Investigating Morphological Complexes Using Informational Dissonance and Bayes Factors: A Case Study in Corbiculate Bees.

Diego S. Porto[1,2,3,*], Eduardo A. B. Almeida[1], Matthew W. Pennell[2,*]

[1] *Laboratório de Biologia Comparada e Abelhas (LBCA), Departamento de Biologia, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP), Universidade de São Paulo, 14040-901, Ribeirão Preto, SP, Brazil*

[2] *Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada*

[3] *Current address: Department of Biological Sciences, Virginia Polytechnic Institute and State University, 926 West Campus Drive, Blacksburg, VA 24061 USA.*

*Correspondence to be sent to D.S.P. (diegosporto@gmail.com) and M.W.P. (pennell@zoology.ubc.ca)

Abstract. It is widely recognized that different regions of a genome often have different evolutionary histories and that ignoring this variation when estimating phylogenies can be misleading. However, the extent to which this is also true for morphological data is still largely unknown. Discordance among morphological traits might plausibly arise due to either variable convergent selection pressures or else phenomena such as hemiplasy. Here we investigate patterns of discordance among 282 morphological characters, which we scored for 50 bee species particularly targeting corbiculate bees, a group that includes the well-known

eusocial honeybees and bumblebees. As a starting point for selecting the most meaningful partitions in the data, we grouped characters as morphological modules, highly integrated trait complexes that as a result of developmental constraints or coordinated selection we expect to share an evolutionary history and trajectory. In order to assess conflict and coherence across and within these morphological modules, we used recently developed approaches for computing Bayesian phylogenetic information allied with model comparisons using Bayes factors. We found that despite considerable conflict among morphological complexes, accounting for among-character and among-partition rate variation with individual gamma distributions, rate multipliers, and linked branch lengths can lead to coherent phylogenetic inference using morphological data. We suggest that evaluating information content and dissonance among partitions is useful step in estimating phylogenies from morphological data, just as it is with molecular data. Furthermore, we argue that adopting emerging approaches for investigating dissonance in genomic datasets may provide new insights into the integration and evolution of anatomical complexes.

Keywords: Apidae, entropy, morphological modules, phenotypic integration, phylogenetic information.

Running heads: INFORMATION CONTENT OF MORPHOLOGICAL DATA

Morphological modules are clusters of characters that co-vary tightly with one other, but that are relatively evolutionarily independent of other characters (Wagner and Altenberg 1996; Armbruster et al. 2014). Over long time periods, characters in the same module are expected to evolve in an integrated fashion (Klingenberg 2014); as such, modules are often considered "quasi-independent" evolutionary units (Lewontin 1978), whose "individuality" are maintained through time by their underlying gene regulatory networks (Wagner 2007; Erwin and Davidson 2009). Morphological modules are therefore historical units – much in the same way that genes are – connected through development by their particular character-identity gene regulatory networks (Rieppel 2005; Wagner 2014). We can learn about this history from the phylogenetic distributions of the morphological characters in different modules (e.g., Geeta 2003; Serb and Oakley 2005; Arendt 2008; Clarke and Middleton 2008). Indeed, this approach has been extensively explored in paleomorphological studies interested in mosaic evolution (reviewed in Clarke and Middleton 2008) and was demonstrated to be a useful way to address evolutionary trends in morphological modules. Furthermore, the emerging use of anatomical ontologies (Mabee et al. 2007, 2012; Yoder et al. 2010; Seltmann et al. 2012; Tarasov and Génier 2015; Wipfler et al. 2016; Tarasov et al. 2019) and novel phylogenetic comparative methods for multivariate phenotypes (e.g., Clavel et al. 2015; Caetano and Harmon 2019; Adams and Collyer 2019) now allow us to investigate these patterns in a coherent way and at an unprecedented scale.

Concurrent with the above-mentioned methodological developments for studying morphological modules, there has been a resurgence of interest in modeling the evolution of morphological data (often, in conjunction with molecular data) to estimate phylogenies (Lewis 2001; Nylander et al. 2004; Clarke and Middleton 2008; Ronquist et al. 2012a; Wright et al. 2015; Klopfstein et al. 2015; Giribet 2015; Wipfler et al. 2016; Tarasov 2019; Wright 2019) particularly in the context of node-dating and tip-dating (Pyron 2011; Ronquist et al.

2012a; Lee et al. 2014; Parins-Fukuchi 2018; Rosa et al. 2019), and to delimit species (e.g., Solís-Lemus et al. 2014). We share the enthusiasm of other researchers about this line of research (Giribet 2015; Wipfler et al. 2016); integrating molecular and phenotypic data will likely provide a richer understanding of evolutionary history and processes than either would on its own. However, as the scale of morphological data increases, we will have to confront new challenges. One such under-theorized issue (but see Serb and Oakley 2005) is that, just like in the well-studied context of molecular phylogenetics (Maddison 1997; Liu and Pearl 2007; Edwards 2009), different parts of the "phenome" may have different evolutionary histories due to discordance in the gene trees underlying phenotypic traits (a phenomenon dubbed "hemiplasy"; Avise and Robinson 2008; Hahn and Nakhleh 2016; Guerrero and Hahn 2018; Mendes et al. 2019a), or similar apparent evolutionary histories due to convergence (Losos 2011; Rosenblum et al. 2014). In molecular phylogenetics, it is typically assumed that though there may be topological conflict *between* gene trees, there is not conflict *within* genes (i.e., intralocus recombination occurs at negligibly low rates; but see, for example, Mendes et al. 2019b); that is to say that the genes are considered evolutionary units. For phenotypic data, we suggest that one biologically meaningful way to partition a dataset is into morphological modules (i.e., anatomical complexes) – as a result of coordinated selection and developmental/genetic constraints, we expect such modules to cohere over evolutionary time (Lewontin 1978; Wagner 1996; Geeta 2003; Serb and Oakley 2005). Such an approach is in alignment with the view of "hierarchical phylogenetics" (Serb and Oakley 2005) and allows for understanding phylogenetic patterns and underlying processes at different scales of interest. It further provides a sound basis for investigating the long-term coordination of traits both within and across modules (Geeta 2003). This opens up the possibility of inferring phylogeny and studying the underlying processes of morphological evolution at the same time. As such, investigating the evolution of morphological modules and inferring phylogenetic

relationships from morphological datasets are not different problems, but the same problem, viewed from two different perspectives.

Our first measure for assessing conflict between anatomical partitions is based on the information content of different hypothesized morphological modules. The entropy H of any (discrete) distribution can be quantified as

$$H = -\Sigma \log(p_i)p_i$$

where p_i is the probability of observing the ith possible configuration of the distribution (Shannon 1948). Information gain I can be expressed as a loss of entropy; in other words, if more information is added to a system, the probability distribution becomes more concentrated. Lindley (1956) proposed a Bayesian interpretation of this information content. If H^* is the entropy of posterior marginal distribution and H is the entropy of prior marginal distribution, then the Lindley's information measure is simply defined as

$$I = H - H^*.$$

Lewis et al. (2016) recently showed that the information content of any particular subset of the data could be quantified as how much more concentrated the posterior distribution of topologies is compared to the prior (assuming a discrete uniform prior on all possible topologies). While Lindley's information metric is useful for comparing the relative ability of different subsets of the data to inform phylogenetic inference, a more meaningful measure for assessing conflict between morphological modules is the related phylogenetic dissonance.

Lewis et al. (2016) defined phylogenetic dissonance as follows: if H^*_{merged} is the entropy of the merged posterior distributions from k subsets of data, $H^*_{average}$ is the average entropy among all k individual subsets of data, then the phylogenetic dissonance D can be calculated as

$$D = H^*_{merged} - H^*_{average}.$$

The average entropy of posterior distributions, thus, must be less than or equal to the entropy obtained by merging posterior distributions from separate analyses of each individual subset (Lewis et al. 2016). In other words, if the average entropy is less than the entropy from the merged distributions, there is conflict (i.e., dissonance) among the information provided by the different subsets (e.g., gene or morphological partitions of the data). If morphological modules appear to be evolving at least partly independently, as organismal biologists often suppose, then investigating the information content of each module would be a sensible way of evaluating how internally integrated they are, whereas investigating the conflict among different modules would indicate how independent different modules are.

