Molecular Phylogenies Map to Biogeography Better than Morphological Ones

Jack Oyston (jwo22@bath.ac.uk)
The University of Bath
Mark Wilkinson
Natural History Museum, London
Marcello Ruta
University of Lincoln
Matthew Wills
University of Bath

Article

Keywords: Molecular Phylogenies Map, Biogeography, Phylogenetic relationships

Posted Date: September 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-846195/v1

Version of Record: A version of this preprint was published at Communications Biology on May 31st, 2022. See the published version at https://doi.org/10.1038/s42003-022-03482-x.

Molecular phylogenies map to biogeography better than morphological ones

Jack W. Oyston*(jwo22@bath.ac.uk), Mark Wilkinson, Marcello Ruta & Matthew A. Wills*(bssmaw@bath.ac.uk)

Abstract

Phylogenetic relationships are inferred principally from two classes of data: morphological and molecular. Most current phylogenies of extant taxa are inferred from molecules, and when morphological and molecular trees conflict the latter are often preferred. Although supported by simulations, the superiority of molecular trees has never been assessed empirically. Here we test phylogenetic accuracy using two independent data sources: biogeographical distributions and fossil first occurrences. For 48 pairs of morphological and molecular trees, we show that molecular trees are, on average, significantly more biogeographically congruent than their morphological counterparts. We also report an increase in the biogeographical congruence of phylogenies over research time. We find no significant differences in stratigraphical congruence between morphological and molecular trees. These findings have implications for understanding homoplasy in morphological data sets, the utility of morphology as a test of molecular hypotheses, and the difficulty of analysing fossil groups for which molecular data are unavailable.

Main

Phylogenies are essential in many areas of analytical biology[1], being widely utilised in evolution[2,3], ecology[4], conservation[5], parasitology[6] and medicine[7]. But what is the best way to produce an accurate phylogeny? Historically[8], putative natural groups and their interrelationships were inferred mainly from morphology, until the advent of affordable molecular sequencing technologies in the 1990s[9]. Since then, molecular sequences and phylogenomic data have overtaken morphology as the preferred basis for phylogenetic inference across extant taxa[10].

Studies of homoplasy and convergence[11] demonstrate that morphological similarity can sometimes be a poor proxy for evolutionary relationships[12]. While phylogenetic hypotheses derived from morphology are sometimes supported by molecular data[13], molecules have overturned many long-standing morphological hypotheses[14]. For example, the traditional placental mammal phylogeny was overhauled by phylogenomic analyses[15] that consistently supported deep relationships at odds with those inferred from morphology[16]. These new major clades of mammals (e.g. Afrotheria, Atlantogenata, Boreoeutheria, Laurasiatheria)[17] are more geographically congruent than those retrieved in morphological trees, and are named accordingly. While some argue that molecules should invariably have primacy[18], morphological and molecular data can be reciprocally illuminating, as shown in large-scale phylogenies of arthropods[19], reptiles and birds[20]. This balanced approach is now common in systematics[21,22].

In the absence of known phylogenies, there can be no definitive assessment of accuracy[23,24]. Here we compare conflicting trees using two independent sources of data, namely biogeographical distribution and first stratigraphic occurrences. Before the cladistic revolution, biogeography was more commonly used to infer the relationships of extant taxa[25,26]. Although highest congruence with stratigraphy can be used as an ancillary criterion to choose between equally optimal trees for groups with a good fossil record, neither biogeographical[27] nor stratigraphic data[28-30] are routinely used to infer phylogeny.

Since Wallace and Darwin, observations on the geographical distributions of species have underpinned the development of evolutionary theory[31]. Numerous studies have demonstrated an association between phylogeny and geographical distribution[32], and phylogenies are routinely used to test biogeographical hypotheses[26,33]. Here, we use biogeographical congruence as an ancillary test of competing phylogenetic hypotheses using a sample of 48 matched pairs of morphological and molecular trees representing animals and plants at multiple taxonomic levels. Biogeographical occurrence point data for extant terminals were compiled from The IUCN Red List of Threatened Species, Version 2019-2[34], the Global Biodiversity Information Facility (GBIF)[35] and The Reptile Database[36]. These distributions were used to define regions consistent with the resolution of the occurrence data, combining adjacent regions that contain identical taxon sets (see Methods and Supplementary Methods). Regional distributions were summarised as a matrix of characters coding for the presence/absence of each taxon in each region (Fig. 1). The fit of these biogeographical characters onto both morphological and molecular trees was assessed using the ensemble Consistency Index (CI) and Retention Index (RI), as well as a modified version of the Homoplasy Excess Ratio[37], the Biogeographical HER (bHER), derived from 10,000 random reassignments of biogeographical distribution data across terminals (see Methods). Importantly, our approach controls for differences in tree size and balance to the extent that these influence our indices of fit.

