WILEY-BLACKWELL

Cladistics 26 (2010) 625–642

Cladistics

10.1111/j.1096-0031.2010.00309.x

Coding behavioural data for cladistic analysis: using dynamic homology without parsimony

Hilton F. Japyassú[a,b]* and Fabio de A. Machado[c,d]

[a]*Universidade Federal da Bahia, Instituto de Biologia, Instituto de Biologia, Rua Barão de Geremoabo s/n, Salvador BA, CEP 40170-115, Brazil;* [b]*Universidade de São Paulo, NeC, Instituto de Psicologia, São Paulo, Brazil;* [c]*Museu de Zoologia da Universidade de São Paulo (USP), Avenida Nazaré, 481, Ipiranga, CEP 04263-000 São Paulo, SP, Brazil;* [d]*Programa de pós-graduação em Zoologia da Universidade de São Paulo, São Paulo, Brazil*

Accepted 8 February 2010

Abstract

Many of the controversies around the concept of homology rest on the subjectivity inherent to primary homology propositions. Dynamic homology partially solves this problem, but there has been up to now scant application of it outside of the molecular domain. This is probably because morphological and behavioural characters are rich in properties, connections and qualities, so that there is less space for conflicting character delimitations. Here we present a new method for the direct optimization of behavioural data, a method that relies on the richness of this database to delimit the characters, and on dynamic procedures to establish character state identity. We use between-species congruence in the data matrix and topological stability to choose the best cladogram. We test the methodology using sequences of predatory behaviour in a group of spiders that evolved the highly modified predatory technique of spitting glue onto prey. The cladogram recovered is fully compatible with previous analyses in the literature, and thus the method seems consistent. Besides the advantage of enhanced objectivity in character proposition, the new procedure allows the use of complex, context-dependent behavioural characters in an evolutionary framework, an important step towards the practical integration of the evolutionary and ecological perspectives on diversity.

For a central theme of comparative biology, homology remains controversial (Griffiths, 2007; Love, 2007). Most controversies result from the differing views of process-orientated developmental biologists and pattern-orientated cladists (Brigandt and Griffiths, 2007), and tentative syntheses of these views include some heterodox solutions, such as the proposition of a continuum between terms that are usually conceived as complementary in comparative biology, such as homology and homoplasy (Hall, 2003). Another area of controversy lies in the tension between the construal of homology as a relation of correspondence between organic parts of different species, or as similarity between these parts, a tension that results from a more fundamental issue: the conception of species, respectively, as historical individuals or as an atemporal class of individuals (Kluge, 2003, Ghiselin, 2005).

Many of these controversies rest on the subjectivity inherent to homology propositions. Some objectivity was provided by de Pinna (1991), who suggested that homology propositions are the result of a two-step procedure, i.e. that primary homology propositions become secondary homology propositions if they survive a test of congruence (Patterson, 1982). Because secondary homologies are primary homologies that have passed a test, there is objectivity in the procedure. Nevertheless, subjectivity remains in the first step, that of properly formulating a primary homology hypothesis. Subjectivity persists because all primary homology hypotheses rely on a previous theory of the organisms to be compared, i.e. on a previous agreement of what

*Corresponding author:
E-mail address: japyassu@ufba.br

renders some organic parts comparable with each other, on a previous understanding of what counts as similarity and on an implicit theory of observation (Rieppel, 1994). This would not be a problem if these agreements and theories were not mostly unspecified, thus entering the analysis as diffuse background knowledge (Harris et al., 2003a), making parts of the observation and data production process unrepeatable (Vogt et al., 2009).

There have been attempts to eliminate this residual subjectivity, for example by implementing a similarity test for the proposition of primary homologies (i.e. a test during the character construction procedure; Patterson, 1982; Rieppel and Kearney, 2002). This test would be mainly based on an objective (thus repeatable), detailed description of the morphology of the comparable parts during character delimitation (a technique for description); furthermore, this test would rely on a clear description and discussion of the connectivities (topology, ontogeny) of the comparable parts (with the surrounding parts) that allow the homology proposition. The absence of justified connectivities would lead to the rejection of the primary homology statement.

We certainly agree that the development of an objective technique for the study of structural complexity would be a welcome improvement, and that careful character analysis in morphological systematics is a necessary step to reduce controversies over competing primary homology statements. Nevertheless, the careful study of characters does not seem to be a proper test of a primary homology hypothesis (Kluge, 2003). The topological relations or connectivities between constituent elements of an organic structure provide a way to identify putative homologues, not a test of homologues. This principle does not state that if two organic parts (in different species) have different connectivities they are not homologous, but only that the recognition of homology is not possible in this particular case. Alternative states of the same character can have different topological relations (connections) with surrounding structures, if only because these topological relations themselves can have an evolutionary history (Harris et al., 2003a,b). Also, similarity is not necessarily tested through additional similarity; to say that a syllable in a courtship song is homologous in two species of crickets due to one of its properties (low pitch) does not entail any compromise regarding future observations of other populations of crickets of these same species. If we find new populations of one of these species with high-frequency syllables (thus, dissimilar syllables), this will not count as a refutation of the previous primary homology proposition, simply because that proposition does not predict the absence of high-pitched syllables; instead, the status of that particular species will simply be modified from "low pitch" to "polymorphic" (low + high pitch). Also, if we discover that one of the species varies regarding other properties of that syllable (e.g. the duration of each syllable, the number of syllable repetitions on a song—i.e. special similarity—or the kind of syllables that appear before/after that specific syllable—i.e. connectivity), this will not necessarily refute the previous primary homology proposition; instead, this could result in *new* homology statements regarding these newly described properties. Different but closely related properties of one same organic character can effectively have diverging evolutionary histories. As an example, among cervids the acoustic structure of a vocalization called "hiss" (coded as present/absent) has phylogenetic signal, despite the fact that in distinct species it has distinct functions (context of emission: mating and agonistic functions). Nevertheless, function (e.g. mating call, also coded as present/absent) shows a phylogenetic history that is distinct from the history of the acoustic structure of the signal, and one same function can be obtained through distinct vocalizations in distinct groups of species (Cap et al., 2008). Thus, the context of usage (function) and the acoustic structure of the vocalization (hiss), although seemingly part of one single character, revealed themselves as two independent characters, with independent evolutionary histories.

Dynamic homology

Less controversial approximations to the testability of primary homology hypotheses involve dynamic homology procedures. Under direct optimization, correspondences between organic parts are no longer primary data, but are themselves the result of a phylogenetic analysis, subject to the same optimality criteria as the trees (Wheeler, 1996, Wheeler 2001). Direct optimization has been a successful procedure to deal with molecular data, and recently there have been proposals for its utilization with morphological structures (Agolin and D'Haese, 2009, Ramírez, 2007, Agnarsson et al., 2007) and developmental sequences (Schulmeister and Wheeler, 2004). Although the test of competing hypotheses of primary homology through direct optimization reduces the subjectivity inherent to primary homology propositions, its implementation has been proposed not as a methodology to be applied to all morphological data, but only to situations where there is no agreement over which of the many competing primary homology propositions should be chosen. This dynamic homology approach for morphological structures involves a huge amount of extra comparative work (Ramírez, 2007), and is not likely to be a practical solution for all but the controversial primary homology propositions.

Here we take a similar approach to the test of competing homology propositions, but using a procedure that does not involve extra comparative work. Our procedure is in the direction of reducing the subjectivity of homology propositions via dynamic homology

procedures, but it does not go so deeply in this direction as to dismiss almost the whole character as a proposition. We test this new procedure with a dataset based on the predatory repertoire in a group of haplogyne spiders. In order to make our method clearer, we first make a brief presentation of the problem, as it appears on the current use of behaviour as a phylogenetic character.

Behavioural transitions as phylogenetic characters

Action patterns, or units of behaviour, have now been routinely used as a basis to infer phylogenetic relations (Alexander, 1991, Pinto, 1984; Coddington, 1990, Shultz, 1990; Proctor, 1992; Gwynne, 1995; Crespi et al., 1998; Cap et al., 2008). The logic underlying the use of these stereotyped behaviours for phylogeny estimation is the same as that underlying the use of morphological data: stereotyped behaviours are as typical of species as morphology, and the criteria for hypothesizing behavioural homologies are also similar to the morphological ones (Wenzel, 1992).

This is not the case when we are dealing with the connection, or transition, between two behavioural units comprising a behavioural sequence. Transitions have been used as characters in the literature only rarely (Japyassú et al., 2006; Robillard et al., 2006; Legendre et al., 2008), and use has been quite straightforward: you simply use the behavioural sequence as a character, and score this sequence of two (or more) behavioural units as present in a species if you have observed this sequence in any specimen of the species. Conversely, you score the transition as absent if it has not been observed in the repertoire of the species. But unlike most morphological characters, and even some behavioural units, behavioural transitions (BeTs) are not usually a yes-or-no phenomenon in any individual. The occurrence of a transition in an individual, for example, does not mean that there is a biological ground for it. Its expression can be the result of chance events alone, for example the casual configuration of external stimuli in a particular temporal order. Species-typical, stereotyped behaviour is conceived as being triggered by species-typical external stimuli; different stimuli trigger different action patterns. Thus, if these different external stimuli are temporally arranged, the result is a sequential organization of the observed action patterns, but the true nature of this observed organization cannot be trivially assigned to endogenous biological factors, because its organization in that particular sequence is a result of the configuration of the external stimuli, of the natural, temporal organization of the environment.

A completely different situation occurs when a BeT is mainly the result of an internal organization, i.e. it is hardwired, inbuilt into the nervous system of the individual. This is expected, for example, when a connection between two neural areas (each controlling the expression of one action pattern) has adaptive value, which would lead to its selection in evolutionary time. For example, the adding of an apomorphic vocal chuck (after a plesiomorphic vocal whine) in the courtship song of *Physalaemus* frogs enhanced female attraction. This enhanced attractiveness helped to establish the transition whine–chuck on the repertoire of a whole clade of frogs, in a process probably driven by sexual selection via sensory exploitation of female preference for that transition (a symplesiomorphic female preference—Ryan et al., 1990; Ryan, 1996). If there is a favoured internal connection between two neural areas organizing the expression of two action patterns (in the example, the BeT whine–chuck), then we expect the transition between these two action patterns to occur at a frequency that is higher than the one expected simply by the casual configuration of external stimuli.

The distinction between internal and external organization just exposed bears directly upon our previous discussion about homology propositions. If we take the frog courtship song transition whine–chuck as a character, we have to score it in each species as present or absent. Obviously, we do not want to score it as present if it is merely the result of the casual configuration of external stimuli, because transitions that are exogenously organized are not genetically inherited. We have more reasons to score it as present if it is inbuilt into the nervous system of the species, because inbuilt structures can be inherited, and inherited features inform the phylogeny.