Bayes factors (BF) are ratios of marginal likelihoods (i.e., the likelihood of data integrated over all parameters values of a model) used to compare different models in a Bayesian framework (Kass and Raftery 1995; Brown and Lemmon 2007). To assess the significance of different values of phylogenetic dissonance, Neupane et al. (2019) proposed using Bayes factors to compare alternative models for concatenated datasets. A similar strategy was taken by Rosa et al. (2019) with morphological data, but they used a partitioning strategy based on among-character compatibility as estimated by their homoplasy (thus not based on morphological/developmental criteria) and did not evaluate the informational dissonance between datasets. Comparing partitioned models employing different priors for tree topology, branch lengths and rate heterogeneity can potentially allow disentangling of how each anatomical partition (i.e., morphological module) can impact the topology of species trees (e.g., Clarke and Middleton 2008; Tarasov and Génier 2015). Indeed, this can be very important when dealing with morphological data, since different portions of the phenotype can potentially exhibit contrasting rates of character change (Harrison and Larsson 2015) or favor different phylogenetic hypotheses (Geeta 2003; Serb and Oakley 2005; Arendt 2008). Using Bayes factors for this purpose is a sound method to select between competing

models (Brown and Lemmon 2007), allowing the selection of an appropriate model to best fit morphological data, accounting for among-character and among-partition rate variation (Rosa et al. 2019).

In this work we evaluate the Bayesian phylogenetic information (BPI) content of morphological data (including 282 characters and 50 taxa representing corbiculate bees as well as closely related non-corbiculate bee species) to discriminate among tree topologies and compute the phylogenetic dissonance among seven anatomical partitions. The corbiculate bee clade (Hymenoptera: Apidae) is certainly one of the best studied groups of bees due to their ubiquity worldwide and relevance in the research of complex social behaviors (e.g., Noll 2002; Kawakita et al. 2008; Almeida and Porto 2014). The clade provides a good case model to investigate the effects of morphological disparity and informational conflicts in phylogenetic inference with morphological data. The group includes the familiar honeybees (*Apis* spp.: Apini) and bumblebees (*Bombus* spp.: Bombini), as well as orchid bees (Euglossini) and the large and diverse group of the stingless bees (Meliponini) (Michener 2007). The relationships among these four lineages have been intensely debated over the last 30 years and are a typical example of discordance between morphological and molecular datasets in phylogenetic inference (reviewed by Almeida and Porto 2014). Depending on the hypothesis considered, complex eusocial behaviors are interpreted to have evolved once (e.g., Roig-Alsina and Michener 1993; Noll 2002) or twice (e.g., Kawakita et al. 2008; Cardinal et al. 2011; Romiguier et al. 2016; Bossert et al. 2019) in bees. Investigating the relationships in the corbiculate clade would thus impact not only how we address issues about incongruence between molecular and morphological data, but also clarify our current understanding about social evolution in bees.

To make better use of information contributions from different morphological modules, deal with possible conflicts, and at the same time infer the phylogenetic relationships of

corbiculate bees, an appropriate partitioned model is essential. We compared different models and partitioning schemes using Bayes factors to assess the best model to employ morphology-based partitioned datasets in phylogenetic inference. We found that despite considerable conflict among information provided by distinct morphological modules, accounting for among-character and among-partition rate variation can lead to coherent and robust inference of phylogeny using morphological data. Our findings support the traditional relationships held with morphological data for corbiculate bees, with Apini sister to Meliponini, in agreement with most previous studies using this same kind of data.

MATERIALS AND METHODS

The Bayesian phylogenetic information (BPI) content and dissonance among seven morphological complexes (Supplementary Material: Fig. S1; all supplementary files and data are available on Dryad at https://doi.org/10.5061/dryad.dz08kprvc) were estimated with the software Galax (Lewis et al. 2016) adopting the same strategy outlined in Lewis et al. (2016) and Neupane et al. (2019). This strategy requires that a marginal posterior distribution of tree topologies be sampled for each partition (e.g., gene or morphological complex) and its entropy estimated and compared to that of the marginal prior distribution to calculate the Lindley's information (Lindley 1956; Lewis et al. 2016). The information values from different partitions were then used to calculate the phylogenetic dissonance among them (Lewis et al. 2016; Neupane et al. 2019). The morphological dataset used for this study was modified from Porto and Almeida (in prep., Supplementary Files S1-S3) and the organization of characters in anatomical partitions is summarized in Table S1 (Supplementary Material: Table S1). This dataset comprises 282 characters from external and internal skeletal anatomy coded for 50 taxa of bees, including 23 species of corbiculate bees and 27 apid outgroups (Supplementary File S3). The anatomical partitions were based on the overall pattern of body

tagmosis in Hexapoda (reviewed in Angelini and Kaufman 2005), but recognizing that in Apocrita the first abdominal segment is morpho-functionally integrated into the thorax (Vilhelmsen et al. 2010), resulting in three main body regions: head (HD), mesosoma (MS) and metasoma (MT). The other anatomical partitions were defined by recognizing that arthropod appendages and their derivatives (i.e., MP: mouthparts, LG: legs, WG: wings and GN/ST: male genitalia and female sting apparatus) have mostly distinct developmental basis (e.g., Rogers et al. 2002; Angelini and Kaufman 2005; Elias-Neto and Belles 2016) and possibly constitute separate morpho-functional modules.

Three sets of analyses with characters organized in different partitioning schemes were performed (Supplementary Material: Table S1): (i) by anatomical complex, with characters organized into seven morphological partitions, as defined above (7-PAR); (ii) by anatomical allocation, with characters organized into two partitions depending on whether they were external or internal structures of the body (2-PAR); and (iii) with all characters comprising one partition (FULL). Additionally, individual analyses for each morphological partition (i.e., HD, MP, MS, WG, LG, MT, and GN) were performed to evaluate the phylogenetic information to tree topology from different body regions and thus obtain structure trees (i.e., the individual trees inferred from each anatomical partition; Serb and Oakley 2005). To obtain the posterior distributions of tree topologies used in Galax, Bayesian analyses were carried out in MrBayes v.3.2.7. (Ronquist et al. 2012b) and ran on the CIPRES platform (Miller et al. 2011). Analyses were performed with 4 runs with 4 chains each for 1.0×10^7 generations, sampling every 100^{th} generation, and discarding the first 25% of each run as burn-in. The only exceptions were the following: for the reduced dataset (7 taxa only, see the discussion section), analyses were run for 5.0×10^6 generations; and for the best model overall (see results section), analysis was run for 2.0×10^7 generations. To ensure better mixing among chains, the temperature was set to 0.025. The following set of priors was used in all analyses:

Tree topology $\tau \sim$ Discrete Uniform (1, $|T|$)

Tree length $L \sim$ Exponential (0.1)

Edge length proportions $\psi \sim$ Dirichlet (1.0, …, 1.0)

Discrete Gamma shape $\alpha \sim$ Exponential (1.0) (when per partition among-character rate variation was allowed)

Rate multiplier $r \sim$ Dirichlet (1.0, …, 1.0) (when among-partition rate variation was allowed)

As suggested by previous authors (e.g., Clarke and Middleton 2008; Tarasov and Génier 2015; Neupane et al. 2019; Rosa et al. 2019), to evaluate partitioning schemes and conflict among partitions, we compared different models for the two concatenated datasets (i.e., 7-PAR and 2-PAR). We evaluated five sets of concatenated models: (a) unlinked topologies and unlinked branch lengths; in this case each partition was allowed to have its own independent tree topology and the total marginal likelihood of the model was obtained as the sum of marginal likelihoods of each partition analyzed separately (i.e., the SEPARATE model of Neupane et al. 2019); (b-c) linked topologies and unlinked branch lengths; (d-e) linked topologies and linked branch lengths with linked or (f-g) unlinked rate multipliers; and (h-i) linked topologies without among-partition rate variation. For all sets of models, among-character rate variation was either allowed (using a per partition discretized gamma function) or disallowed in order to evaluate its effect (i.e., b *vs.* c, d *vs.* e, f *vs.* g, h *vs.* i). In total, nine concatenated models were evaluated per partitioning scheme in addition to the two FULL models (a and b, with and without among-character variation accommodated), thus resulting in 20 models studied. The models evaluated embraced a broad range of possible combinations of among-partition and among-character rate variation in the concatenated datasets.