Results

Phylogenies tend to be significantly congruent with biogeography

The overall congruence of phylogenies with biogeographical data is good: 54% of morphological and 65% of molecular trees have a better fit than random with $p<0.05$ (and 69% of groups have one or both trees with $p<0.05$). Biogeography and phylogeny are often thought to be correlated for major clades at large geographical scales (e.g., the distribution of placental

mammal orders on continents [19]: Fig. 2a), and we find compelling evidence for similar patterns at other taxonomic levels and geographical scales (Fig. 2b). Our findings support the use of biogeographical distribution data as an ancillary criterion for choosing between otherwise equally optimal trees, similar to the widespread practice adopted for stratigraphical congruence[38].

Molecular trees are more congruent with biogeography than morphological trees

Overall, biogeographical congruence is higher for our sample of molecular trees than for their morphological counterparts (means of 0.322 vs. 0.305 for CI; 0.263 vs 0.228 for RI; 0.188 vs 0.121 for bHER). Morphological and molecular phylogenies show significantly different distributions of CI (Wilcoxon paired signed ranks test: V = 305, p = 0.027), RI (V = 295, p = 0.0199) and bHER (V = 288, p = 0.002) across the 48 pairs of trees, with molecular trees being superior on average for each index (Fig. 3). Sign tests also show selecting the tree from each pair with highest biogeographical congruence results in significantly more molecular trees being selected than their morphological counterparts (Fig. 4). Our samples of molecular and morphological trees do not differ significantly in balance (how symmetrical or pectinate they are), the degree to which CI/RI differed from random or any stratigraphical congruence measure tested (Table 1). The bHER is our preferred index, since it controls for tree size, balance and the number of biogeographical regions. Considering only the 28 groups with significantly structured ($p<0.05$) region matrices (Supplementary Table 14), we recover a similar result for bHER (Wilcoxon paired signed ranks test: V = 101, p = 0.020). In order to further ensure that the observed differences in congruence were not the result of conflating factors, we also modelled CI, RI and bHER as a function of tree size (the number of terminals), tree resolution (number of internal nodes / (number of terminals – 2)), tree balance (using Colless's index [39]) and the number of geographical regions recognised (Table 2). Residuals from these models (and from minimum adequate models selected by the AIC) show a similar pattern to those above, with residual CI, RI and bHER all demonstrating better biogeographical congruence for molecular trees, and these differences being significant for the residual CI (V

= 887, p = 0.001) and bHER (V = 751, p = 0.048). Morphological trees contain more polytomies (Supplementary Table 2) and significantly fewer resolved nodes (Supplementary Table 16), but there is still a significant difference between morphological and molecular bHER when groups with polytomous morphological trees were omitted (n = 16, V = 179, p = 0.01459).

Morphological and molecular trees have similar stratigraphical congruence

Of our 48 pairs of morphological and molecular trees, 23 had at least 50% of terminals with a fossil record, and these were assessed for stratigraphical congruence (Methods and Supplementary Methods). Our preferred index is the modified Gap Excess Ratio (GER*) [28], since it is relatively insensitive to differences in tree shape (balance), tree size, and the distribution of first occurrence dates (although the latter two variables are constant for each of our pairs). Morphological and molecular trees (Supplementary Figure 13) had similar GER* values overall (0.774 and 0.780 respective means; 0.871 and 0.871 respective medians), and a Wilcoxon signed-rank test revealed no significant difference between the distributions of GER* values (V = 79, p = 0.925). We note that the highest stratigraphical congruence is more often for morphological (n = 10) than molecular trees (n = 7) (Supplementary Figure 14), but this difference is not significant (sign test, p = 0.134). We observed similar results for the Gap Excess Ratio (GER, V = 121, p = 0.305), modified Manhattan Stratigraphic Metric (MSM*, V = 102, p = 0.486) and Stratigraphic Consistency Index (SCI, V = 59.5, p =0.159) (Table 1).