Thus, one of the problems ethologists face is to uncover inherited BeTs. We could simply decide that high-frequency transitions must have been inherited; if they are above-chance events, there must be a reason explaining the strong connectivity between these action patterns, and the reason could certainly be inheritance. Highly stereotyped, fixed action patterns are usually conceived as inherited, and if they are present in all individuals of a population, one usually seeks for its adaptive value. If a transition has a high frequency it could be itself a complex, but stereotyped pattern. Nevertheless this is not always the case, simply because the reason behind above-chance transitions could be the above-chance configuration of external stimuli. Let us consider a hypothetical example. Suppose we have in the repertoire of a male frog a low-frequency transition, the two-syllable song AB, and a high-frequency two-syllable song,- XW (where A, B, X and W are song syllables). In this case, following the reasoning above, we should simply discard AB and keep XW, based on the presupposition that high-frequency transitions are hard-wired, inbuilt mechanisms, while low-frequency transitions are mainly the result of the chance configuration of external stimuli. But let us further suppose that the rare AB song is used by the male as a response to giant females, and that the common XW song is the result of

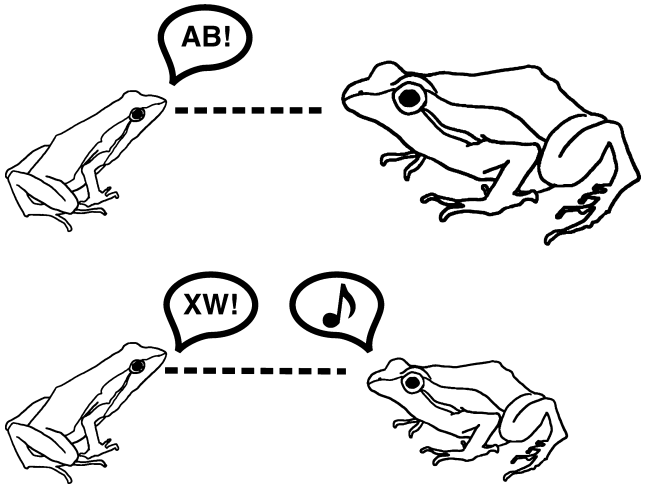

Fig. 1. Hypothetical example showing that the high frequency of a behavioural transition (BeT) should not be taken as the single criterion to score it as present in a species. The rare AB male courtship song is an inbuilt mechanism responsive to rare giant females, while the common (high-frequency) XW song is not inbuilt, but merely the result of the usual and sequential organization of two external triggering stimuli: female receptivity call (which elicits X) followed by visual contact between partners (which elicts W).

the male response to two external stimuli, for example that X occurs as a male response to a hypothetical female receptivity call, and that W occurs as a male response to establishing visual contact with the female (Fig. 1). In this hypothetical case, the AB song could be hardwired in the male, and be rare only because of the rarity of giant females (the single external stimulus that triggers the AB song). Also, the XW song may not be hardwired, and occur frequently in the male repertoire only because the female receptivity call (which triggers X) occurs when the male is nearby, usually right before the visual contact between the male and the female (which triggers W). In this case, the common XW song would simply be a side-effect of the temporal organization of two triggering stimuli (female receptivity call is usually followed by visual contact). In this situation, we would better consider just the opposite of our previous conclusion, i.e. that the rare AB song should be scored as present, and the common XW song as absent, in the repertoire of the male frogs. This counterintuitive result shows that it is not safe to rely only on the frequency of the BeT in order to score it as present or absent in a phylogenetic matrix.

The analytical procedures that have been proposed for the use of behavioural sequences in evolutionary studies (Japyassú et al., 2006; Robillard et al., 2006; Legendre et al., 2008) do not cope with the problem we have just discussed, namely the existence of a sequential organization of external stimuli and the possible noise that it can introduce in the analysis of homology hypotheses. To deal with this problem, we considered that when studying complex behaviour in an evolutionary framework, we have more information than merely the frequency of the BeTs. We also have the between-species congruence of frequency information, something that can be analysed through a kind of direct optimization method. Accordingly, we develop in this paper a new dynamic homology procedure, which we term ''soft Dynamic Homology''.

The rationale for the soft Dynamic Homology procedure

Our approach builds on the realization that homology hypotheses are the result of a three-step procedure (Brower and Schawaroch, 1996). Topographic identity is the first step and involves discovering comparable features among the taxa, i.e. establishing the commonalities between the organic parts of the different species. The next step is establishing character state identity (Fig. 2). In order to make clear their point, Brower and Schawaroch (1996) state that ''a convenient way to conceptualize this distinction is to imagine an empty character matrix; identifying comparable characters by lining up the columns of the matrix is clearly a separate operation from filling in the individual cells''. Together, these two steps represent the primary homology described by de Pinna (1991). The third step is the congruence between the characters in the data matrix in order to obtain the most parsimonious cladogram; in this step synapomorphies are obtained, and some of the homology propositions obtained in the first two steps are partially refuted, becoming homoplasies.

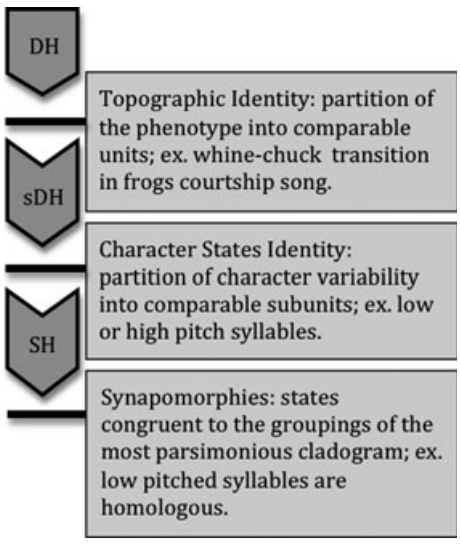

Fig. 2. Three steps in homology assessment: topographic identity, identity of character states, and synapomorphy. Homology can be tested in all three steps: SH, static homology test of congruence; DH, dynamic homology iterative procedure for the simultaneous assessment of synapomorphies and phylogeny; sDH, soft dynamic homology iterative procedure proposed herein.

Homology is usually tested through the construction of the cladogram. If cladogram construction occurs after the topographic and character state identity (first two steps) are fully established, we are dealing with static homology (SH in Fig. 2). If the cladogram is constructed before all three steps, i.e. if the topographic identity of characters is unknown from the beginning, and is obtained after successive parsimonious reconstructions of the phylogeny, we are dealing with dynamic homology (DH).

Our method is positioned between dynamic and static procedures. Because it is similar to a dynamic homology procedure, and because it is not so radical as to dismiss from the start the majority of the putative primary homology propositions, we call it soft Dynamic Homology (sDH). In our proposal, the parsimonious construction of the cladogram occurs after the topographic identity step, but before deciding on character state identity. Thus, under sDH the iterative dynamic homology procedure does not change the topographic identity of character propositions, which is known from the start; instead, it changes character state identity. Building on the previous hypothetical example of the frequency of courtship song syllable transitions (Fig. 1), we could imagine some species with a high frequency of the AB song (transition from syllable A to syllable B), some other species without AB, and still many others showing all the intermediate frequencies. In this case, we are sure about the topographic identity of the character (the character is the transition AB, topographically identical whenever it appears), but we are uncertain about character state identity, as we have all intermediate frequencies, from totally absent (no individual performs AB in the species), to fully present (all individuals perform AB). We have no means to decide whether a species with a low AB frequency should better be scored AB-present or AB-absent; as we have discussed above, some of the observed BeTs could be merely triggered by external, naturally occurring, consecutive stimuli. In this case, where we have no outright reason to consider a specific case as belonging to one or another character state, we should let a dynamic homology procedure decide where each species' data best fit, using congruence with other characters as an arbiter. Thus, in our method we rely on the congruence between characters to define the state (present or absent) of each character in the matrix.

In order to test the sDH procedure we chose a complex, context-dependent behaviour (predatory sequence), in a group of basal Araneomorph spiders. Predatory behaviour is very complex among spiders, and usually comprises a large repertoire of simple action patterns, with responses that are adjustable to prey size and type, to the level of hungriness of the predator, to web characteristics and to previous experience (Japyassú and Viera, 2002; Garcia & Japyassú, 2005; Japyassú & Jotta, 2005; Cross & Jackson, 2006; Japyassú & Caires, 2008). This plasticity is usually expressed as changes in the frequencies of the BeTs. All this plasticity and context sensitivity allows predatory BeTs to be structured by external stimuli (e.g. the behaviour of the prey), making it a perfect test of our methodology.

One could argue that we should better use the frequency information directly in the phylogenetic matrix through one of the methods available for continuous data (e.g. Smith and Gutberlet, 2001; Wiens, 2001; Goloboff et al., 2006). However, much of the behavioural expression in an individual is the result of complex interactions between genetic, neural and environmental networks, a complexity that results in the use of non-linear models to predict behavioural expression (Fitch et al., 2002; Corrado et al., 2005; Freeman and Vitiello, 2006; Patterson et al., 2008). This argues against the additivity of these characters, thus preventing the use of these frequencies as quantitative characters in cladistic analysis (Goloboff et al., 2006). Furthermore, the frequencies of a BeT are thought of as an expression of a neurological trait, namely the connectivity between neural modules (each related to one of the behavioural patterns, Japyassú, 2008), and are used to evaluate the existence of these connectivities; in this way they act not as quantitative characters *per se*, but more as diagnostic features for the underlying neural organization. Although our approach could be considered simplistic (Felsenstein, 1988), as these connections could themselves be a continuum, we prefer not to presuppose continuity from the start, remaining on the qualitative (presence/absence) aspect of these connections, because high levels of precision in coding, especially when applied to complexly structured systems, may be illusory due to the noise of non-heritable variation (Lawing et al., 2008).

Methods

We focused on the predatory behaviour of scytodid spiders. *Scytodes* (Scytodidae) is a genus of mainly tropical and subtropical nocturnal spiders that present a remarkable characteristic: their cephalic glands are greatly specialized and allow them to produce and secrete a glue-like substance through their chelicerae (Nentwig, 1985), ensnaring and envenoming their prey at a safe distance (Li et al. 1999). The spitting mechanisms are highly complex and the evolution of the morphological adaptations and the associated behaviour are not fully understood. The understanding of the phylogenetic structure and signal embedded in this complex behaviour could give insight into the processes leading to such an innovation. As representatives of this genus, we observed the predatory behaviour of three

species: *S. fusca* Walckenaer 1837, *S. globula* Nicolet 1849 and *S. itapevi* Brescovit & Rheims 2000.

Other taxa entering the analysis include *Pholcus phalangioides* (Fuesslin 1775) and *Smeringopus pallidus* (Blackwall 1858) (Pholcidae), and *Nephila clavipes* (Linnaeus, 1767) (Nephilidae). The predatory behaviour of these three non-spitting species was previously studied by us (Japyassú and Macagnan, 2004; M. Silveira and H.F. Japyassú, unpublished data). All the species were studied as part of one research programme, so that all were treated with the same behavioural protocols (see below), with equivalent sampling effort, and were analysed with the same behavioural repertoire (see Appendix).

Pholcid spiders are phylogenetically close to Scytodids (Platnick et al., 1991a; Ramírez, 2000), and both are distant from the orb-weaving nephilids (Coddington and Levi, 1991). Thus, from these previous cladistic analyses, we expect the two pholcid and the three scytodid spider species to form monophyletic groups, with the following family-level relationships: (Nephilidae (Pholcidae (Scytodidae))). If this is indeed the real phylogenetic structure behind these taxa, departures from these predictions imply either the failure of the sDH procedure in recovering the phylogenetic relationships, or some inadequacy in data collecting, or even an excess of noise in the database (excess of homoplasy).

All *Scytodes* individuals were maintained in acrylic boxes with a removable frontal panel to allow the introduction of prey. The frontal panels were sprayed with liquid silicon to decrease silk adherence, thus minimizing any disturbance to the web caused by handling. The spiders were maintained under a 12/12-h inversed photoperiod, with a controlled temperature of 23 ± 2 °C and were starved for 2 weeks or until they became active during the dark period.