Models were compared using Bayes factors (BF) with marginal likelihoods estimated via stepping-stones sampling (Xie et al. 2010) as implemented in MrBayes (Ronquist et al.

2012b). BFs were calculated as $2 \times (\ln(M_0) - \ln(M_1))$ and the resultant values were interpreted using the scale proposed by Kass and Raftery (1995), where BF > 10 is interpreted as strong evidence in favor of model M_0. The stepping-stones sampling was performed with 4 runs of 4 chains each including 50 steps of 1.0×10^6 generations totalizing 5.0×10^7 generations. In all Bayesian analyses executed, convergence among chains and parameters was assessed by manually evaluating the uniformity of log-likelihood trace plots, standard deviation split values less than 0.01, effective sample size (ESS) values greater than 200 and values of Potential Scale Reduction Factor (PSRF) equal to 1.0 (Gelman and Rubin 1992).

RESULTS AND DISCUSSION

The results from the estimation of Bayesian phylogenetic information (BPI) content and dissonance of the dataset are presented in Table 1 and Table S2 (Supplementary Material: Table S2). The overall information content estimated for the full dataset was between 94.2% and 95.3% (Supplementary Material: Table S2, FULLa and FULLb, respectively). The information content varied from 58.9% (MT) to 78.0% (LG) considering individual anatomical partitions and was higher for external (92.3%, EXT) than that for internal characters (82.0%, INT) (Supplementary Material: Table S2). The estimated values of dissonance among individual runs in the analyses of each anatomical partition indicate that convergence of posterior distributions of tree topologies has been achieved for most partitions (Supplementary Material: Table S2, lower values of D (%), between 0.6% and 1.3%), but not for the smaller ones (i.e., WG and MT: < 20 characters). Since dissonance values are considerably higher in these cases (WG and MT: 2.9% and 5.7%, respectively), this may indicate that more samples from the posterior distributions would be needed to achieve convergence or that the number of characters in those partitions (i.e., 16 chars for WG and 11 chars for MT) was not sufficient to provide a reasonable topological inference, thus resulting

in sparse posterior distributions. The estimated value of dissonance among anatomical partitions was about 14.0% and 11.7% using the 7-PAR and the 2-PAR partitioning schemes, respectively (Table 1). This shows that, despite the high information content provided by each individual partition (as shown by I (%) values > 50% in Supplementary Material: Table S2), there is considerable conflict within the morphological dataset.

<TABLE 1>

The results from the comparisons among concatenated models with different partitioning schemes, numbers of parameters and alternative priors on tree topology, branch lengths and character rates are presented in Table 2 and Table S3 (Supplementary Material: Table S3). The best model overall was the 7-PARf, which has seven partitions, per partition gamma, a linked tree topology shared across all partitions, and branch lengths linked through individual rate multipliers. This model allows for among-character and among-partition rate variation (i.e., heterogeneity inside and across the seven anatomical modules) indicating that this particular parameterization is important to accommodate the degree of phylogenetic dissonance observed in this dataset.

<TABLE 2>

Concatenated models with a linked tree topology (i.e., same topology inferred for all partitions; 2PARb-i, 7-PARb-i) were favored over unlinked (2-PARa and 7-PARa). For models with linked topology and unlinked branch lengths (2-PARb-c, 7-PARb-c), equal rates (*rates* = *equal*) was favored over among-character rate variation. For all concatenated models with linked branch lengths (with or without among-partition rate variation accommodated

through rate multipliers, 2-PARd-i, 7-PARd-i), per partition among-character rate variation was favored (*rates = gamma*, *unlink shape (all)*); the same pattern was found for the unpartitioned models (FULLa-b). For all concatenated models (2-PARa-i, 7-PARa-i), among-partition rate variation accommodated through linked branch lengths and individual (unlinked) rate multipliers (2-PARf-g, 7-PARf-g) was favored over linked rate multipliers (2-PARd-e, 7-PARd-e) or unlinked branch lengths (2-PARb-c, 7-PARb-c).

It is important to note that concatenated models including unlinked topology or unlinked branch lengths were not favored despite the overall dissonance found across anatomical partitions using both partitioning schemes (2-PAR: 11.7% and 7-PAR: 14.0%). It is indeed intriguing, but one possible interpretation would be that models allowing for unlinked topology or unlinked branch lengths were too costly in terms of marginal likelihoods since a tree topology and all associate branch lengths should be estimated for each partition. ESS and PSRF values (i.e., higher than 200 and close to 1.0, respectively) indicate that the additional parameters were reasonably estimated, so the reason why models with unlinked topologies and/or branch lengths were not favored for this dataset may lie in other unknown factors. Nonetheless, among-partition rate variation can be accommodated simply by using linked branch lengths and rate multipliers, as have been demonstrated in previous works dealing with partitioned morphological datasets (Clarke and Middleton 2008; Tarasov and Génier 2015; Rosa et al. 2019). Another interesting result is that individual (unlinked) rate multipliers were favored over linked rate multipliers when associated with linked branch lengths. Although interpretation of this particular finding is not straightforward, a similar result was found in Tarasov and Génier (2015) when analyzing a partitioned morphological dataset for dung beetles.

The tree obtained from the Bayesian analysis with the best model overall (7-PARf) is shown in Figure 1. All corbiculate bee tribes (i.e., Apini, Bombini, Euglossini and

Meliponini) and their closest relatives (i.e., Centridini) were recovered with posterior probability close to 1.0. These results show that despite the considerable conflict among information provided by different morphological modules (i.e., 12~14%), accounting for among-character and among-partition rate variation can lead to coherent and robust inference of phylogeny using morphological data in a Bayesian framework. Our main result supports the traditional relationships held with morphological data for corbiculate bees, with Apini and Meliponini as sister-groups, in agreement with most previous studies using this same kind of data (e.g., Roig-Alsina and Michener 1993), but in contrast with hypotheses based in molecular data (e.g., Bossert et al. 2019; Supplementary Material: Fig. S2). It is remarkable that the conflict between morphology and molecules still persists even after exploring a wide range of partitioning schemes and models that account for heterogeneity in evolutionary rates.

<FIGURE 1>

The violin plots depicted in Figure 2 show that one possible explanation for the dissonance in this dataset would be due to differing rates of character evolution, since heterogeneity in rates can be found both inside each partition (Fig. 2a) and, to a lesser degree, among partitions (Fig. 2b). The second best model overall (2-PARf) also includes per partition gamma and linked branch lengths with individual rate multipliers, thus reinforcing the idea that accounting for rate heterogeneity among characters and partitions is indeed an important feature of the model, independent from the number of partitions. Interestingly, models allowing for each anatomical partition to have its own free tree topology and associated branch lengths (e.g., 7-PARa and 2-PARa; the total marginal likelihood of the partitioned model, in this case, was calculated as the sum of the marginal likelihoods estimated in the analyses of each individual partition, as suggested in Neupane et al. 2019 in

their SEPARATE model) were not favored over models with a linked tree topology (e.g., 7-PARb-i and 2-PARb-i). Completely partitioned models winning over shared topology models would be expected if, for example, the phylogenetic information provided by different partitions was in high degree of conflict. Models treating the whole dataset as a single partition (FULLa-b) were not favored as well (Bayes factors against them >100). This latter scenario would be expected, for example, if rates of character evolution were considerably homogeneous across the dataset. Therefore, the scenario found here seems to fall into an intermediate situation.

<FIGURE 2>

The trees obtained from analyses of individual morphological partitions (Fig. 3) demonstrate that many clades recovered with posterior probability greater than 0.9 are also found in the analysis obtained with the best model overall (7-PARf) (indicated as color-filled circles at nodes of the online version of Fig. 1). When inspecting the trees summarized in Figure 3, it is possible to observe that individual partitions were decidedly informative to the trees they imply (as for example, values of I (%) > 50%), but despite their unique individual majority rule tree topologies, many clades are shared among them and with the tree from the best model overall (e.g., Centridini was recovered with the partitions HD, LG, and GN; Euglossini with the partitions HD, MP, MS, LG, and GN). The dissonance inferred for this dataset, therefore, is likely not the result of pervasive conflict among tree topologies of individual partitions. Instead, the dissonance may stem from a mixture of factors, as for example, very low resolution in trees obtained from some partitions (e.g., MT) or focal clades varying in position in different trees (e.g., HD vs. WG).