More recently published trees tend to be more biogeographically congruent

The history of systematic research is broadly one of greater volumes of data being analysed with increasingly sophisticated methods and models[40]. All other factors being equal, we might therefore expect phylogenetic accuracy to increase over research time[41]. Across all 96 morphological and molecular trees, we observed significant positive correlation between publication year and bHER (r_s = 0.257, p = 0.012) and negative correlation between publication year and p-values from our biogeographical CI and RI (r_s = -0.284, p = 0.005). Hence, more recent trees tend to have higher biogeographical congruence (Supplementary Figure 7). A

similar pattern was found for the bHER of the morphological trees considered alone (r_s = 0.292, p = 0.044), but not for the molecular trees alone (bHER, r_s = 0.184, p = 0.210: p-values from CI and RI fit, r_s = -0.274, p = 0.060). A significant minority (22 from 48) of our tree pairs had different publication dates, but we found no significant difference in the median publication years of the morphological and molecular partitions (Wilcoxon signed rank V = 32, p = 0.362). Hence, while we infer an overall improvement in phylogenetic accuracy with research time, this is not straightforwardly a function of the analysis of increasing volumes of data, preponderance of molecular analyses or potential biases in the nature of the trees in our sample.

Discussion

The significant biogeographical congruence of the majority of our clades (69% had one or both trees with CI and RI p < 0.005) supports the use of biogeographic data as an ancillary test of phylogenetic accuracy. Moreover, median biogeographical congruence for our 48 molecular trees was significantly higher than for their morphological counterparts, and this was not a function of a differential distribution of any known biases (e.g., tree size and balance). Indeed, if our results are representative, biogeographical distribution may be a better ancillary test than the established criterion of stratigraphical congruence[38]. It has been shown that morphological trees constructed using maximum parsimony often show greater stratigraphic congruence than Bayesian equivalents[42], and this effect combined with small sample size (n = 23) might explain why morphological and molecular trees in the sample we analysed did not differ significantly in their stratigraphic congruence.

Molecular data offer a number of advantages. Firstly, molecular characters can be acquired in vastly greater numbers and more economically than morphological ones, and with less taxonomic expertise[43] . Secondly, published sequence data can be readily searched, repurposed and reanalysed alongside novel sequences. Despite efforts to systematically archive morphological character matrices and character descriptions[44], there is no way to automatically produce iteratively larger morphological matrices in a manner analogous to that

possible for molecular data[45]. Thirdly, morphological systematists must make judgements concerning the homology of their characters and the manner in which they are coded[46]. Morphological variation is unlikely to be atomised in precisely the same manner by different systematists[47], whereas *a priori* rules mitigate against subjectivity and promote repeatability in molecular systematics. Nucleobases are identified using automated methods based on their molecular structure, while empirical data on substitution, insertion and deletion rates inform expectations of how likely they are to mutate. Fourthly, a well-developed body of theory and empirical data facilitate sophisticated models of molecular evolution[48], while mathematical models for morphological evolution are still in their infancy[49,50].

All other things being equal, where molecular and morphological data yield conflicting trees, our results suggest that molecular trees are likely to be more accurate. Phylogenetic signals across multiple gene alignments are typically much stronger, and lead to higher bootstrap branch support and posterior probabilities than signals from morphology[51]. Most morphological characters are two-state and may be more prone to saturation than nucleotides. Convergence in morphological character states is also common[52], even though they pass conventional tests of homology[53] and may have persisted in the literature for decades[54]. Despite these expectations, we find several cases where morphological trees have better fit than their molecular counterparts, such as dogs (Canidae), squirrels (Sciuridae), bats (Chiroptera, Macropodidae), conifers as a whole (Pinales) and pines (Pinaceae). However, in these cases values (and specifically bHER) are similar for both trees in the pair and only slightly higher for morphology. Members of some these clades, such as conifers and bats, are able to disperse or travel over long distances and so may have large geographic ranges which limit the number of region characters and hence impact the power of the tests employed here. Some morphological datasets may also contain many characters that have evolved in response to particular environmental conditions (e.g., the pine dataset was based on cone morphology), which may increase congruence with biogeography when the regions within the clade's range broadly correspond to these environmental zones. Some clades (e.g., Canidae)

had many more regions than taxa, indicating that the same taxa shared their range with different species in the clade in different areas (few taxa but a high number of different sets of taxa). This might indicate cases where species ranges are not continuous but have become fragmented over time, obscuring the original biogeographic signal. Other problems that may impact accuracy, such as long branch attraction and incomplete lineage sorting, are not unique to morphological data. While simulations suggest that likelihood and Bayesian analyses are more resilient to some of these issues[55], these methods are increasingly being applied to morphological data. Therefore, while either morphological or molecular trees may show better fit in particular cases, biogeographical congruence still provides a valuable ancillary test of phylogenetic accuracy.