Prey was offered approximately 30 min after the opening of the box. To control for alterations in the predatory sequence due to different prey taxa or size (Japyassú and Viera, 2002) we used only crickets (*Gryllus* sp.) that were as large as the spider's body (abdomen + cephalothorax). This is the same procedure as that used in previous studies with nephilids and pholcids (Japyassú and Macagnan, 2004; M. Silveira and H.F. Japyassú, unpublished data). Previous experience with prey items was not controlled. Although it is well known that learning mechanisms contribute to the performance of spiders (Ades, 1982; Punzo, 2002), these mechanisms result in a more efficient and quick capture, but do not alter the sequence of the behavioural units employed in the attack (Herberstein et al., 1998). That is the reason why we focus on the sequence of predatory events and do not consider other measures of behaviour, such as the timing of the units, or the success of the whole performance. The observation of the predatory sequence lasted for 1 h or until the spider returned to its resting place (with or without the prey).

All the observations were filmed with a digital camera (Cannon XL1) and transcribed to text files in the software EthoLog (Ottoni, 2000). For the transcription, we used the behavioural categories described in the Appendix. We obtained a total of 29 successful capture sequences [*S. globula* ($n = 13$); *S. fusca* ($n = 9$); *S. itapevi* ($n = 7$)], and we present a table with the range of variation for the frequencies of first-order transitions as supplementary online material (see Variation of behavioural transitions).

Steps in the sDH procedure

Transition matrices for each species were obtained from the text files with the program EthoSeq (Japyassú et al., 2006). From each of the first-order BeT matrices obtained for each species, we selected a sub-matrix containing only the behavioural units common to all species (13 behavioural units were common to all the six species included in the study). Discarding these transitions implies a reduction of the potentially informative characters, but it is a way to keep track of the most clearly comparable portion of all the transitions, and also a way to avoid the excessive use of non-applicable data, as this could result in phylogenetic resolutions not potentially supportable by any character (Platnick et al., 1991a,b). The list of all the 156 characters and states employed in the analysis is presented as supplementary online material (see List of characters).

Next, the main diagonal of each matrix was set to zero. This is a usual way of computing transitions, and it means that we do not measure self-repetitions: if an animal keeps doing the same thing repeatedly, we consider that nothing new has happened, so that no transition has occurred. There is no doubt that self-repetitions can be informative (Robillard et al., 2006), but we discard them simply because we are focusing on the transitions between alternative behavioural states; long cycles of self-repetitions could enhance the expected values for other transitions, possibly preventing the identification of meaningful connections between alternative behavioural states. These matrices contained a total of 156 valid cells (13×13 behavioural units, minus the 13 cells of the main diagonal). Each cell shows the frequency of a specific BeT, and each BeT is a character for the cladistic analysis (thus, there are 156 characters). These matrices are then joined in a single matrix and submitted to the iterative part of the analysis, in order to decide the states that these characters should assume: present or absent. This full matrix with the untransformed raw data on the frequency of BeTs for all the species is presented as a supplementary online material (see Frequency of behavioural transitions).

We are interested in unravelling an evolutionarily sensible threshold for separating strong and weak

associations between preceding and following acts in a BeT. We expect that, if selection has favoured a BeT at a specific point in the phylogeny, i.e. if a BeT appears at a node of the phylogeny, the association between preceding and following acts appears or becomes stronger from that node up in the topology, with the reverse being also true for the disappearance of an unfavoured BeT. Thus, we should inspect BeT frequencies for changes between species. Nevertheless, BeT frequencies also reflect non-causal associations; for example, the transition from A to B can be more frequent in a species simply because the overall frequency of A and of B have increased, independently of each other, in the whole repertoire (without any increase in the neural connection between them). A usual solution to this problem is to calculate expected values for the transitions and use the residuals between observed and expected values as a measure of the connection between the two behavioural units. The association between the preceding and the following act of a transition gets stronger as the residuals get larger (van Hooff, 1982). Also, different BeTs can have different scales of variation, a situation that could possibly cause different characters to be treated differently by the algorithm. In order to reduce this problem, we standardized cell frequencies computing the adjusted residuals (Haberman, 1973) for each cell of the complete sample. This resulted in the main phylogenetic matrix (Table 1).

The adjusted residuals, main matrix (Table 1), was next employed for building successive phylogenetic matrices (sDH matrices). In order to build each sDH matrix, we ranked all the residuals from larger to smaller before proceeding to the iterative procedure. The ranked residuals were separated in 100 blocks, from the larger to the smaller, each block comprising 1% of the data; the lower residual on each block provided a threshold residual for each step in the iterative procedure.

In the first step, only those cells with residuals equal to or higher than the first, highest block threshold were inspected for presence or absence in the original frequency data matrix; the remaining transitions were scored as absent. As a result, in this first step, only those cells in which the observed frequencies deviated most from the expected ones (the biggest positive residuals) were considered as present in the phylogenetic matrix.

In the second step, the threshold residual was lowered, and the cells with residuals equal to or higher than the first two blocks were inspected for presence or absence in the original frequency data matrix. Thus, this step is a repetition of the preceding one, differing only in the threshold between strong and weak association, which has been lowered 1% more. This step resulted in a second phylogenetic matrix for cladistic analysis.

This procedure was iterated until all the cells were considered for inspection (the threshold for inspection was the lowest possible), which means simply that in the 100th step all BeTs that occurred (either with high or low frequencies) were scored as present. Thus, as we step forward in the procedure, we get less rigorous on each occasion as to whether that a BeT is present in a taxon.

Each matrix produced by this algorithm represents a set of primary homology propositions. Character delimitation (topographic identity) is the same for all the matrices; it is only the delimitation of the states within each character (character state identity, Fig. 2) that change from one matrix to another. All the matrices were analysed with the default TNT (Goloboff et al., 2008a) traditional search parameters (traditional search, random seed = 1, repls. = 10, multiple TBR with 10 trees per replication). An example of these individual matrices is offered as supplementary online material (see sDH example matrix).

Under usual dynamic homology procedures, the most-parsimonious reconstruction is the one to be chosen among the competing set of homology propositions. This solution could be possible for our data if there were no differences in the number of characters between the competing matrices. But in our analysis the number of characters changes dramatically along the procedure, so the length of the most-parsimonious tree is not a useful guide (the number of steps in any tree always increases with the number of characters in a matrix). For example, if one is very stringent as to consider a BeT as present, i.e. if one chooses a very high threshold for the residuals, most characters will be scored as absent in all species, thus reducing the effective number of characters available for building the phylogeny. However, if one lowers the threshold, some of the characters will now be scored as present on at least one of the species, thus increasing the effective number of characters. So, our iterative procedure cannot rely on tree length, as is usually the case under dynamic homology, and we must find another arbiter for choosing between the alternative homology propositions.

Under standard dynamic homology procedures, we usually do not have any theoretical reason to prefer or to reject any of the available homology propositions being tested through parsimony. This is the very reason for using parsimony to decide between the alternatives. Fortunately, this is not exactly the case in our sDH procedure, and this is part of our solution to the problem of finding an arbiter to choose the best topology.

It is clear to behavioural scientists that behavioural sequences are rarely the result of deterministic processes. Instead, behavioural sequences are by and large the result of stochastic processes. This means that an endogenous behavioural organization will only rarely result in a total linkage between one act and another (i.e., only rarely behaviour A will always preceed B). This is because there are random factors affecting the

Table 1
Main phylogenetic matrix for the sDH procedure. The numbers are the standardized residuals of each behavioural transition in each species

Behavioral transition		Species					
		Nephila	*Pholcus*	*Smeringopus*	*Itapevi*	*Fusca*	Globula
0	Approach_Cut-thread	−1.5588684	−1.3150538	−0.2814263	−1.2057115	−1.454315	−2.6741332
1	Approach_Detection	11.9918906	−1.2540381	−0.4372813	17.4387259	18.1226253	21.2653192
2	Approach_Wrap	−1.2874126	−1.8775499	−0.2999845	−1.3942189	−1.4796324	−2.8408943
3	Approach_Pay-out-line	−0.2404157	−0.5757828	−0.0479463	−0.9176754	−1.0942255	−0.9177264
4	Approach_Fix	−0.6167656	−0.5985408	−0.2492287	−0.6826331	−1.6885867	−0.3570253
5	Approach_Fix-prey	−1.0844279	−0.5996604	−0.0479463	−1.2161307	−1.2676251	−1.3056937
6	Approach_Grooming	−1.1131919	−0.8717036	−0.0338837	−0.9805945	−2.1442451	−2.7304117
7	Approach_Manipulation	4.90798029	−0.8377545	−0.1608775	−2.4867382	−4.0133328	−4.5158678
8	Approach_Bite	−2.2910981	−0.9673252	−0.2644486	−1.2873743	−2.2756329	−2.0579533
9	Approach_Pause	−0.2946135	−0.6664882	−0.2192585	−0.5364566	−1.3312378	1.84783673
10	Approach_Touch	8.19515079	0.67365102	−0.1277369	5.60443379	2.45349435	0.92899653
11	Approach_Transport	−0.7680679	−1.0268547	−0.2221156	−0.2960574	−0.2799939	−1.4941565
12	Cut-thread_Approach	−1.4598247	−1.3150538	−0.2814263	−1.2359613	−1.8637714	−2.5558754
13	Cut-thread_Detection	−0.7554535	−2.1023445	−3.6319083	−0.3306069	−1.6426343	−2.5897281
14	Cut-thread_Wrap	6.66827245	15.8296421	6.94914038	6.73742328	3.61847775	11.6220999
15	Cut-thread_Pay-out-line	−0.4351801	−0.9652767	−0.3982252	−0.2145115	0.11198419	−1.0209346
16	Cut-thread_Fix	−0.1215678	−2.3424113	6.78840864	−0.3939363	−1.422008	−0.6600743
17	Cut-thread_Fix-prey	13.7797916	1.17045469	2.31779527	2.8474483	1.72095075	0.10296437
18	Cut-thread_Grooming	−2.015005	−1.461376	−0.2814263	−1.0016281	−1.3427402	1.3798941
19	Cut-thread_Manipulation	0.76932392	−1.4044618	−1.3361937	−0.4267577	5.65255932	1.20556031
20	Cut-thread_Bite	−3.3959934	−1.6216819	−2.196419	−2.7580704	−2.052349	−2.2893923
21	Cut-thread_Pause	−0.5332842	−1.1173408	−1.8210855	−1.1493046	−1.2006174	−0.9120613
22	Cut-thread_Touch	−1.2795858	−1.3958088	−1.0609387	−1.8922211	1.17169457	−2.4614677
23	Cut-thread_Transport	0.22027989	0.24209553	7.97545558	−0.6342732	−0.252521	−1.6621905
24	Detection_Approach	21.2048781	10.6450019	2.18790774	25.336963	19.0758781	21.8772262
25	Detection_Cut-thread	−1.2005903	−2.493405	−3.8005174	−1.1728803	−1.2988046	−2.657725
26	Detection_Wrap	−0.9915238	−3.5599246	−4.0511375	−1.3562547	−1.6931816	−3.2478854
27	Detection_Pay-out-line	−0.1851605	−1.0917117	−0.6474898	−0.8926873	−1.2521505	−1.0492013
28	Detection_Fix	−0.475013	−1.7487941	−2.6026203	−1.4198	0.75468772	−0.7388024
29	Detection_Fix-prey	−0.8351915	−1.1369848	−0.6474898	−1.1830158	−1.4505762	−1.4927495
30	Detection_Grooming	0.38615358	−1.6527917	−0.4575818	−0.9538932	−1.3365872	−1.943149
31	Detection_Manipulation	−1.3325894	−1.5884227	−2.1725689	−2.419025	−4.5925603	−5.1628183
32	Detection_Bite	−2.4577522	−1.834095	−3.5712425	−1.2523194	−2.6040655	−2.352779
33	Detection_Pause	−0.2269018	−1.2636936	1.74668011	1.46296142	1.45507424	2.64537016
34	Detection_Touch	−0.5444383	−1.711301	1.12444616	−0.3677446	−0.5149576	−0.1856347
35	Detection_Transport	−0.5915412	−1.9469656	−2.999558	−0.2879958	−0.3204042	6.32782427
36	Wrap_Approach	−1.205616	−1.8775499	−0.2999845	−1.2275497	−1.39386	−2.6615888
37	Wrap_Cut-thread	15.1669851	14.6435622	9.78135305	1.57759793	1.87385647	10.7375575
38	Wrap_Detection	−0.6239015	−3.0015934	−3.8714098	−1.2275497	−1.6284596	−3.1197384
39	Wrap_Pay-out-line	−0.3593994	−0.5270644	−0.4244857	13.3123058	14.9977872	8.46026274
40	Wrap_Fix	0.24954624	−1.4667705	2.51477469	−0.7499424	0.75251324	−1.3677779
41	Wrap_Fix-prey	4.5585332	5.92855062	−0.4244857	2.43881064	0.38791691	−0.0071182
42	Wrap_Grooming	−1.6641193	−1.513787	−0.2999845	−0.4284834	−1.7563849	0.72358694
43	Wrap_Manipulation	−0.7488815	4.53318641	4.85559276	0.73217238	0.19005873	2.16775703
44	Wrap_Bite	−3.2131802	−0.7554905	−1.3451971	−2.7392999	−1.5348917	−2.3840836
45	Wrap_Pause	−0.44042	−1.5952679	−0.175919	−0.1646369	−0.8979066	−0.949785
46	Wrap_Touch	−1.0567633	1.0980459	2.78686292	−2.7392999	−1.101702	−2.5632763
47	Wrap_Transport	−1.1481908	7.38769499	−1.9664702	−0.6299566	−0.1888531	−1.7309402
48	Pay-out-line_Approach	−0.2251408	−0.5757828	−0.0479463	−0.8019715	−1.0652921	−0.8710051
49	Pay-out-line_Cut-thread	−0.4351801	−0.9652767	−0.3982252	−1.6760925	−1.0427754	−1.0137917
50	Pay-out-line_Detection	−0.1165095	−0.9204899	−0.618764	−0.8019715	−1.2445906	−1.0209346
51	Pay-out-line_Wrap	−0.3593994	−1.3781605	−0.4244857	−1.9381419	−0.7627442	−1.0770126
52	Pay-out-line_Fix	−0.1721788	10.7972962	5.68412818	7.66108173	8.51455705	5.93128555
53	Pay-out-line_Fix-prey	−0.3027333	−0.4401631	−0.0678452	−1.6905765	−0.6534554	−0.4950021
54	Pay-out-line_Grooming	−0.3107632	−0.6398485	−0.0479463	−1.3631512	0.34634628	−1.0351275
55	Pay-out-line_Manipulation	−0.483026	−0.6149292	−0.2276458	−3.059386	−2.068856	−1.7120125
56	Pay-out-line_Bite	−0.8908657	−0.7100367	−0.3742014	−1.7896142	−1.1730791	−0.7801915
57	Pay-out-line_Pause	−0.0822455	−0.4892162	−0.3102562	−0.745743	−0.6862474	−0.3108172
58	Pay-out-line_Touch	−0.1973435	−0.8645105	−0.1807509	−1.7896142	−1.3423607	0.45144429
59	Pay-out-line_Transport	−0.214417	−0.7537326	−0.3142992	−0.4115574	−0.1443357	−0.5664503