<FIGURE 3>

As pointed out before, dissonance may also stem from differing rates of character evolution among partitions. The relatively slowly evolving partitions (e.g., Fig. 2b: HD, MP, MS, and GN) share many clades among themselves and with the best model tree (7PARf). As for example, Apini, Meliponini, and Apini + Meliponini are always recovered, despite the lower support in some cases (e.g., Meliponini in HD and MP). Some of the sharpest disagreements among datasets, however, occur in relatively fast evolving partitions (e.g., Fig. 2b: WG and LG) or those with the fewest characters (e.g., WG and MT). For example, monophyly of the well-supported corbiculate tribes Meliponini and Euglossini is not recovered with WG nor Apini with LG. Furthermore, some controversial relations are recovered with fast evolving partitions, such as Bombini + Euglossini sister to Manueliini with WG, and Tetrapediini + Ctenoplectriini sister to Centridini with LG. These spurious results suggest that convergence due to fast evolving characters can be a possible explanation for dissonance at least in some partitions (WG and LG). The effects of fast evolving partitions can be further linked to the fact that characters may plausibly be under distinctive selective pressures or morpho-functional constraints.

Despite our finding that there is substantial dissonance between different morphological modules, each partition is highly informative for a particular subclade (or set of subclades) (Fig. 3; Supplementary Material: Table S2). Different clades (e.g., Apini, Bombini, Centridini, Euglossini, and Meliponini) are shared among the individual partition trees and are also recovered in the tree from the best model overall (Fig. 1). This shows that despite dissonance, different morphological modules provide complementary information. Some clades are recovered only with particular partitions (e.g., Centridini: HD, LG, and GN) or only with low posterior probability (e.g., Meliponini: HD, MP, and GN). When all the

information from different anatomical modules is put together, the resolution of the final tree is improved, both with an increase in number of clades recovered, and higher posterior probabilities for those clades (Fig. 1). Therefore, trees obtained from individual anatomical modules, as expected, are different, but not necessarily in complete conflict. In this regard, we have demonstrated that estimating the Bayesian phylogenetic information (BPI) content and dissonance in morphological datasets can be used to assess the contributions of individual partitions (i.e., modules) and understand the conflicts among them, in addition to detect where and explain how information and conflicts influence the final species tree. This kind of informational analysis would be viewed as a desirable step in any phylogenetic inference from partitioned morphological data, just as it is with molecular data.

We have demonstrated that emerging approaches applied to evaluate conflict among partitions in molecular datasets (Lewis et al. 2016; Neupane et al. 2019) can be explored to understand incongruence in morphological datasets as well. The use of a measure of entropy as a proxy to the information content of data is a natural choice in a Bayesian framework (Lindley 1956; Lewis et al. 2016). Since the prior distribution of tree topologies is known (i.e., the discrete uniform prior probability depends only on the total number of taxa N in a given tree: $1/N$) and the posterior can be sampled via Metropolis-coupled Markov-chain Monte-Carlo, the total information can be accurately estimated with conditional clade distributions (Larget 2013) as discussed by Lewis et al. (2016). One potential problem with this approach, as stated by Lewis et al., is that as the total number of taxa increases, the coverage (i.e., unique tree topologies sampled in the posterior relative to the maximum possible number of rooted tree topologies under a discrete uniform prior) steadily decreases. In this scenario, the Bayesian phylogenetic information is systematically overestimated (Lewis et al. 2016). We should note as well that in some real datasets the approximations of Larget used by Lewis et al. may not hold (Whidden et al. 2015). To evaluate these effects in our dataset, we further

estimated the information content of the morphological matrix of Porto and Almeida (in prep.) using a reduced dataset including only seven taxa (Supplementary Files S4 and S5), thus ensuring that the coverage of the posterior would be around 100% (Supplementary Material: Tables S4 and S5; Fig. S3). We have observed that the overall information estimated for each morphological partition in the larger dataset (i.e., 50 taxa) was considerably higher (5.3~209.6%) than that in the scenario with only seven taxa (e.g., WG, LG, MT and GN) or, in some cases (e.g., HD, MP and MS), slightly lower (0.3~7.3%). The exceptionally high information values estimated for the larger dataset were for particular cases involving smaller partitions (e.g., WG: 16 chars and MT: 11 chars) or partitions with more missing information for non-corbiculate taxa (e.g., LG and GN). The values of dissonance estimated were also different, with about 14% and 12% dissonance found respectively for the 7-PAR and 2-PAR schemes in the analysis of the larger dataset (Table 1) and about 24% and 8% respectively for the smaller one (Supplementary Material: Table S5).

The relationship between information content and number of characters (i.e., partition size) is not so straightforward. Lewis et al. (2016) have shown that as the partition size increases to a given threshold (particular to each dataset), the information provided by data turns redundant with that already available (as for example, see the simulations presented in Fig. 1 of Lewis et al. 2016); in other words, more and more characters favor the same few tree topologies. The relationship between information content and number of taxa is also entangled in this problem. Lewis et al. (2016) demonstrated that information is systematically overestimated when the number of taxa grows more than seven terminals. But another possible effect of this bias would be that the distribution of information among clades also changes, since the overall information can be split into clade-specific components (Larget 2013; Lewis et al. 2016). For a dataset with a fixed number of characters, if the number of taxa increases, the total information initially available becomes potentially scattered among

more and more clades through smaller clade-specific information components, as estimated with the conditional clade distributions (Larget 2013; Lewis et al. 2016).

We emphasize here that exploring emerging approaches for investigating dissonance in genomic datasets may provide new insights into integration and evolution of anatomical complexes. Morphological characters and groups of characters are expected to be linked by underlying hidden processes maintaining their identity and continuity through time and across species (Rieppel 2005; Wagner 2007; Tarasov 2020). Understanding how the information is partitioned in different anatomical modules and how the individual phylogenetic hypotheses implied by them conflict with each other can help us to not only make more reliable phylogenetic inferences, but also to provide new insights into the evolution of phenotypic integration.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.dz08kprvc.

FUNDING

This work was supported by the São Paulo Research Foundation, FAPESP (grants 2015/15347-1 and 2018/19277-6 to D.S.P.; and 2018/09666-5 to E.A.B.A. and D.S.P.) and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* - Brazil (CAPES) - Finance Code 001. D.S.P. was also partly supported by a National Science Foundation Grant (DBI-1661516). E.A.B.A. was partly supported by the Brazilian National Council for Scientific and Technological Development (CNPq grants 310111/2019-6 and 422019/2018-6). M.W.P. was supported by a NSERC Discovery Grant.

ACKNOWLEDGMENTS

We would like to thank the Editor in Chief Bryan Carstens, Associate Editor Robert Thomson, Paul O. Lewis and two anonymous reviewers for many valuable comments and suggestions. We are thankful to Josef Uyeda, Suman Neupane and all members of Uyeda lab and Pennell lab for useful discussions on an earlier version of this work. Portions of the analyses were run on the UBC Zoology Computing Cluster.

REFERENCES

Adams D.C., Collyer M.L. 2019. Phylogenetic comparative methods and the evolution of multivariate phenotypes. Annu. Rev. Ecol. Evol. Syst. 50:405–425.

Almeida E.A.B., Porto D.S. 2014. Investigating eusociality in bees while trusting the uncertainty. Sociobiology. 61:355–368.

Angelini D.R., Kaufman T.C. 2005. Comparative developmental genetics and the evolution of arthropod body plans. Ann. Rev. Genet. 39:95–119.

Arendt D. 2008. The evolution of cell types in animals: emerging principles from molecular studies. Nat. Rev. Genet. 9:868–882.

