylogenomics. Annu. Rev. Ecol. Evol. Syst. 36, 541–562.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., Baurein, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9, e1000602.

Pierce, W.D., 1913. Miscellaneous contributions to the knowledge of the weevils of the families Attelabidae and Brachyrhinidae. Proc. US Nat. Mus. 45, 365–426.

de Pinna, M.C.C., 1991. Concepts and tests of homology in the cladistic paradigm. Cladistics 7, 367–394.

Pogue, M.G., Mickevich, M.F., 1990. Character definitions and character state delineation: the bête noire or phylogenetic inference. Cladistics 6, 319–361.

Pol, D., Escapa, I.H., 2009. Unstable taxa in cladistic analysis: identification and the assessment of relevant characters. Cladistics 25, 515–527.

Proctor, H.C., 1996. Behavioral characters and homoplasy: perception versus practice. In: Sanderson, M.J., Hufford, L. (Eds.), Homoplasy: The Recurrence of Similarity in Evolution. Academic Press, San Diego, CA, pp. 131–149.

Putnam, H., 1974. The "corroboration" of theories. In: Schilpp, P.A. (Ed.), The Philosophy of Karl Popper. Open Court, La Salle, IL, Vol. 1, pp. 221–240.

Ramírez, M.J., 2006. Further problems with the incongruence length difference test: "hypercongruence" effect and multiple comparisons. Cladistics 22, 289–295.

Ramírez, M.J., 2007. Homology as a parsimony problem: a dynamic homology approach for morphological data. Cladistics 23, 588–612.

Ramírez, M.J., Coddington, J.K., Maddison, W.P., Midford, P.E., Prendini, L., Miller, J., Griswold, C.E., Hormiga, G., Sierwald, P., Scharff, N., Benjamin, S.P., Wheeler, W.C., 2007. Linking of digital images to phylogenetic data matrices using a morphological ontology. Syst. Biol. 56, 283–294.

Richards, R., 2003. Character individuation in phylogenetic inference. Philos. Sci. 70, 264–279.

Rieppel, O., 1988. Fundamentals of Comparative Biology. Birkhäuser Verlag, Basel, Switzerland.

Rieppel, O., 2003. Semaphoronts, cladograms and the roots of total evidence. Biol. J. Linn. Soc. 80, 167–186.

Rieppel, O., 2005. Modules, kinds, and homology. J. Exp. Zool. B Mol. Dev. Evol. 304, 18–27.

Rieppel, O., 2007a. The performance of morphological characters in broad-scale phylogenetic analyses. Biol. J. Linn. Soc. 92, 297–308.

Rieppel, O., 2007b. Parsimony, likelihood, and instrumentalism in systematics. Biol. Philos. 22, 141–144.

Rieppel, O., 2009. "Total evidence" in phylogenetic systematics. Biol. Philos. 24, 607–622.

Rieppel, O., Kearney, M., 2002. Similarity. Biol. J. Linn. Soc. 75, 59–82.

Rieppel, O., Kearney, M., 2007. The poverty of taxonomic characters. Biol. Philos. 22, 95–113.

Rodríguez-Ezpeleta, N., Brinkmann, H., Roure, B., Lartillot, N., Lang, B.F., Phillippe, H., 2007. Detecting and overcoming systematic errors in genome-scale phylogenies. Syst. Biol. 56, 389–399.

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

Scotland, R.W., 2010. Deep homology: a view from systematics. BioEssays 32, 438–449.

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

Seltmann, K.C., Yoder, M.J., Mikó, I., Forshage, M., Bertone, M.A., Agosti, D., Austin, A.D., Balhoff, J.P., Borowiec, M.L., Brady, S.G., Broad, G.R., Brothers, D.J., Burks, R.A., Buffington, M.L., Campbell1, H.M., Dew, K.J., Ernst, A.F., Fernández-Triana, J.L., Gates, M.W., Gibson, G.A.P., Jennings, J.T., Johnson, N.F., Karlsson, D., Kawada, R., Krogmann, L., Kula, R.R., Mullins, P.L., Ohl, M., Rasmussen, C., Ronquist, F., Schulmeister, S., Sharkey, M.J., Talamas, E., Tucker, E., Vilhelmsen, L., Ward, P.S., Wharton, R.A., Deans, A.R., 2012. A hymenopterists' guide to the Hymenoptera Anatomy Ontology: utility, clarification, and future directions. J. Hymenopt. Res. 27, 67–88.