Table 1
(Continued)

Behavioral transition		Species					
		Nephila	*Pholcus*	*Smeringopus*	*Itapevi*	*Fusca*	Globula
60	Fix_Approach	−0.577579	−0.5985408	−0.2492287	−1.2359613	−0.9788361	−0.1639233
61	Fix_Cut-thread	−1.1164164	−1.3438487	1.25189823	0.16356649	−1.4779202	−1.8378695
62	Fix_Detection	−0.2988948	−0.6731269	−3.2163872	−1.2359613	1.44348552	−1.2314386
63	Fix_Wrap	−0.9220077	−2.2178005	−1.6819246	−2.5817912	−1.4935814	−1.3591621
64	Fix_Pay-out-line	5.73487828	8.64767858	5.68412818	-1.9660275	−0.1079149	−0.6307319
65	Fix_Fix-prey	0.58975309	−1.0681321	−0.3526649	−2.6054382	−1.2795755	0.2888419
66	Fix_Grooming	−0.7972358	−0.8295025	−0.2492287	0.0975692	5.58134926	2.41084337
67	Fix_Manipulation	−1.239161	−1.4922322	−1.1833218	3.1127306	−1.3601022	2.07862927
68	Fix_Bite	4.53265853	−1.7230272	−1.9451301	−2.3251458	−2.297086	−1.4143834
69	Fix_Pause	−0.2109937	−1.1871678	−0.232424	−0.1779075	0.30961882	−0.5634702
70	Fix_Touch	−0.5062676	1.19992637	−0.9395583	6.33334682	−0.3480354	0.67221469
71	Fix_Transport	−0.5500681	−0.5857195	12.6903044	2.8503143	−0.2826335	1.06145806
72	Fix-prey_Approach	−1.0155281	−0.5996604	−0.0479463	−1.0407876	−1.2244325	−1.2392211
73	Fix-prey_Cut-thread	10.8644711	1.17045469	2.31779527	2.05865957	1.54448992	3.25088132
74	Fix-prey_Detection	−0.5255318	−0.9586624	−0.618764	−1.0407876	−1.4305159	−1.452533
75	Fix-prey_Wrap	1.12539269	−1.4353126	2.1489587	4.97953464	2.76458235	2.19279985
76	Fix-prey_Pay-out-line	3.15929378	−0.4401631	−0.0678452	2.39424004	2.58304062	−0.4950021
77	Fix-prey_Fix	−0.7766359	−1.0681321	−0.3526649	−2.1817771	−0.4664286	−0.9118467
78	Fix-prey_Grooming	−1.4017397	−0.6663828	−0.0479463	0.13707149	0.67536826	−1.472726
79	Fix-prey_Manipulation	−1.6429263	0.9951299	−0.2276458	3.06808122	0.89406286	2.44473691
80	Fix-prey_Bite	−4.0183711	0.69121728	−0.3742014	−2.3225367	−1.3483214	−1.1100162
81	Fix-prey_Pause	−0.3709795	−0.5095039	−0.3102562	−0.9678151	−0.7887636	−0.4422147
82	Fix-prey_Touch	−0.8901448	−0.9003615	−0.1807509	−2.3225367	0.67536826	−1.1934473
83	Fix-prey_Transport	−0.967157	11.4060147	−0.3142992	−0.5341136	−0.1658975	−0.8059162
84	Grooming_Approach	−1.0155281	−0.8717036	−0.0338837	−0.8800941	−1.4385703	−2.6266403
85	Grooming_Cut-thread	−1.9629389	−0.699956	−0.2814263	0.00275974	−0.3711768	−0.2606724
86	Grooming_Detection	1.47777738	−1.3935712	−0.4372813	−0.8800941	−1.537661	−2.2843123
87	Grooming_Wrap	−1.6211197	−0.3684422	3.33732798	−1.0399607	−1.7312739	−0.5842685
88	Grooming_Pay-out-line	−0.3027333	−0.6398485	−0.0479463	−1.3999543	−1.2803207	−1.0492013
89	Grooming_Fix	−0.7766359	−0.8295025	−0.2492287	0.91558811	2.48459702	1.64907473
90	Grooming_Fix-prey	−1.3655198	−0.6663828	−0.0479463	−0.6362265	−0.7172103	−1.4927495
91	Grooming_Manipulation	−0.5712698	−0.930969	−0.1608775	2.78007974	1.55303729	7.57758986
92	Grooming_Bite	−4.0183711	−1.0749567	−0.2644486	−1.9639463	−2.6626504	−2.352779
93	Grooming_Pause	2.45734313	−0.7406464	−0.2192585	3.0905504	5.75544899	0.25691428
94	Grooming_Touch	8.69065937	−0.4706102	−0.1277369	0.93956237	2.60177686	0.75195894
95	Grooming_Transport	0.13887382	−1.1411098	−0.2221156	−0.4516486	−0.3276125	−1.0385421
96	Manipulation_Approach	−1.6105828	−0.8377545	−0.1608775	−2.211228	−3.8683422	−4.3005089
97	Manipulation_Cut-thread	0.80684856	−1.4044618	−1.3361937	5.82090831	5.18716892	2.38482179
98	Manipulation_Detection	−0.8334703	−1.3392977	−2.0761834	−2.211228	−4.51942	−5.0407723
99	Manipulation_Wrap	0.66035171	−2.0052009	0.93065528	3.35491818	4.37764312	3.3457337
100	Manipulation_Pay-out-line	−0.4801217	−0.6149292	−0.2276458	−2.7339732	−2.0482742	−0.9569436
101	Manipulation_Fix	−1.2317103	−1.4922322	−1.1833218	−1.9271625	−3.1492075	−1.4654793
102	Manipulation_Fix-prey	−1.088863	−0.6404301	−0.2276458	4.78943681	4.2402586	5.13373854
103	Manipulation_Grooming	−1.1716643	−0.930969	−0.1608775	−1.1774424	0.64082976	2.99396419
104	Manipulation_Bite	9.99040029	21.8326736	10.1077576	15.1066934	15.5266394	12.4992385
105	Manipulation_Pause	3.21464747	−0.7118015	−1.041025	−2.0561927	−1.4396661	−1.5346318
106	Manipulation_Touch	−0.6065747	−1.2578485	1.11254824	−4.6439457	−4.2938922	−0.4727129
107	Manipulation_Transport	−1.5338684	−1.0966685	−1.0545907	0.03414795	−0.5241189	−1.8438811
108	Bite_Approach	−2.8650602	−0.5996604	−0.2305143	−1.1054649	−2.0072898	−1.8876167
109	Bite_Cut-thread	−2.8424045	−1.0053065	2.24024073	−1.8078806	−1.9648624	−1.667512
110	Bite_Detection	−1.4826574	−0.9586624	−2.9748716	−1.1054649	−2.3451353	−2.2125396
111	Bite_Wrap	3.36221956	5.92855062	0.20872094	−2.6716011	0.11040298	−2.3340702
112	Bite_Pay-out-line	−0.8540868	−0.4401631	−0.3261836	−1.7584487	−1.0628532	−0.7540012
113	Bite_Fix	4.75750179	−0.0348575	−1.6955296	−1.5110674	0.00549399	−1.3889508
114	Bite_Fix-prey	−3.8524749	1.79297856	−0.3261836	−1.3327486	−0.3363222	−1.072754
115	Bite_Grooming	−3.5931562	−0.6663828	−0.2305143	2.34402288	−0.6437497	3.48403568
116	Bite_Manipulation	2.27554889	0.9951299	3.84081107	8.08620591	10.3559631	4.73346397
117	Bite_Pause	−1.0466266	−0.5095039	−1.4916389	1.1046352	−1.2930702	−0.6735939
118	Bite_Touch	−2.5113224	0.29723474	0.36255679	−2.4668654	−2.0579573	−0.5751152
119	Bite_Transport	−0.1719702	−0.7849897	1.41391196	−0.5673049	−0.2719662	0.547696

Table 1
(Continued)