Armbruster W.S., Pélabon C., Bolstad G.H., Hansen T.F. 2014. Integrated phenotypes: understanding trait covariation in plants and animals. Philos. Trans. R. Soc. B Biol. 369:20130245.

Avise J.C., Robinson T.J. 2008. Hemiplasy: a new term in the lexicon of phylogenetics. Syst. Biol. 57:503–507.

Bossert S., Murray E.A., Almeida E.A.B., Brady S.G., Blaimer B.B., Danforth B.N. 2019. Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. Mol. Phylogenet. Evol. 130:121–131.

Brown J.M., Lemmon A.R. 2007. The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. Syst. Biol. 56:643–655.

Caetano D.S., Harmon L.J. 2019. Estimating correlated rates of trait evolution with uncertainty. Syst. Biol. 68:412–429.

Cardinal S., Danforth B.N. 2011. The antiquity and evolutionary history of social behavior in bees. PLOS ONE. 6:e21086.

Clarke J.A., Middleton K.M. 2008. Mosaicism, modules, and the evolution of birds: results from a Bayesian approach to the study of morphological evolution using discrete character data. Syst. Biol. 57:185–201.

Clavel J., Escarguel G., Merceron G. 2015. mvMORPH: an R package for fitting multivariate evolutionary models to morphometric data. Methods Ecol. Evol. 6:1311–1319.

Edwards S.V. 2009. Is a new and general theory of molecular systematics emerging? Evolution. 63:1–19.

Elias-Neto M., Belles X. 2016. Tergal and pleural structures contribute to the formation of ectopic prothoracic wings in cockroaches. Roy. Soc. Open Sci. 3:160347.

Erwin D.H., Davidson E.H. 2009. The evolution of hierarchical gene regulatory networks. Nat. Rev. Genet. 10:141.

Geeta R. 2003. Structure trees and species trees: what they say about morphological development and evolution. Evol. Dev. 5:609–2621.

Gelman A., Rubin D.B. 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7:457–472.

Giribet G. 2015. Morphology should not be forgotten in the era of genomics—a phylogenetic perspective. Zool. Anz. 256:96–103.

Guerrero R.F., Hahn M.W. 2018. Quantifying the risk of hemiplasy in phylogenetic inference. Proc. Natl. Acad. Sci. USA. 115:12787–12792.

Hahn M.W., Nakhleh L. 2016. Irrational exuberance for resolved species trees. Evolution. 70:7–17.

Harrison L.B., Larsson H.C. 2015. Among-character rate variation distributions in phylogenetic analysis of discrete morphological characters. Syst. Biol. 64:307–324.

Kass R.E., Raftery A.E. 1995. Bayes factors. J. Am. Stat. Assoc. 90:773–795.

Kawakita A., Ascher J.S., Sota T., Kato M., Roubik D.W. 2008. Phylogenetic analysis of the corbiculate bee tribes based on 12 nuclear protein-coding genes (Hymenoptera: Apoidea: Apidae). Apidologie. 39:163–175.

Klingenberg C.P. 2014. Studying morphological integration and modularity at multiple levels: concepts and analysis. Philos. Trans. R. Soc. B Biol. 369:20130249.

Klopfstein S., Vilhelmsen L., Ronquist F. 2015. A nonstationary Markov model detects directional evolution in hymenopteran morphology. Syst. Biol. 64:1089–1103.

Larget B. 2013. The estimation of tree posterior probabilities using conditional clade probability distributions. Syst. Biol. 62:501–511.

Lee M.S., Cau A., Naish D., Dyke G.J. 2014. Morphological clocks in paleontology, and a mid-Cretaceous origin of crown Aves. Syst. Biol. 63:442–449.

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

Lewis P.O., Chen M.H., Kuo L., Lewis L.A., Fučíková K., Neupane S., Wang Y.-B., Shi D. 2016. Estimating Bayesian phylogenetic information content. Syst. Biol. 65:1009–1023.

Lewontin R. 1978. Adaptation. Sci. Am. 239:212–230.

Lindley D.V. 1956. On a measure of the information provided by an experiment. Ann. Math. Stat. 27:986–1005.

Liu L., Pearl D.K. 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. Syst. Biol. 56:504-514.

Losos J.B. 2011. Convergence, adaptation, and constraint. Evolution. 65:1827–1840.

Mabee P.M., Ashburner M., Cronk Q., Gkoutos G.V., Haendel M., Segerdell E., Mungall C., Westerfield M. 2007. Phenotype ontologies: the bridge between genomics and evolution. Trends Ecol. Evol. 22:345–350.

Mabee P., Balhoff J.P., Dahdul W.M., Lapp H., Midford P.E., Vision T.J., Westerfield M. 2012. 500,000 fish phenotypes: the new informatics landscape for evolutionary and developmental biology of the vertebrate skeleton. J. Appl. Ichthyol. 28:300–305.

Maddison W.P. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

Mendes F.K., Fuentes-González J.A., Schraiber J.G., Hahn M.W. 2019a. A multispecies coalescent model for quantitative traits. Elife 7:e36482.

Mendes F.K., Livera A.P., Hahn M.W. 2019b. The perils of intralocus recombination for inferences of molecular convergence. Philos. Trans. R. Soc. B Biol. 374:20180244.

Michener C.D. 2007. The bees of the world. Baltimore (MD): John Hopkins University Press.

Miller M.A., Pfeiffer W., Schwartz T. 2011. The CIPRES science gateway: a community resource for phylogenetic analyses. In Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery, 18 Jul. 2011, Salt Lake City, UT, p 41.

Neupane S., Fučíková K., Lewis L.A., Kuo L., Chen M.H., Lewis P.O. 2019. Assessing combinability of phylogenomic data using Bayes factors. Syst. Biol. 68:744–754.

Noll F.B. 2002. Behavioral phylogeny of corbiculate Apidae (Hymenoptera; Apinae), with special reference to social behavior. Cladistics. 18:137–153.

Nylander J.A.A., Ronquist F., Huelsenbeck J.P., Nieves-Aldrey J. 2004. Bayesian phylogenetic analysis of combined data. Syst Biol. 53:47–67.

Parins-Fukuchi C. 2018. Bayesian placement of fossils on phylogenies using quantitative morphometric data. Evolution 72:1801–1814.

Pyron R.A. 2011. Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. Syst. Biol. 60:466–481.

Rieppel O. 2005. Modules, kinds, and homology. J. Exp. Zool. Part B 304:18–27.

Rogers B.T., Peterson M.D., Kaufman T.C. 2002. The development and evolution of insect mouthparts as revealed by the expression patterns of gnathocephalic genes. Evol. Dev. 4:96–110.

Roig-Alsina A., Michener C.D. 1993. Studies of the phylogeny and classification of long-tongued bees (Hymenoptera: Apoidea). Univ. Kans. Sci. Bull. 55:123–162.

Romiguier J., Cameron S.A., Woodard S.H., Fischman B.J., Keller L., Praz C.J. 2016. Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. Mol Biol Evol. 33:670–678.

Ronquist F., Klopfstein S., Vilhelmsen L., Schulmeister S., Murray D.L., Rasnitsyn A.P. 2012a. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. Syst. Biol. 61:973–999.

Ronquist F., Teslenko M., Van Der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012b. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61:539–542.

Rosa B.B., Melo G.A., Barbeitos M.S. 2019. Homoplasy-based partitioning outperforms alternatives in Bayesian analysis of discrete morphological data. Syst. Biol. 68: 657–671.

Rosenblum E.B., Parent C.E., Brandt E.E. 2014. The molecular basis of phenotypic convergence. Annu. Rev. Ecol. Evol. Syst. 45:203–226.

Seltmann K.C., Yoder M.J., Miko I., Forshage M., Bertone M.A., Agosti D., Austin A.D., Balhoff J.P., Borowiec M.L., Brady S.G., Broad G.R., Brothers D.J., Burks R.A., Buffington M.L., Campbell H.M., Dew K.J., Ernst A.F., Fernández-Triana J.L., Gates M.W., Gibson G.A.P., Jennings J.T., Johnson N.F., Karlsson D., Kawada R., Krogmann L., Kula R.L.,

Mullins P.L., Ohl M., Rasmussen C., Ronquist F., Schulmeister S., Sharkey M.J., Talamas E., Tucker E., Vilhelmsen L., Ward P.S., Wharton R.A., Deans A.R. 2012. A hymenopterists' guide to the Hymenoptera Anatomy Ontology: utility, clarification, and future directions. J. Hymenopt. Res. 27:67–88.