The biogeographical distribution of extant species arises by two main processes: vicariance and dispersal[56]. Vicariance is the division of an ancestral area of sympatry by a physical barrier to create allopatric populations that may ultimately speciate. Pure vicariance results in phylogenetic and geospatial patterns that are entirely congruent, assuming that the geological histories of the areas are taken into account[57]. Dispersal, by contrast, implies the migration or diffusion of individuals from some centre of endemism. This can ultimately lead to speciation through physical or reproductive isolation[58]. Such patterns are typical where there are, for example, repeated raftings or other migrations of species away from an island or some other reservoir[59], as well as biotic interchanges resulting from the merging of landmasses[60]. Dispersal tends to yield area cladograms in which the colonists repeatedly derive from a paraphyletic or polyphyletic centre of endemism[58], and hence there will be less consilience between species and area cladograms. Species distribution patterns are rarely purely vicariant or purely dispersive, but usually the result of both mechanisms[61] and may be shaped by related phenomena such as range expansions[62], migrations[63] and extinctions[64]. Our tests make no assessment of the relative contributions of these processes, and the geographic regions assessed here are analogous to the areas that would form the starting point of any traditional study of historical biogeography using area cladograms[65]. Groups with limited dispersal ability,

223 a stepping-stone pattern of colonisation or which are subject to strong vicariant barriers might
224 be expected to show greater congruence between phylogeny and geography than groups that
225 migrate extensively, have large ranges or are distributed over relatively homogenous or
226 contiguous habitats. While all of our indices would be likely to yield higher values for a purely
227 vicariant than a purely dispersive pattern, both mechanisms are likely to yield non-random
228 distributions and there is no reason why morphological or molecular trees should be
229 preferentially more congruent with either pattern. It is possible, however, that selection
230 pressures that cause similar adaptations to evolve in similar environments might result in a
231 bias in favour of morphological trees where 'convergent' geographical transitions have
232 occurred.

233 Despite the superiority of molecular trees, the reciprocal illumination of morphological and
234 molecular data[66] and the simultaneous "total evidence" analysis of multiple data types has
235 been instrumental in resolving the deep relationships of many otherwise recalcitrant clades
236 including arthropods[19], echinoderms[67], angiosperms[68] and embryophytes[69]. Approximately
237 98% of species are extinct, and morphology remains the only source of data for those species
238 known only from fossils[70]. Fossils often realise combinations of character states that are
239 unknown from the extant biota[71], sample otherwise extinct or sparsely populated branches of
240 the tree, and preserve the order in which character states have evolved and thereby enable a
241 better understanding of evolutionary transitions (e.g., fish-tetrapod transition[72] or theropod-
242 bird transition[73]). A better understanding of morphological evolution and fossilisation biases[74]
243 is key to calibrating molecular trees[75] because despite the development of increasingly
244 sophisticated clock models[76], there is often a paucity of good fossil calibration dates[77]. We
245 therefore hope that our study will stimulate further ancillary biogeographical and stratigraphical
246 tests of phylogenies inferred from a variety of morphological, molecular and combined data
247 sets using different methodologies.

248

249

Methods

Dataset Compilation

We initially obtained 106 animal and plant phylogenetic trees from 61 papers published between 1981 and 2015. These were reduced to 48 pairs of morphological and molecular trees for the same clades (Supplementary Table 1), derived from the same paper whenever possible. Trees with the greatest possible overlap in taxon sets were selected, subsequently pruning unique leaves in order to yield identical taxon sets (46% of trees had different sources, 24% of trees had one or more taxa pruned, and these had a mean of 63% of leaves pruned). The majority of clades (73%) were terrestrial and freshwater vertebrates with strong patterns of endemism, but insect (13%) and plant (15%) clades were also included. Only 10% of clades contained any marine taxa, partly a function of the difficulties of accurately ascertaining and coding regions in these environments.

Coding Biogeographical Distributions

Biogeographical data were obtained from The IUCN Red List of Threatened Species, Version 2019-2 (http://www.iucnredlist.org), the Global Biodiversity Information Facility, 29th Dec 2019 (https://www.gbif.org) and The Reptile Database, 24 Dec 2019 (http://www.reptile-database.org) (Supplementary Table 3). Biogeographical data were collected in two forms: taxon presences defined at the highest resolution available (state/district/province: e.g., 'California', country: 'U.S.A.' or larger geographic region: 'North America') and point occurrences and recorded sightings. Point occurrences were synthesised into a list of presences for areas equivalent to the highest resolution of areas available in the online database. Our approach to coding was inclusive insofar as taxa known from multiple regions were coded as present in all of these regions. For each clade, lists were combined to create a biogeographical character matrix of presence/absence characters for each recognised region (column). A matrix of characters, rather than a single multistate character, allowed for taxa that were observed from more than one region. As the areas being combined are often