Behavioral transition		Species					
		Nephila	*Pholcus*	*Smeringopus*	*Itapevi*	*Fusca*	Globula
120	Pause_Approach	−0.275895	−0.6664882	−0.2192585	−0.4755635	−1.0987926	−0.7781204
121	Pause_Cut-thread	−0.5332842	−1.1173408	−1.8210855	−0.9939111	−1.0755678	−0.9056802
122	Pause_Detection	−0.1427746	5.0759166	9.96275514	3.86547742	2.2032176	5.19239963
123	Pause_Wrap	−0.44042	−1.5952679	0.41249961	−0.1779075	−0.7867304	−0.9621591
124	Pause_Pay-out-line	−0.0822455	−0.4892162	−0.3102562	−0.7564726	−0.581807	−0.3108172
125	Pause_Fix	−0.2109937	−1.1871678	−0.922581	−1.2031535	0.69736702	1.26223123
126	Pause_Fix-prey	−0.3709795	−0.5095039	−0.3102562	−1.0025	0.88215328	−0.4422147
127	Pause_Grooming	2.38089669	−0.7406464	−0.2192585	0.50928326	0.25477838	−0.9247407
128	Pause_Manipulation	3.19295322	−0.7118015	−1.041025	0.56106587	−0.9248766	−1.5294421
129	Pause_Bite	−1.0916967	−0.8218918	−1.7112244	−1.0612287	−1.2099692	−0.6969913
130	Pause_Touch	−0.2418314	−1.0007004	−0.8265746	1.01458126	2.71380744	−0.7493786
131	Pause_Transport	−0.2627538	−0.8724712	−1.4372916	−0.2440507	−0.1488746	1.55219134
132	Touch_Approach	−0.6619948	−1.177774	−0.1277369	−0.1739039	−2.4003432	−1.5018362
133	Touch_Cut-thread	−1.2795858	−1.3958088	−1.0609387	−2.3851614	0.15983728	−1.4155604
134	Touch_Detection	−0.3425797	−1.8828785	−1.648491	−1.1412445	−1.0777476	−1.4380367
135	Touch_Wrap	−1.0567633	2.40374518	1.80742197	−1.8922211	−1.7186347	−2.5458057
136	Touch_Pay-out-line	−0.1973435	−0.8645105	−0.1807509	−1.8153628	−1.2709737	−0.8224005
137	Touch_Fix	−0.5062676	−1.5482494	0.20922937	5.0387058	−1.6437608	0.68341643
138	Touch_Fix-prey	−0.8901448	0.29723474	−0.1807509	−1.9202541	−0.7018355	−0.2362215
139	Touch_Grooming	6.10262828	−1.3088214	−0.1277369	0.99630758	2.6575481	−1.4824422
140	Touch_Manipulation	4.18894339	3.96222292	−0.6064865	1.77160635	2.82174489	5.47261962
141	Touch_Bite	−2.6194656	2.35280304	2.27496103	−2.5467082	−2.6432115	−1.8441901
142	Touch_Pause	−0.2418314	−1.0007004	−0.8265746	−1.0612287	0.66067904	−0.7346991
143	Touch_Transport	−0.630463	−1.5417734	−0.8373458	−0.5856663	−0.3252208	−0.5169264
144	Transport_Approach	−0.7192683	−1.0268547	−0.2221156	−0.2935959	−0.3732553	−1.2392211
145	Transport_Cut-thread	−1.3902911	0.24209553	−1.2310493	−0.6136052	−0.365366	−1.4423706
146	Transport_Detection	−0.3722185	−1.6416075	−2.8664833	−0.2935959	−0.4360776	8.6582001
147	Transport_Wrap	−1.1481908	−0.9809929	−1.3849226	0.84969553	−0.2672491	−1.532318
148	Transport_Pay-out-line	−0.214417	0.65527201	−0.3142992	−0.4670192	−0.1976374	−0.4950021
149	Transport_Fix	−0.5500681	−1.2073911	−1.6337536	5.2670136	2.30666149	3.76344021
150	Transport_Fix-prey	−0.967157	−0.7849897	−0.3142992	−0.6189077	−0.2289566	−0.7042634
151	Transport_Grooming	0.0871736	−1.1411098	−0.2221156	−0.4990397	−0.4703342	0.06545449
152	Transport_Manipulation	−1.5431469	−1.0966685	−1.0545907	−1.2655395	−0.7248824	−2.435763
153	Transport_Bite	−0.3374813	−1.266284	7.97999397	−0.6551647	−0.4110215	−1.1100162
154	Transport_Pause	−0.2627538	−0.8724712	−1.4372916	−0.2730111	−0.2404463	1.89599537
155	Transport_Touch	−0.630463	−1.5417734	−0.8373458	−0.6551647	−0.4703342	−1.1934473

transitions; there is noise in the system. The noise also means that some of the transitions do not result from any endogenous organization (Haccou and Meelis, 1995; Lehner, 1996).

Due to this structural uncertainty inbuilt into behavioural processes, we should not be too stringent when considering a transition as present in a species; we should not expect that all individuals in a species perform it in order to say that this transition is present in that taxon—i.e. we should not use very high thresholds in our stepwise procedure. High thresholds are expected to result in false negatives, i.e. they are expected to score a transition as absent even when it is the result of an evolutionary process.

For the same reason, the converse is also true. We should not be too permissive when considering a transition as present in a species; its mere occurrence in an individual of a whole population does not render it a species typical behaviour. There is noise and uncertainty inbuilt into the behavioural systems, and one should not consider extremely low frequency transitions as evolutionarily inbuilt connections. This means that one should expect that, in our iterative procedure, extremely low thresholds provide many false positives, that is, they are expected to score many transitions as present even when they are not the result of any evolutionary process.

Thus, we have strong theoretical and empirical reasons to discard the extreme, i.e. the very high and the very low thresholds, on the basis that they will provide too frequently false homology hypotheses (due to the inherent stochasticity of behavioural systems—Lehner, 1996; Haccou and Meelis, 1995). One line of reasoning states that one should not care about these false homology hypotheses, because they will result in homoplasy, and homoplasy does not correlate with phylogeny; but the point here is that if most or all of the homology propositions are erroneously delimited,

there will not be enough phylogenetic signal in the matrix to produce a reliable phylogeny. We therefore expect that our iterative procedure will produce multiple, almost random phylogenies at the extreme thresholds, and single or at least stable solutions under midway thresholds, whereupon erroneous character delimitation is less frequent. If this expectation holds true, we will have a criterion to choose among the alternative phylogenies produced during the iterative procedure. If there is phylogenetic signal in the data, we will find preferred midway stable topologies between non-stable solutions.

Results

Dynamic homology

The sDH procedure was successful in identifying a preferred cladogram. As predicted, building phylogenies only with BeTs that occur at a frequency greatly above or greatly below the expected (highest and lowest residual thresholds) resulted in unstable relationships between species, while midway thresholds produced stable topologies.

At the very beginning of the iterative procedure, the use of unreasonably high frequency thresholds to consider a BeT as present in a species set most of the putative 156 characters to absence in all terminals, so that the number of characters was insufficient to specify the phylogenetic relationships between taxa. For example, the first step (1% threshold) resulted in only three phylogenetically informative characters. We analysed cladistically only the matrices with at least six informative characters. As we lowered the threshold in the following steps of the iterative procedure, i.e. as we got less rigorous when considering a BeT as present in a taxon, we soon got a sufficient number of characters to specify the topology of the tree. The general trend for any character, as we lower the threshold, is to enter the analysis at the tips of the phylogeny (as an autapomorphy), and then migrate to its inner portion in the following steps, turning into a synapomorphy and, sometimes, later in the procedure, into a symplesiomorphy.

As we proceeded in the iterative procedure, there was an almost constant influx of characters (approximately five per step), until nearly 100 were included in the analysis (Fig. 3). Throughout this procedure, the level of phylogenetic information varied strongly. The proportion of phylogenetically informative characters (in relation to the total number of characters in the matrix) showed a triphasic pattern, from the first to the last matrix analysed; it grew almost linearly, then stabilized for a long series of successive matrices, and finally grew again in a nearly linear trend (Fig. 3).

As predicted, the trees specified by the initial matrices changed from step to step (Fig. 4a). For example, consider the result of some of the initial steps: using the 7% threshold (34 characters, eight of them phylogenetically informative) we got one single tree, with pholcids and scytodids as monophyletic groups; using the 10% threshold (45 characters, 16 informative), pholcids were no longer monophyletic; using the 15% threshold (61 characters, 27 informative) both groups were again monophyletic, but the internal relationship is changed in *Scytodes*; add to this picture the next 1% highest BeTs, and we already have two most-parsimonious trees.

Clade instability eventually peaked, and one tree was stable as we stepped from the 18 to 52% threshold

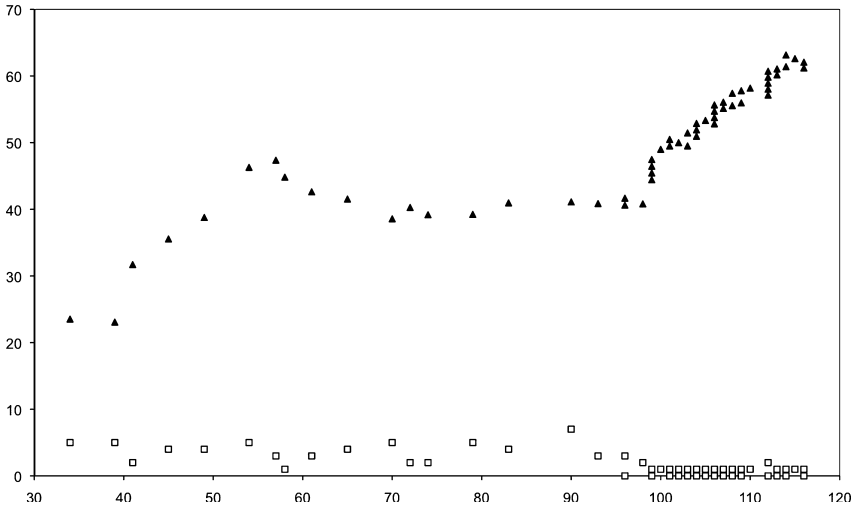

Fig. 3. Proportion of phylogenetically informative characters (black triangles), and number of new characters per step (white squares), as a function of the total number of characters per step along the iterative procedure.