Serb J.M., Oakley T.H. 2005. Hierarchical phylogenetics as a quantitative analytical framework for evolutionary developmental biology. Bioessays. 27:1158–1166.

Shannon C.E. 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27:379–423.

Solí s-Lemus C., Knowles L.L., Ané C. 2014. Bayesian species delimitation combining multiple genes and traits in a unified framework. Evolution 69:492–507.

Tarasov S. 2019. Integration of anatomy ontologies and evo-devo using structured Markov models suggests a new framework for modeling discrete phenotypic traits. Syst. Biol. 68:698–716.

Tarasov S. 2020. The invariant nature of a morphological character and character state: insights from gene regulatory networks. Syst. Biol. 69:392–400.

Tarasov S., Génier F. 2015. Innovative Bayesian and parsimony phylogeny of dung beetles (Coleoptera, Scarabaeidae, Scarabaeinae) enhanced by ontology-based partitioning of morphological characters. PLoS One. 10(3): e0116671.

Tarasov S., Miko I., Yoder M.J., Uyeda J. 2019. PARAMO pipeline: reconstructing ancestral anatomies using ontologies and stochastic mapping. Insect Syst. Divers. 3:1.

Vilhelmsen L., Mikó I., Krogmann L. 2010. Beyond the wasp-waist: structural diversity and phylogenetic significance of the mesosoma in apocritan wasps (Insecta: Hymenoptera). Zool. J. Linn. Soc. 159:22–194.

Wagner G.P. 1996. Homologues, natural kinds and the evolution of modularity. Am. Zool. 36:36–43.

Wagner G.P. 2007. The developmental genetics of homology. Nat. Rev. Genet. 8:473–479.

Wagner G.P. 2014. Homology, genes, and evolutionary innovation. Princeton (NJ): Princeton University Press.

Wagner G.P., Altenberg L. 1996. Complex adaptations and the evolution of evolvability. Evolution. 50:967–976.

Whidden C., Matsen IV F.A. 2015. Quantifying MCMC exploration of phylogenetic tree space. Syst. Biol. 6:472–491.

Wipfler B., Pohl H., Yavorskaya M.I., Beutel R.G. 2016. A review of methods for analysing insect structures—the role of morphology in the age of phylogenomics. Curr. Opi. Insect Sci. 18:60–68.

Wright A.M. 2019. A systematist's guide to estimating Bayesian phylogenies from morphological data. Insect Syst. Divers. 3:2.

Wright A.M., Lloyd G.T., Hillis D.M. 2015. Modeling character change heterogeneity in phylogenetic analyses of morphology through the use of priors. Syst. Biol. 65:602–611.

Xie W., Lewis P.O., Fan Y., Kuo L., Chen M.H. 2010. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Syst. Biol. 60:150–160.

Yoder M.J., Mikó I., Seltmann K.C., Bertone M.A., Deans A.R. 2010. A gross anatomy ontology for Hymenoptera. PlOS One. 5:e15991.

TABLE CAPTIONS

TABLE 1. Bayesian phylogenetic information content and dissonance estimated among all morphological complexes.

TABLE 2. Comparisons among models with different partitioning schemes and parameters.

FIGURE CAPTIONS

FIGURE 1. Tree topology obtained from the analysis with the best model overall (7-PARf). Numbers next to nodes indicate posterior probabilities of clades. Filled circles indicate clades with posterior probability greater than 0.9 in the analyses of individual morphological partitions (as shown in Fig. 3). Top-left box presents a summary of the organization of characters into anatomical modules (as shown in Fig. S1); numbers denote the amount of characters in each anatomical partition (percentages shown in parentheses). Codes making reference to morphological partitions follow those presented in Tables 1 and S1. For the color figure, refer to the online version of this paper available at Systematic Biology.

FIGURE 2. Violin plots showing the posterior distributions of a) rate variances (1/alpha) and b) rate multipliers for each morphological partition in the analysis with the best model overall (7-PARf). Discretized gamma distributions with a four-rate shape alpha parameter were used to accommodate among-character rate variation in each individual partition. Rate multipliers and linked branch lengths were employed to accommodate among-partition rate variation, as suggested by Rosa et al. (2019). Codes and colors making reference to morphological partitions follow those presented in Figure 1 (top-left box) and Table 1. For the color figure, refer to the online version of this paper available at Systematic Biology.

FIGURE 3. Tree topologies obtained from analyses of each individual morphological partition. Codes for partitions follow those of Table S1 and Figure S1 and are indicated in the bottom-right corner of each tree. Color-coding of branches and clades indicating certain bee tribes follows Figure 1. Circles: black denotes clades with posterior probability greater than 0.90; grey stands for clades with posterior between 0.90 and 0.75; white indicates clades with

posterior between 0.75 and 0.50. Abbreviations: N: number of characters in each partition; I: total information estimated in favor of a given tree topology expressed as a percentage of the maximum. For the color figure, refer to the online version of this paper available at Systematic Biology.

TABLE 1. Bayesian phylogenetic information content and dissonance estimated among all morphological complexes.

Analysis [a]	N[b]	C[c]	H*[d]	I[e]	I (%)[f]	D[g]	D (%)[h]
7 partition scheme							
Run 1							
HD	74975	0.000	40.863	130.571	76.164	-	-
MP	74991	0.000	38.927	132.508	77.293	-	-
MS	74981	0.000	37.987	133.447	77.842	-	-
WG	74999	0.000	58.582	112.852	65.828	-	-
LG	74981	0.000	37.206	134.228	78.297	-	-
MT	75000	0.000	66.419	105.015	61.257	-	-
GN	75000	0.000	48.037	123.397	71.979	-	-
average	74989.6	0.000	46.860	124.574	72.666	-	-
merged	524927	0.000	54.494	116.940	68.213	7.634	14.009
Run 2							
HD	74963	0.000	40.926	130.508	76.127	-	-
MP	74988	0.000	38.945	132.489	77.283	-	-
MS	74987	0.000	38.007	133.427	77.830	-	-
WG	75000	0.000	58.436	112.998	65.913	-	-
LG	74969	0.000	37.195	134.239	78.304	-	-
MT	75000	0.000	66.600	104.835	61.151	-	-
GN	75000	0.000	47.910	123.524	72.053	-	-
average	74986.7	0.000	46.860	124.574	72.666	-	-
merged	524907	0.000	54.499	116.935	68.210	7.639	14.017
Run 3							
HD	74975	0.000	40.703	130.732	76.257	-	-
MP	74986	0.000	39.008	132.426	77.246	-	-
MS	74989	0.000	38.095	133.340	77.779	-	-
WG	75000	0.000	58.580	112.854	65.829	-	-
LG	74976	0.000	37.141	134.293	78.335	-	-
MT	75000	0.000	66.623	104.812	61.138	-	-
GN	74998	0.000	47.938	123.496	72.037	-	-
average	74989.1	0.000	46.870	124.565	72.660	-	-
merged	524924	0.000	54.512	116.922	68.202	7.642	14.019
Run 4							
HD	74966	0.000	40.787	130.647	76.208	-	-
MP	74994	0.000	38.961	132.473	77.273	-	-
MS	74981	0.000	38.058	133.376	77.800	-	-
WG	75000	0.000	58.398	113.036	65.936	-	-
LG	74980	0.000	37.255	134.179	78.269	-	-
MT	75000	0.000	66.703	104.731	61.091	-	-
GN	75000	0.000	47.969	123.466	72.019	-	-
average	74988.7	0.000	46.876	124.558	72.657	-	-
merged	524921	0.000	54.531	116.903	68.191	7.655	14.038
Mean		0.000	54.509	116.925	68.204	7.642	14.021

TABLE 1. Continued.