Sereno, P.C., 2007. Logical basis for morphological characters in phylogenetics. Cladistics 23, 565–587.

Sereno, P.C., 2009. Comparative cladistics. Cladistics 25, 624–659.

Sharkey, M.J., Leathers, J.W., 2001. Majority does not rule: the trouble with majority-rule consensus trees. Cladistics 17, 282–284.

Smith, B., 2006. From concepts to clinical reality: an essay on the benchmarking of biomedical terminologies. J. Biomed. Inform. 39, 288–298.

Smith, B., Ceusters, W., 2010. Ontological realism: a methodology for coordinated evolution of scientific ontologies. Appl. Ontol. 5, 139–188.

Sober, E., 1988. Reconstructing the Past. Parsimony, Evolution, and Inference. MIT Press, Cambridge, MA.

Song, H., Bucheli, S.R., 2010. Comparison of phylogenetic signal between male genitalia and non-genital characters in insect systematics. Cladistics 26, 23–35.

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

Thompson, R.T., 1992. Observations on the morphology and classification of weevils (Coleoptera, Curculionoidea) with a key to major groups. J. Nat. Hist. 26, 835–891.

Ting, P., 1936. The mouth parts of the coleopterous group Rhynchophora. Microentomology 1, 93–114.

Travillian, R., Malone, J., Pang, C., Hancock, J., Holland, P.W.H., Schofield, P., Parkinson, H., 2011. The Vertebrate Bridging Ontology (VBO). In: Proceedings of the Second International Conference on Biomedical Ontology, July 26–30, 2011. Buffalo, NY, pp. 416–421. Available at http://icbo.buffalo.edu/ICBO-2011_Proceedings.pdf (accessed 5 June 2012).

Vaurie, P., 1961. A review of the Jamaican species of the genus *Exophthalmus* (Coleoptera, Curculionidae, Otiorhynchinae). Am. Mus. Novit. 2062, 1–41.

Vergara-Silva, F., 2009. Pattern cladistics and the "realism-antirealism debate" in the philosophy of biology. Acta. Biotheor. 57, 269–294.

Vogt, L., 2009. The future role of bio-ontologies for developing a general data standard in biology: chance and challenge for zoomorphology. Zoomorphology 28, 201–217.

Vogt, L., Bartolomaeus, T., Giribet, G., 2010. The linguistic problem of morphology: structure versus homology and the standardization of morphological data. Cladistics 26, 301–325.
Wägele, W., 2005. Foundations of Phylogenetic Systematics. Translated from the German Second Edition by Stefen, C., Wägele, J.-W. and Revised by Sinclair, B. Verlag Dr Friedrich Pfeil, München, Germany.
Wenzel, J.W., Carpenter, J.M., 1994. Comparing methods: adaptive traits and tests of adaptation. In: Eggleton, P., Vane-Wright, R.I. (Eds.), Phylogenetics and Ecology. Academic Press, London, UK, pp. 79–101.
Wenzel, J.W., Siddall, M.E., 1999. Noise. Cladistics 15, 51–64.
Wheeler, W.C., 2001. Homology and the optimization of DNA sequence data. Cladistics 17, S3–S11.
Wheeler, Q.D., 2004. Taxonomic triage and the poverty of phylogeny. Philos. Trans. R. Soc. Lond. B Biol. Sci. 359, 571–583.
Wiens, J.J., 2003. Incomplete taxa, incomplete characters, and phylogenetic accuracy: is there a missing data problem? J Vertebr Paleontol 23, 297–310.
Wilkinson, M., 1995. Coping with abundant missing entries in phylogenetic inference using Parsimony. Syst. Biol. 44, 501–514.
Williams, D.M., Ebach, M.C., 2008. Foundations of Systematics and Biogeography. Springer, New York, NY.
Windsor, M.P., 2000. Species, demes, and the Omega Taxonomy: Gilmour and *The New Systematics*. Biol. Philos. 15, 349–388.
Winther, R.G., 2009. Character analysis in cladistics: abstraction, reification, and the search for objectivity. Acta Biotheor. 57, 129–162.
Yoder, M.J., Mikó, I., Seltmann, K.C., Bertone, M.A., Deans, A.R., 2010. A gross anatomy ontology for Hymenoptera. PLoS One 5, e15991.