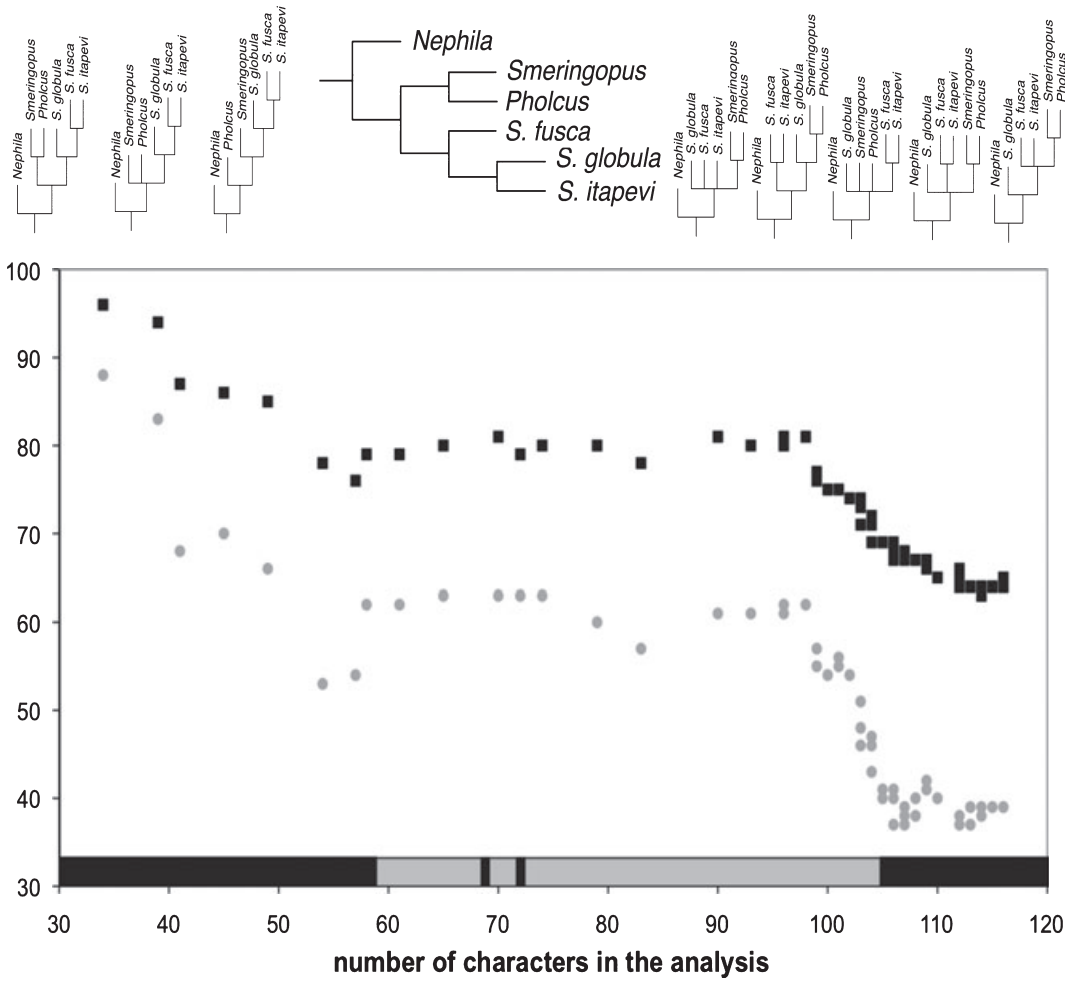

Fig. 4. Triphasic decrease of the consistency and retention indices throughout the steps of the dynamic homology procedure (graphic below, d), and the phylogenies resulting from the first unstable phase (left, a), the second stable phase (central, b) and the last, unstable phase (right, c). The consistency (black squares) and retention indices (grey circles) decreased progressively until nearly 55 characters were included in the analysis. From there up to nearly 100 characters, these indices showed little variation, but decreased again thereafter. The bar at the bottom shows the cladistic stability phase (grey) between the cladistic instability phases (black). At the beginning of the stable phase, two spikes of instability still occur, but both produce simply a tree less resolved than the stable one.

(Fig. 4b). During all these steps clade stability was disrupted only at two points, which nevertheless showed a cladogram [*Nephila* ((*Pholcus*, *Smeringopus*) (*Globula*, *Fusca*, *Itapevi*))] only less resolved than the stable one [*Nephila* ((*Pholcus*, *Smeringopus*) (*Globula* (*Fusca*, *Itapevi*)))]. After this long stability phase, we had again changing topologies from step to step, and multiple equally parsimonious trees for most steps, a situation that persisted until the end of the procedure, when all behavioural transitions that occurred were considered as present (regardless of its high or low frequency in the raw data, Fig. 4c). Thus, we have again a general triphasic pattern: a large phase of cladistic stability between two phases of cladistic instability. The cladogram obtained in the large stability phase is referred to as the stable cladogram.

As we stepped through the procedure, the level of homoplasy and homology present in the most-parsimonious cladograms (as measured by the ensemble consistency and retention indices) also showed a triphasic pattern. First, the indices decreased, they then showed an intermediate phase of stability and, finally, a new decreasing trend (Fig. 4d).

The stability phases described above are clearly correlated. The stable phase of the phylogenetic information (Fig. 3) corresponds to the cladistic stability phase, i.e. stability of topological relationships between taxa (Fig. 4b). It also corresponds to the stability in the consistency and retention indices (Fig. 4d). This stable phase is surprising as it is also, in some respects, the more variable of the three phases of the iterative procedure. It is in this cladistically stable phase, for

example, that most of the changes are introduced to the matrices. From the beginning to the end of the iterative procedure there is a total of 350 character state changes in the successive matrices, and most of them (51%) are introduced during this cladistic stability phase. In the stability phase, most of the between-matrices character-state changes lead to new grouping evidence (109 new grouping information entered the analysis); others disrupt previous grouping information (15 previously informative characters become symplesiomorphic during the stepwise changes of the stable phase). Despite this strong turbulence in the input data, the topological relationships between taxa remain unaltered for as much as 35 steps of the iterative procedure. This means that the phylogenetic signal for these relationships is indeed very robust in the dataset.

There are other, smaller areas of relative clade stability along the procedure. This becomes clear when we plot the number of node changes in the trees (NC) obtained in successive iterations, as a function of the steps in the sDH procedure (Fig. 5). The NC is simply the number of topological alterations in the cladograms during the successive iterations, and it is a measure of instability. We plot four NC levels. The bottom graph shows that there are various stable areas ($NC_1 = 0$) during the sDH procedure (areas with no topological change), if we use only one iteration (NC_1) to count the changes. At stronger criteria for stability (higher NC levels, NC_5 to NC_{15}), we have progressively fewer areas of stability. When we sum the topological changes over 15 successive iterations (NC_{15}), we have only one area of clade stability. Thus, at more rigorous stability criteria there is only one stable solution: the one we adopt here (Fig. 4b).

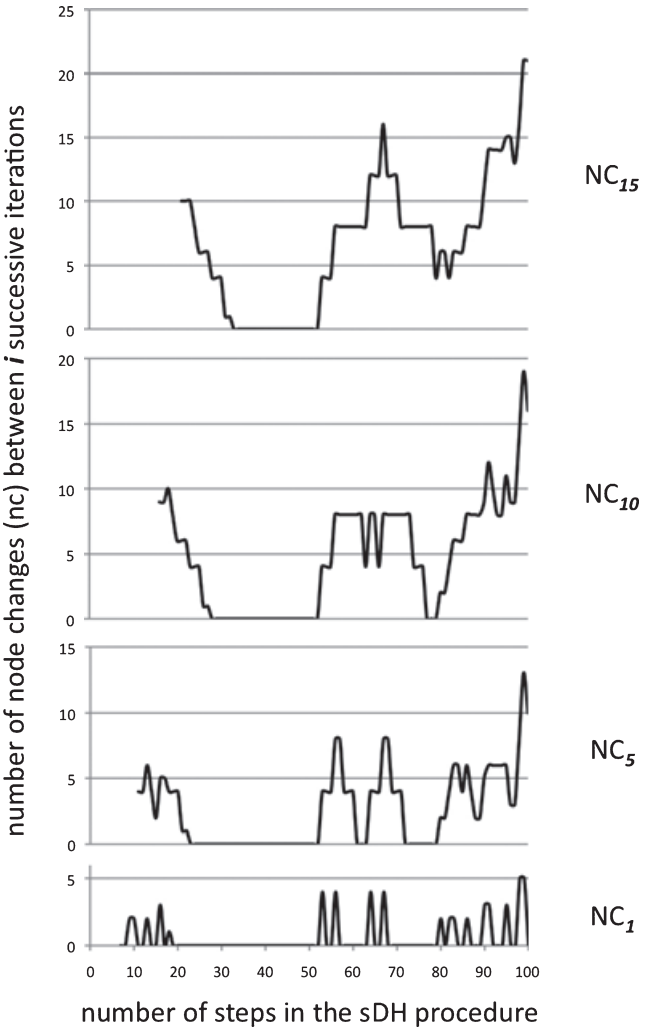

Fig. 5. Clade instability between successive iterations. The number of node changes (NC) after successive (1, 5, 10 and 15) iterations varies along the steps of the sDH procedure. One single stable phase is obtained on the more rigorous stability criterion (NC15), the one with the preferred cladogram.

Discussion

Dynamic homology

Our results show precisely what we expected from the successive sets of homology propositions along the sDH procedure. The sets of homology propositions at the extreme sDH thresholds produced unstable trees (Fig. 4a, c), and the midway thresholds produced one single and stable most-parsimonious reconstruction. The stable reconstruction (Fig. 4b) is also fully compatible with previous cladistic analyses (Coddington and Levi, 1991; Ramírez, 2000), so that the method seems appropriate for detecting the phylogenetic signal within the data.

There are other, non-phylogenetic signals within the dataset, but they are not as strong as the phylogenetic signal itself. It is obvious that most datasets should include non-phylogenetic signals, patterns that result from other structuring factors, such as ecological relationships or ontogenetical influences, or even random patterns resulting, for example, from erroneous character delimitation (Rieppel and Kearney, 2007). However, phylogenetic systematics is mostly used to tease apart these patterning factors, and our results show that the sDH procedure is apt to choose the strongest signal in the dataset, that it detects unambiguously this signal, and that it allows the reconstruction of the phylogenetic pattern. The method is well suited for any frequency behavioural data, no matter if the sequences are small dyads or longer entrainments of behavioural units, no matter if they are stereotyped or plastic, or performed by solitary individuals or by individuals in a social context. In order to use sDH with longer sequences (third, fourth or higher order chains of

actions), we would simply need an extension of the method to *n*-dimensional transition matrices.

One could argue that clade stability is not a proper measure of support for a phylogenetic hypothesis, and that it certainly could not adequately substitute parsimony as a criterion for choosing among alternative cladograms. However, we would note that it appears clear that one desirable property of any cladogram is stability under data perturbation. As Hovenkamp (2009) points out, stability is a desirable property of any observation, and it is related to repeatability, a founding theme of all scientific endeavour. It is not a coincidence that the usual procedure for measuring support on any clade is to disturb progressively the original data (either by permutation, bootstrap or jackknife methods) in order to see how strong a perturbation a clade resists. Clades that resist longer are usually considered better supported than other, less resistant clades, so that stability and support are somehow related (Goloboff et al., 2008b; but see Grant and Kluge, 2003). Giribet (2003), for example, shows that, despite interesting exceptions, there is an overall positive relationship between stability and some resampling measures of support, such as the jackknife. We could think of our sDH procedure as a perturbation of the "real" homology hypotheses present in the data. Consider the hypothetical possibility that we knew from the start the real homology hypotheses for the characters at hand. In this case, trying different combinations of character state partitions is a perturbation of the real homology hypotheses; if there is phylogenetic signal in the real homology hypotheses, we expect that progressively stronger errors of character state partition will progressively erode this signal; we surely do not expect small random perturbations to produce a signal that is different from, and stronger than, the phylogenetic one. This hypothetical example is simply the reverse of the sDH procedure. Under sDH we do not know the real homology hypotheses for the data at hand, but we are sure that they are among the progressive perturbations we make; our task is basically to search the most stable solution.

Also, we have independent, ethological reasons for choosing the stable midway threshold cladograms. It is well known from the ethological literature that there are sampling effects on the discovery of behavioural transitions, and that these transitions are rarely the result of deterministic processes (Lehner, 1996; Haccou and Meelis, 1995). It is thus clear for behavioural scientists that our procedure will frequently result in poor homology hypotheses at extremely high thresholds, because stochastic processes usually prevent the occurrence of a transition in all individuals of a species. The opposite also holds true: extremely low thresholds are expected to produce taxa-grouping hypotheses that are too inclusive, because stochastic processes guarantee that some BeTs will occur by chance alone in some individuals of some species. We therefore have ethological reasons to expect random factors to abound at the extreme thresholds, leading to the production of random, conflicting hypotheses of taxa grouping. Consequently, we expect, simply from these ethological considerations, the production of multiple and unstable phylogenetic patterns at these extreme thresholds. We should discard the extreme and unstable cladograms, and prefer those resulting from midway thresholds; and the midway cladogram is exactly the stable and single solution we prefer in the present study (Fig. 4b).