Analysis [a]	N[b]	C[c]	H*[d]	I[e]	I (%)[f]	D[g]	D (%)[h]
2 partition scheme							
Run 1							
EXT	51018	0.383	13.248	158.186	92.272	-	-
INT	74849	0.000	30.604	140.831	82.148	-	-
average	62933.5	0.191	21.926	149.509	87.210	-	-
merged	125867	0.034	24.829	146.605	85.517	2.903	11.694
Run 2							
EXT	50998	0.381	13.274	158.160	92.257	-	-
INT	74841	0.000	30.608	140.827	82.146	-	-
average	62919.5	0.190	21.941	149.493	87.201	-	-
merged	125839	0.033	24.838	146.596	85.511	2.897	11.665
Run 3							
EXT	51116	0.381	13.262	158.172	92.264	-	-
INT	74822	0.000	30.502	140.932	82.207	-	-
average	62969	0.191	21.882	149.552	87.236	-	-
merged	125938	0.034	24.774	146.661	85.549	2.891	11.671
Run 4							
EXT	51078	0.385	13.229	158.205	92.283	-	-
INT	74864	0.000	30.609	140.825	82.145	-	-
average	62971	0.193	21.919	149.515	87.214	-	-
merged	125942	0.033	24.829	146.605	85.517	2.910	11.719
Mean		0.033	24.817	146.617	85.523	2.900	11.687

[a]Codes making reference to morphological partitions and partitioning schemes: 7-PAR: characters organized using the 7-partition scheme (i.e., HD, MP, …GN); 2-PAR: characters organized using the 2-partition scheme (i.e., EXT and INT); HD: characters from head (not including mouthparts); MP: characters from mouthparts; MS: characters from mesosoma (not including wings and legs); WG: characters from wings; LG: characters from legs; MT: characters from metasoma (not including male genitalia and female sting apparatus); GN: characters from male genitalia and female sting apparatus; EXT: all external characters; INT: all internal characters.

[b]Number of unique tree topologies sampled in the posterior.

[c]Estimated posterior coverage (Larget 2013).

[d]Entropy of marginal posterior tree topology distribution.

[e]Phylogenetic information (Lindley 1956).

[f]Phylogenetic information expressed as percentage of maximum.

[g]Phylogenetic dissonance (Lewis et al. 2016).

[h]Phylogenetic dissonance expressed as percentage of total.

TABLE 2. Comparisons among models with different partitioning schemes and parameters.

Model [a]	N.Ch. [b]	N.Part. [c]	N.Par. [d]	ACV [e]	APV [f]	Br.L. [g]	Tree [h]	MgLk [i]	BF [j]
HD	42	1	5	gamma	-	-	-	-864.64	-
MP	52	1	5	gamma	-	-	-	-860.60	-
MS	57	1	5	gamma	-	-	-	-988.84	-
WG	16	1	5	gamma	-	-	-	-469.52	-
LG	49	1	5	gamma	-	-	-	-870.88	-
MT	11	1	5	gamma	-	-	-	-276.00	-
GN	55	1	5	gamma	-	-	-	-718.55	-
EXT	181	1	5	gamma	-	-	-	-3086.19	-
INT	101	1	5	gamma	-	-	-	-1503.71	-
7-PARa	282	7	-	-	-	-	unlinked	-5049.03	1729.85
7-PARb	282	7	23	gamma	-	unlinked	linked	-4530.43	692.66
7-PARc	282	7	16	equal	-	unlinked	linked	-4434.17	500.14
7-PARd	282	7	11	gamma	linked mult	linked	linked	-4430.65	493.10
7-PARe	282	7	4	equal	linked mult	linked	linked	-4523.30	678.40
7-PARf *	282	7	17	gamma	unlinked mult	linked	linked	-4184.10	0 *
7-PARg	282	7	10	equal	unlinked mult	linked	linked	-4269.10	170.64
7-PARh	282	7	11	gamma	-	linked	linked	-4430.07	491.94
7-PARi	282	7	4	equal	-	linked	linked	-4499.36	630.52
2-PARa	282	2	-	-	-	-	unlinked	-4589.90	811.60
2-PARb	282	2	8	gamma	-	unlinked	linked	-4466.73	565.26
2-PARc	282	2	6	equal	-	unlinked	linked	-4440.65	513.10
2-PARd	282	2	6	gamma	linked mult	linked	linked	-4435.33	502.46
2-PARe	282	2	4	equal	linked mult	linked	linked	-4474.89	581.58
2-PARf	282	2	7	gamma	unlinked mult	linked	linked	-4262.96	157.72
2-PARg	282	2	5	equal	unlinked mult	linked	linked	-4368.60	369.00
2-PARh	282	2	6	gamma	-	linked	linked	-4434.13	500.06
2-PARi	282	2	4	equal	-	linked	linked	-4543.79	719.38
FULLa	282	1	5	gamma	-	-	-	-4432.65	497.10
FULLb	282	1	4	equal	-	-	-	-4543.97	719.74

[a]Codes referring to morphological partitions and partitioning schemes follow those presented in Table 1 and S1: FULL: characters organized using the 1-partition scheme (i.e., all 282 chars comprising one partition).

[b]Number of characters in each morphological partition/partitioning scheme.

[c]Number of partitions in the model.

[d]Number of parameters in the model.

[e]Among-character rate variation was accommodated using per partition discretized gamma distribution with four rate categories.

[f]Among-partition rate variation was accommodated using unlinked branch lengths (see branch lengths prior) or linked branch lengths and rate multipliers (linked or unlinked); linked rate multipliers (linked mult) means that multipliers for each partition are allowed to vary constraining the mean across all partitions to be 1.0; unlinked rate multipliers (unlinked mult) means that each partition is allowed to have its own multiplier estimated independently.

[g]Branch lengths prior: unlinked means that branch length proportions were allowed to differ across different partitions; linked means that all partitions share the same branch length proportions.

[h]Tree topology prior: unlinked means that tree topology is allowed to differ across different partitions (i.e., each partition can have its own inferred tree); linked means that all partitions share the same inferred tree topology.

[i]Marginal likelihood of the model.

[j]Bayes factors calculated relative to the best model overall (indicated with *).

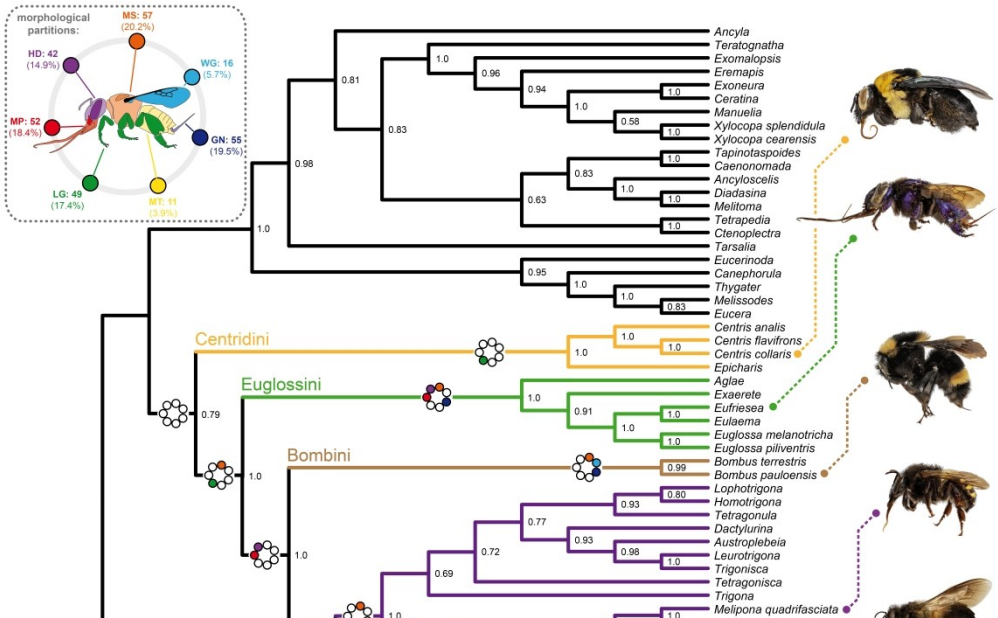

FIGURE 1. Tree topology obtained from the analysis with the best model overall (7-PARf). Numbers next to nodes indicate posterior probabilities of clades. Filled circles indicate clades with posterior probability greater than 0.9 in the analyses of individual morphological partitions (as shown in Fig. 3). Top-left box presents a summary of the organization of characters into anatomical modules (as shown in Fig. S1); numbers denote the amount of characters in each anatomical partition (percentages shown in parentheses). Codes making reference to morphological partitions follow those presented in Tables 1 and S1. For the color figure, refer to the online version of this paper available at Systematic Biology.