It is important to note that the sDH method recovers phylogenies, but it does not lead directly to a specific set of homology propositions. This is because it recovers the most stable phylogenetic pattern, but this pattern is the most stable reconstruction for many matrices, i.e. for many alternative homology propositions. If one is interested in the evolution of the characters, one should follow the data partitioning throughout the sDH procedure for the particular characters of interest, keeping track of the alternative character state partitions allowed in the procedure, during the stable phase. In our dataset we have two main situations: the characters enter the stable phase with a character partition that remains unchanged until the final steps of the stable phase, or alternatively the character partition changes around the beginning of the stable phase and remains unchanged thereafter. In these cases, we clearly have a preferred character partition: the one that remains unchanged for most of the stable phase. We suspect that this will also be the bulk of any other phylogenetic analysis with behavioural data, and if this proves true, our method will provide reliable homology propositions for any dataset at hand. For less obvious situations, one should apply a measure of adherence of the alternative character state partitions to the preferred, stable phylogeny; the partition that fits better the phylogeny should be preferred. It should be noted that any dynamic homology procedure leaves an amount of uncertainty about the chosen homology, and any transformational series drawn from this kind of study should be interpreted carefully and possibly scrutinized further (Agolin and D'Haese, 2009).

Conclusions

Although behavioural units are now becoming a respectable source of phylogenetical characters (Pinto, 1984; Coddington, 1989; Shultz, 1990; Alexander, 1991; Proctor, 1992; Gwynne, 1995; Crespi et al., 1998; Cap et al., 2008), the same cannot be said for more complex behaviours, such as sequences of these units. A few recent studies have dealt with behavioural sequences in a

cladistic framework (Japyassú et al., 2006; Robillard et al., 2006; Legendre et al., 2008), and our present study goes in this same direction, aiming to solve a problem that appears when using transitions as characters, namely the problem of the origin of the organization of the sequential pattern. It seems clear that in evolutionary studies we are searching for sequences of behavioural units that are shared between taxa because of a common history, but it is also clear that behavioural sequences can be the result of exogenously organized stimuli (Donahoe and Palmer, 1994), and this exogenous organization can be the result of non-historical factors. By "exogenous origin of patterns" we mean not only learned, conditioned behavioural sequences, but also simpler phenomena such as sequences of stereotyped units that are ordered together by ordered sequences of triggering stimuli, in an unconditioned chain of reactions. Our method allows the distinction between historical and non-historical sources of behavioural organization because it selects behavioural configurations based on the congruence between homology propositions in the data set. Because congruence is built into the dataset through evolutionary processes, its use will guarantee that we are dealing with evolutionary phenomena.

Behavioral sequences form the bulk of animal activities, and they promise a large number of phylogenetic characters (Japyassú and Viera, 2002; Japyassú et al., 2006). We think that our method will not only help behavioural scientists interested in evolution, but also that the great amount of phylogenetic information that it uncovers will improve the cladistic resolution of many hypotheses of relationships between taxa. The sDH algorithm can be easily implemented in software packages, so that its use can become routine. The method allows the evolutionary study of the plastic aspects of performance, an important advance as most, if not all, complex behaviours involve some degree of plasticity in its expression. Including the plastic, context-dependent aspects of performance on a cladistic framework is an important step towards the practical integration of theoretically distinct areas of biological reasoning, such as the evolutionary and ecological perspectives on diversity.

Acknowledgements

We would like to thank Martin Ramírez, Camilo Mattoni, Frédéric Legendre, Gustavo Hormiga and two anonymous reviewers for their comments. We also thank Paulo Miranda Nascimento for providing some of the illustrations for this paper. This work was partially financed by a CNPq grant to H.F.J. and by FAPESP (proc. no. 04/00296-8) and FAPESB (PNX 0016/2009).

References

Ades, C., 1982. Immediate stimulation and recent experiences as controlling factors in the predatory sequence of the spider *Argiope argentata*. Revista Interamericana de Psicologia 16, 57–63.

Agnarsson, I., Coddington, J., Knoflach, B., 2007. Morphology and evolution of cobweb spider male genitalia (Araneae, Theridiidae). J. Arachnol. 35, 334–395.

Agolin, M., D'Haese, C.H., 2009. An application of dynamic homology to morphological characters: direct optimization of setae sequences and phylogeny of the family Odontellidae (Poduromorpha, Collembola). Cladistics 25, 353–385.

Alexander, B.A., 1991. Phylogenetic analysis of the genus *Apis* (Hymenoptera: Apidae). Ann. Entomol. Soc. Am. 84, 137–149.

Brigandt, I., Griffiths, P.E., 2007. The importance of homology for biology and philosophy. Biol. Philos. 22, 633–641.

Brower, A.V.Z., Schawaroch, V., 1996. Three steps of homology assessment. Cladistics 12, 265–272.

Cap, H., Deleport, P., Joachim, J., Reby, D., 2008. Male vocal behavior and phylogeny in deer. Cladistics 24, 917–931.

Coddington, J.A., 1989. Spinneret silk spigot morphology: evidence for the monophyly of orb weaving spiders, Cyrtophorinae (Araneidae), and the group Theridiidae plus Nesticidae. J. Arachnol. 17, 71–95.

Coddington, J.A., 1990. Cladistics and spider classification: araneomorph phylogeny and the monophyly of orbweavers. Acta Zool Fennica 190, 75–87.

Coddington, J.A., Levi, H.W., 1991. Systematics and evolution of spiders (Araneae). Annu. Rev. Ecol. Syst. 22, 565–592.

Corrado, G.S., Sugrue, L.P., Seung, H.S., Newsome, W.T., 2005. Linear-nonlinear-Poisson models of primate choice dynamics. J. Exp. Anal. Behav. 84, 581–617.

Crespi, B.J., Carmean, D.A., Mound, L.A., Worobey, M., Morris, D., 1998. Phylogenetics of social behavior in Australian gall-forming thrips: evidence from mitochondrial DNA sequence, adult morphology and behavior, and gall morphology. Mol. Phylogenet. Evol. 9, 163–180.

Cross, F.R., Jackson, R.R., 2006. From eight-legged automatons to thinking spiders. In: Fujita, K., Itakura, S. (Eds.), Diversity of Cognition. Kyoto University Press, Kyoto, Japan, pp. 188–215.

Donahoe, J.W., Palmer, D.C. 1994. Learning and Complex Behavior. Boston: Allyn and Bacon.

Felsenstein, J., 1988. Phylogenies and quantitative characters. Annu. Rev. Ecol. Syst. 19, 445–471.

Fitch, W.T., Neubauer, J., Herzel, H., 2002. Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. Anim. Behav. 63, 407–418.

Freeman, W.J., Vitiello, G., 2006. Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. Phys. Life Rev. 3, 93–118.

Garcia, C.R.M., Japyassú, H.F. 2005. Stereotypy and plasticity in the predatory sequence of Theridion evexum Keyserling 1884 (Araneae: Theridiidae). Biota Neotropica 5(1a), http://www.biotaneotropica.org.br/v5n1a/en/abstract?article+BN001051a2005.

Ghiselin, M.T., 2005. Homology as a relation of correspondence between parts of individuals. Theory Biosci. 124, 91–103.

Gilbert, C., Rayor, L.S., 1985. Predatory behavior of spitting spiders (Araneae: Scytodidae) and the evolution of prey wrapping. J. Arachnol. 13, 231–241.

Giribet, G., 2003. Stability in phylogenetic formulations and its relationship to nodal support. Syst. Biol. 52, 554–564.

Goloboff, P., Mattoni, C.I., Quinteros, A.S., 2006. Continuous characters analyzed as such. Cladistics 22, 589–601.

Goloboff, P., Farris, J., Nixon, K., 2008a. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.

Goloboff, P., Carpenter, J.M., Arias, J.S., Esquivel, D.R.M., 2008b. Weighting against homoplasy improves phylogenetic analysis of morphological data sets. Cladistics 24, 758–773.

Grant, T., Kluge, A.G., 2003. Data exploration in phylogenetic inference: scientific, heuristic, or neither. Cladistics 19, 379–418.

Griffiths, P.E., 2007. The phenomena of homology. Biol. Philos. 22, 643–658.

Gwynne, D.T., 1995. Phylogeny of the Ensifera (Orthoptera): a hypothesis supporting multiple origins of acoustical signaling, complex spermatophores and maternal care in crickets, katydids, and Weta. J. Orthoptera Res. 4, 203–218.

Haberman, S.J., 1973. The analysis of residuals in cross-classified tables. Biometrics. 29, 205–220.

Haccou, P., Meelis, E. 1995. Statistical Analysis of Behavioural Data: An Approach Based on Time-Structured Models. Oxford University Press, Oxford.

Hall, B.K., 2003. Descent with modification: the unity underlying homology and homoplasy as seen through an analysis of development and evolution. Biol Rev 78, 409–433.

Harris, S.R., Gower, D.J., Wilkinson, M., 2003a. Intraorganismal homology, character construction, and the phylogeny of Aetosaurian Archosaurs (Reptilia, Diapsida). Syst. Biol. 52 (2), 239–252.

Harris, S.R., Gower, D.J., Wilkinson, M., 2003b. Phylogenetic methods and Aetosaur interrelationships: a rejoinder. Syst. Biol. 52 (6), 851–852.

Herberstein, M.E., Abernethy, K.E., Backhouse, K., Bradford, H., de Crespigny, F.E., Luckock, P.R., Elgar, M.A., 1998. The effect of feeding history on prey capture behaviour in the orb-web spider *Argiope keyserlingi* Karsch (Araneae: Araneidae). Ethology 104, 565–571.

van Hooff, J. 1982. Categories and sequences of behaviour: methods of description and analysis. In: Sherer, K.R., Ekman, P. (Eds.), Handbook of Methods in Nonverbal Behaviour Research. Cambridge University Press, Cambridge, pp. 97–112.

Hovenkamp, P., 2009. Stability and support. Cladistics 25, 107–108.

Japyassú, H.F. 2008. Cognições mínimas. In: Vianna, B. (Ed.), Biologia da Libertação. Mazza Edições, Belo Horizonte, pp. 97–108.

Japyassú, H.F., Caires, R.A., 2008. Hunting tactics in a cobweb spider (Araneae-Theridiidae) and the evolution of behavioral plasticity. J Insect Behav 21, 258–284.

Japyassú, H.F., Jotta, E.G., 2005. Foraging in Achaearanea cinnabarina Levi 1963 (Araneae, Theridiidae) and the evolution of hunting in spiders with irregular webs. Biota Neotropica 5(1a), http://www.biotaneotropica.org.br/v5n1a/en/abstract?article+BN003051a2005.

Japyassú, H.F., Macagnan, C.R., 2004. Fishing for prey: the evolution of a new predatory tactic among spiders (Araneae, Pholcidae). Revista de Etologia. 6, 79–94.

Japyassú, H.F., Viera, C., 2002. Predatory plasticity in *Nephilengys cruentata* (Araneae, Tetragnathidae): relevance for phylogeny reconstruction. Behaviour. 139, 529–544.

Japyassú, H.F., Alberts, C.C., Izar, P., Sato, T., 2006. EthoSeq: a tool for phylogenetic analysis and data mining on behavioral sequences. Behav. Res. Methods. 38, 549–556.

Kluge, A.G., 2003. The repugnant and the mature in phylogenetic inference: atemporal similarity and historical identity. Cladistics 19, 356–368.

Lawing, A.M., Meik, J.M., Schargel, W.E., 2008. Coding meristic characters for phylogenetic analysis: a comparison of step-matrix gap-weighting and generalized frequency coding. Syst. Biol. 57, 167–173.

Legendre, F., Robillard, T., Desutter-Grandcolas, L., Whiting, M.F., Grandcolas, P., 2008. Phylogenetic analysis of non-stereotyped behavioural sequences with a successive event-pairing method. Biol. J. Linn. Soc. 94, 853–867.

Lehner, P.N. 1996. Handbook of ethological methods. Cambridge University Press, Cambridge, UK.

Li, Q., Jackson, R.R., Barrion, A.T., 1999. Parental and predatory behaviour of *Scytodes sp*., an araneophagic spitting spider (Araneae: Scytodidae) from the Philippines. J. Zool. (Lond.) 247, 293–310.

Love, A.C., 2007. Functional homology and homology of function: biological concepts and philosophical consequences. Biol. Philos. 22, 691–708.

Nentwig, W., 1985. Feeding ecology of the tropical spitting spider *Scytodes longipes* (Araneae, Scytodidae). Oecologia 65, 284–288.

Ottoni, E.B., 2000. EthoLog 2.2: a tool for the transcription and timing of behavior observation sessions. Behav. Res. Methods. Instrum. Comput. 32, 446–449.

Patterson, C. 1982. Morphological characters and homology. In: Joysey, K.A., Friday, A.E. (Eds.), Problems of Phylogenetic Reconstruction. Academic Press, London, pp. 21–74.

Patterson, T.A., Thomas, L., Wilcox, C., Ovaskainen, O., Matthiopoulus, J., 2008. State-space models of individual animal movement. Trends Ecol. Evol. 23, 87–94.

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

Pinto, J.D., 1984. Cladistic and phenetic estimates of relationship among genera of eupomphine blister beetles (Coleoptera: Meloidae). Syst. Entomol. 9, 165–182.

Platnick, N.I., Coddington, J.A., Forster, R.R., Griswold, C.E., 1991a. Spinneret morphology and the phylogeny of haplogyne spiders. Am. Mus. Novit. 3016, 1–73.

Platnick, N.I., Griswold, C.E., Coddington, J.A., 1991b. On missing entries in cladistic analysis. Cladistics 7, 337–343.

Proctor, H.C., 1992. Sensory exploitation and the evolution of male mating behavior: a cladistic test using water mites (Acari: Parasitengona). Anim. Behav. 44, 745–752.

Punzo, F., 2002. Food imprinting and subsequent prey preference in the lynx spider, *Oxyopes salticus* (Araneae: Oxyopidae). Behav. Processes 58, 177–181.

Ramírez, M.J., 2000. Respiratory system morphology and the phylogeny of haplogyne spiders (Araneae, Araneomorphae). J. Arachnol. 28, 149–157.

Ramírez, M.J., 2007. Homology as a parsimony problem: a dynamic homology approach for morphological data. Cladistics 23, 1–25.

Rieppel, O. 1994. Homology, topology, and tyoplogy: the history of modern debates. In: Hall, B.K. (Ed.), Homology: The Hierarchical Basis of Comparative Biology. Academic Press Inc., San Diego, CA, pp. 63–99.

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.

Robillard, T., Legendre, F., Desutter-Grandcolas, L., Grandcolas, P., 2006. Phylogenetic analysis and alignment of behavioral sequences by direct optimization. Cladistics 22, 602–633.

Ryan, M.J. 1996. Phylogenetics in behavior: some cautions and expectations. in: Martins, E.P. (ed.), Phylogenies and the Comparative Method in Animal Behavior. Oxford University Press, Oxford, UK, pp. 1–21.

Ryan, M.J., Fox, J.H., Wilczynski, W., Rand, A.S., 1990. Sexual selection for sensory exploitation in the frog *Physalaemus pustulosus*. Nature 343, 66–67.

Schulmeister, S., Wheeler, W.C., 2004. Comparative and phylogenetic analysis of developmental sequences. Evol. Dev. 6, 50–57.

Shultz, J.W., 1990. Evolutionary morphology and phylogeny of Arachnida. Cladistics 6, 1–38.

Smith, E.N., Gutberlet, R.L. Jr, 2001. Generalized frequency coding: a method of preparing polymorphic multistate characters for phylogenetic analysis. Syst. Biol. 50, 156–169.

Vogt, L., Bartolomaeus, T., Giribet, G., 2009. The linguistic problem of morphology: structure versus homology and the standardization of morphological data. Cladistics 25, 1–25.

Wenzel, J.W., 1992. Behavioral homology and phylogeny. Ann Rev Ecol Syst 23, 361–381.
Wheeler, W., 1996. Optimization alignment: the end of multiple sequence alignment in phylogenetics? Cladistics 12, 1–9.
Wheeler, W., 2001. Homology and the optimization of DNA sequence data. Cladistics 17, S3–S11.
Wiens, J.J., 2001. Character analysis in morphological phylogenetics: problems and solutions. Syst. Biol. 50, 689–699.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Frequency of behavioural transitions.

Data S2. sDH example matrix.

Data S3. Variation of behavioural transitions in *Scytodes*. Range of variation of the frequency of dyads in *Scytodes fusca* (1a), *S. globula* (1b) and *S. itapevi* (1c).

Data S4. List of characters.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Appendix

Description and discussion of the behavioural categories of the spiders of the genus Scytodes

Approach. The spider moves towards the prey. In *S. fusca* it is usually speedy, but in the long-legged species (*S. globula* and *S. itapevi*) spiders also move very slowly, even when very close to the prey.

Detection. The spider uses legs I and II to hold a thread, forming an arc with both of them. Legs I and II can move simultaneously, or the movement can be made only with legs II (less frequent), or only with legs I (more frequent). The movement is very similar to that of the "touch" category (see below), including the distinction between *S. fusca* and the other species.

Touch. This category consists of any directed touching of the prey using legs I and II. The scytodid removes its leg from the substrate, elevates the femur and simultaneously contracts the femur–patella articulation. While the femur is still elevated, the spider extends all of the legs' articulations and then lowers the femur, placing some distal portion of the leg (usually the tarsus) over the prey. Although this category was performed by all investigated species, *S. fusca* performed a variation of this behaviour at a higher frequency: instead of extending all of the legs' articulations (before lowering the femur), these actions occurred simultaneously, so that the tarsus performed a nearly linear trajectory, when compared with the semicircular trajectory of the variant behaviour. In general, scytodids are very cautious and tend not to touch moving prey for too long.

Spit. During the execution of this behaviour, the spider slightly elevates its cephalothorax and exposes the chelicerae, which forms an angle ranging from 20° to 90°. The spider then shoots a glue-like substance from the tip of its chelicerae while simultaneously moving rapidly towards the prey, in a strike-like movement. Both the ejection of the glue and the strike were observed to occur faster than 0.03 s (a single video frame) in all of the investigated species. The glue forms a zigzag banded pattern similar to those described by Gilbert and Rayor (1985). A scytodid is able to perform consecutive spits: *S. globula* spat up to five times in the same prey, and *S. fusca* spat up to six times in a row. It is not uncommon for the spider to hit its own legs with the glue.

Manipulation. The spider approaches the subdued prey with its palps extended, performing small lateral movements to touch it. Next, it performs small exploratory tapping movements with the palps. This is usually followed by a bite (see below).

Bite. The spider uses the chelicerae to break through the prey exoskeleton. As this is not an immobilization strategy for the genus *Scytodes*, it is unclear if there is any injection of glue or venom during the performance of this category. However, it was not uncommon to observe a line of glue attached to both the spider's chelicerae and the prey immediately after the bite.

Cut thread. The studied sytodids extend their palps with oscillatory movements similar to those performed before the manipulation, slowly approaching the thread. Upon touching it, they use the palps as tweezers to take the thread of silk or glue towards the chelicerae to cut it. This can occur after spitting, in order to free the prey from glue or silk, or to release the spider's own legs from the spit. Spiders were observed performing this behaviour also before the capture as a means to have better access to the prey.

Wrap. Members of the genus *Scytodes* support their body with legs I and II, holding the immobilized prey with legs III, and using legs IV to deposit silk threads from the spinnerets directly over the prey's body or in the proximities. There is no rotation of the prey, so the spider usually places the thread in a slightly different area from where it was previously placed. When capture occurs outside the web, legs III can also be used to support the body of the spider.

Pay out a line. After attaching a thread to or near the wrapped prey, the spider moves away from it, leaving behind a dragline.

Fix. After leaving a dragline behind (walking onto the substrate or web), the spider lowers its abdomen and touches the web (or the substrate) with the spinnerets. While on the web, the spider can also do this by using legs III to bring the web closer to the spinnerets.

Fix prey. Similar to "fix" (above), but the thread is now attached to the immobilized prey. If the spider is on the web, it can flex legs III and IV (holding the prey), instead of lowering its abdomen.

Fix and rotate. The spider rotates its body sagitally and, while turning, the spinnerets are dabbed against the web in an arc.

Grooming. Scytodids perform this category basically by three means: by rubbing the palps against each other and taking them to the chelicerae; by placing the proximal portion of the leg's metatarsus between the chelicerae and by gently pushing the leg anteriorly (in the case of legs I and II) or posteriorly (in the case of legs III), until the tarsus passes through the cheliceae; and by one leg rubbing another, or onto the abdomen. This set of different behaviours occurs in repetitive bouts.

Pull. The spider detaches the prey from the surrounding silk or glue. Scytodids perform this by seizing the prey with the chelicerae and extending legs I, II and III, projecting backwards and away from the body of the spider. Legs IV can be used likewise, although this was more evident in *S. globula* and *S. itapevi*. Usually the spider transported the prey (see below) while there were still some treads attached to it.

Retrieve. The spider takes back the prey after being apart from it. *S. fusca* flexes legs I and II to bring the prey closer. *S. itapevi* and *S. globula* perform the same movements but, instead of bringing the prey closer, it is the spider that comes closer to the prey. This difference could be due not only to differences in the relative size of the legs in these species, but also to characteristics of the web (e.g. silk resistance) and the number of threads attached to the prey and the substrate.

Leg spinning. Scytodids move legs I in a sequence that results in aerial and circular displacements of the tips of each leg. In each leg the femur is elevated and then the femur–patella, patella–tibia, and tibia–metatarsus articulations are extended; next the femur is lowered and the above articulations are contracted, bringing the tarsal regions to a more proximal position. Legs I make these movements alternately, and the cycle is repeated many times. Leg spinning was observed only after contact with prey.

Leg spinning is different from "touch" or "detection" because it does not involve legs II. Also, it clearly occurs on the air, without

touching either the web or the prey. Leg spinning was often followed and preceded by touch or detection.

Gilbert and Rayor (1985) described a category that is similar to "leg spinning", which they termed "reach and roll". Nevertheless, reach and roll is used to cover the prey with glue that is dispersed over the substrate, and it involves the synchronized (as opposed to alternate) movement of both legs (i.e. both legs perform the same movement at the same time). We did not observe reach and roll in any specimen.

Transport. The spider returns to the resting site with the prey (subdued). Scytodids accomplish this by holding the prey with the chelicerae while walking. The prey is sometimes still attached to the spinnerets during the transport, either because the spider has just finished a wrapping bout, or because she has fixed a line to the prey before seizing it. Usually the spiders seize and pull (see above) the prey before transporting it, but larger spiders, like some *S. globula*, can seize the entangled prey and transport it directly, without pulling.