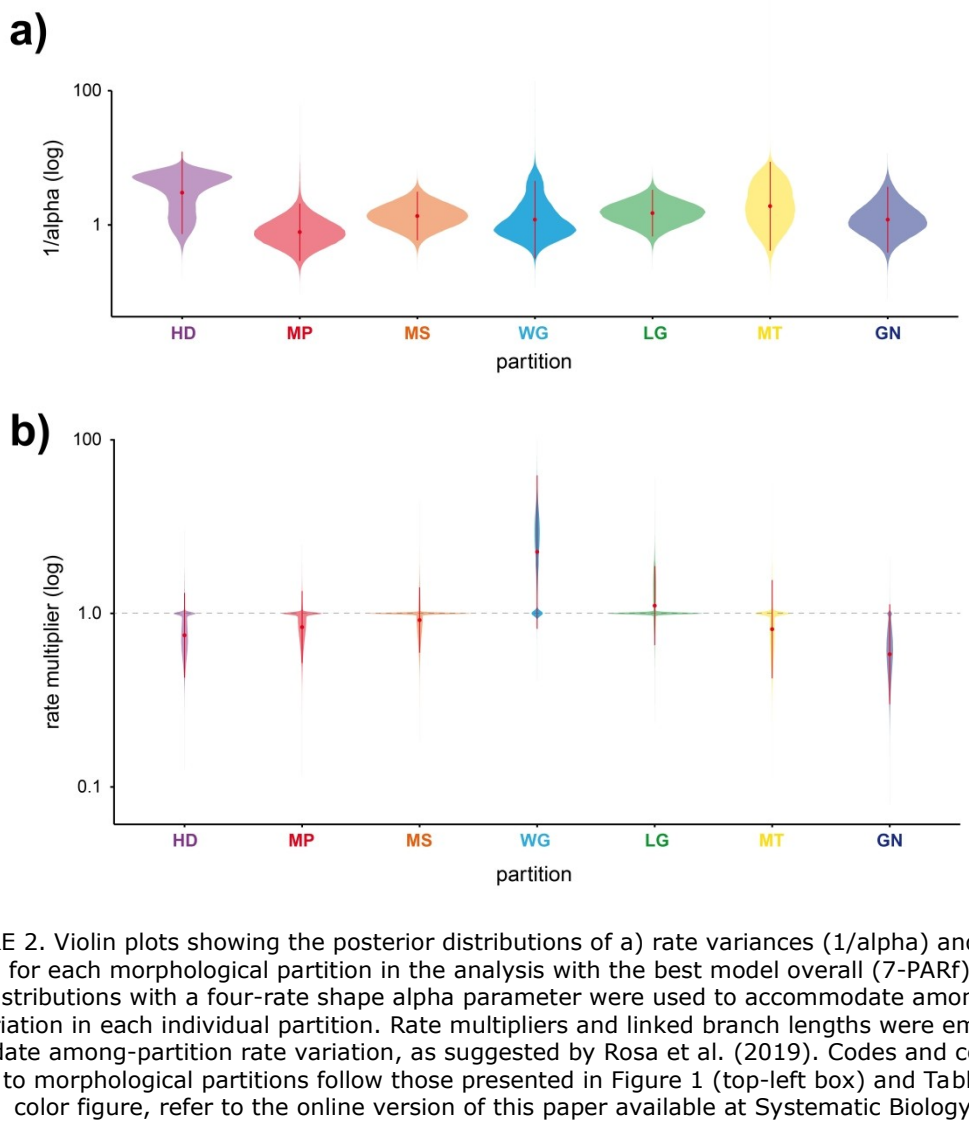

FIGURE 2. Violin plots showing the posterior distributions of a) rate variances (1/alpha) and b) rate multipliers for each morphological partition in the analysis with the best model overall (7-PARf). Discretized gamma distributions with a four-rate shape alpha parameter were used to accommodate among-character rate variation in each individual partition. Rate multipliers and linked branch lengths were employed to accommodate among-partition rate variation, as suggested by Rosa et al. (2019). Codes and colors making reference to morphological partitions follow those presented in Figure 1 (top-left box) and Table 1. For the color figure, refer to the online version of this paper available at Systematic Biology.

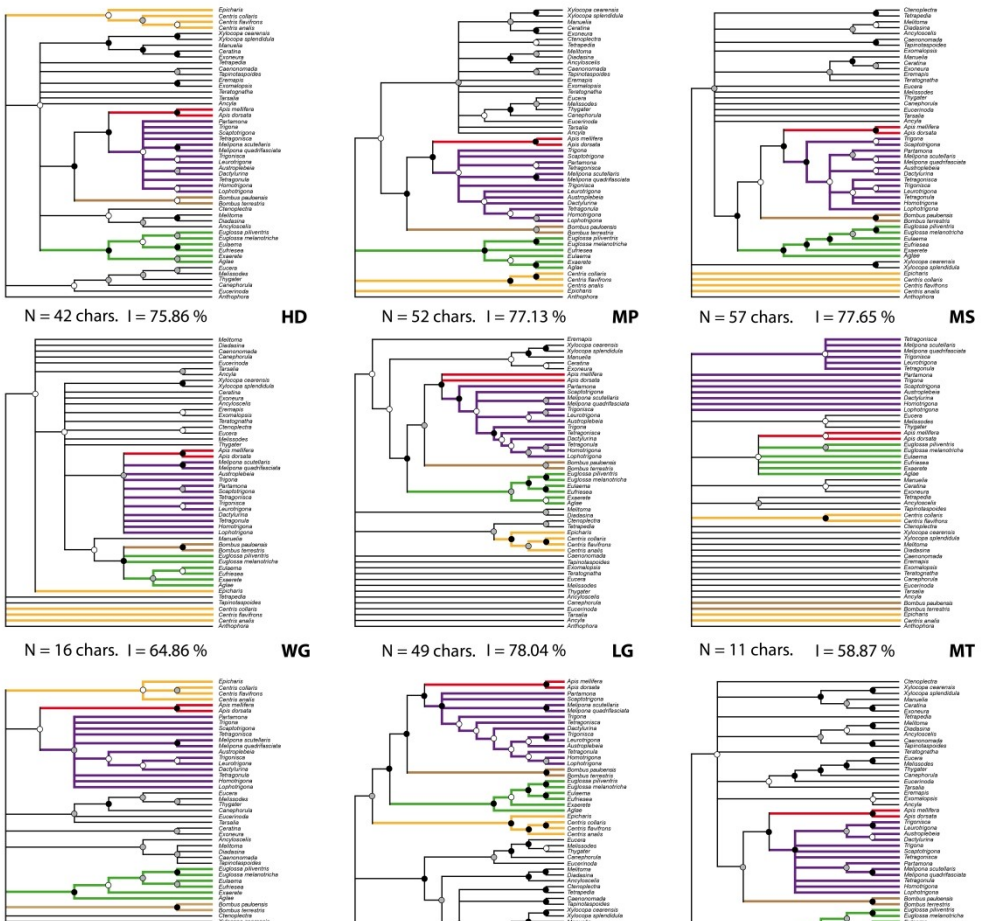

FIGURE 3. Tree topologies obtained from analyses of each individual morphological partition. Codes for partitions follow those of Table S1 and Figure S1 and are indicated in the bottom-right corner of each tree. Color-coding of branches and clades indicating certain bee tribes follows Figure 1. Circles: black denotes clades with posterior probability greater than 0.90; grey stands for clades with posterior between 0.90 and 0.75; white indicates clades with posterior between 0.75 and 0.50. Abbreviations: N: number of characters in each partition; I: total information estimated in favor of a given tree topology expressed as a percentage of the maximum. For the color figure, refer to the online version of this paper available at Systematic Biology.