defined geopolitically or represent the limited spatial resolution of data, regions derived from them are only biogeographically meaningful if they contain unique information about how taxa are grouped in space. Therefore, to avoid over-splitting of regions, we combined pairs of closest geographically neighbouring regions with identical taxon presence/absences into a single larger region and continued this process until all regions had unique taxon presence/absences. As it was not uncommon for biogeographical region matrices to contain more regions than taxa after this process (as a difference in presence for one taxon is sufficient to define a distinct region) we merged regions with single unique taxa (autapomorphic region characters) into their geographically closest neighbours. Each resulting biogeographical region matrix was assessed for non-random structure using Monte Carlo permutation tail probability (MCPTP) compatibility tests[78] (Supplementary Methods). Two characters are incompatible if it is not possible to map them onto the same evolutionary tree without homoplasy. Fewer incompatibilities indicate a more highly structured character matrix which is more likely to be phylogenetically informative.

Testing Biogeographical Congruence

We assessed the fit of the biogeographical matrices onto both morphological and molecular trees using the ensemble consistency index (CI) and ensemble retention index (RI). We note that the CI is biased by tree size, and by tree shape and balance with certain types of characters[79] (e.g., irreversible and ordered). We therefore also measured congruence using a modification of the homoplasy excess ratio (HER) of Archie[37]. Our biogeographical HER (bHER) was calculated by comparing the additional step length (over and above the minimum necessary – MINL – for the number and nature of characters) for our observed data (L) with the mean additional step length for biogeographically random data (MEANNS) (randomly reassigning rows in the data matrix to the taxa 10,000 times, while holding tree topology constant). The bHER (or, more precisely, our modified MEANNS) therefore differed from the HER in its original form by permuting rows of the matrix across taxa (rather than the entries within each column separately) and by calculating the length of the original and permuted

biogeographical matrices on the morphological or molecular tree (rather than inferring a tree from these data). Specifically, bHER = 1-(L-MINL)/(MEANNS-MINL) (see Supplementary Materials for full details). A similar procedure was also used to produce a distribution of randomised tree length values, against which the original tree length could be compared in order to yield approximate p-values (the probability that a length as short or shorter could be observed for biogeographical data distributed at random on the tree). This is equivalent to a randomisation test for both CI and RI and will yield the same p-values for both metrics by definition. By permuting rows of codes across taxa (rather than each column of data across taxa independently), we ensured that there were no unrealised or unlikely combinations of regional distribution patterns. The bHER, CI, RI and the p-values from CI/RI randomisation tests for morphological and molecular tree samples were compared using paired Wilcoxon signed-rank tests. In addition, sign tests were used to test whether selecting the most biogeographically congruent tree in each pair resulted in significantly more molecular or morphological trees being chosen than expected by chance.

Testing Stratigraphical Congruence

Data on the fossil record of each of the 48 clades in this study were collated from the Palaeobiology database (PBDB) and Benton (1993), as well as data within the source papers (Supplementary Methods). 23 Clades had published fossil data for at least 50% of their leaves, and so were judged suitable for tests of stratigraphical congruence. First and last occurrences were assigned at the stage-level, after O'Connor *et al.*, (2016)[38]. Low preservation potential and scarcity often ensure that first fossil occurrences lag behind true times of origin, while scarcity prior to the actual point of extinction mean that lineages are lost from the record prematurely (the 'Signor-Lipps effect'). Where stratigraphy was unresolved at the stage level, taxa were therefore assigned to the first stage in the time interval given for their first occurrence and the last interval of the time period for their last occurrence. Stratigraphical congruence was assessed using a number of previously published and commonly utilised metrics, namely the stratigraphic consistency index (SCI), modified Manhattan stratigraphic measure (MSM*),

the gap excess ratio and its modification (GER and GER*). The stratigraphical congruence of morphological and molecular trees was assessed using paired Wilcoxon signed-rank tests as well as sign tests, in a similar manner to that detailed for the biogeographical congruence tests.

Data Availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

References

1 Harvey, P. H. & Pagel, M. D. *The Comparative Method in Evolutionary Biology*. (Oxford University Press, 1991).

2 Oyston, J. W., Hughes, M., Wagner, P. J., Gerber, S. & Wills, M. A. What limits the morphological disparity of clades? *Interface Focus* **5**, 20150042 (2015).

3 Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).

4 Webb, C. O. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist* **156**, 145-155 (2000).

5 Purvis, A., Gittleman, J. L. & Brooks, T. (Eds.) *Phylogeny and Conservation*. (Cambridge University Press, 2005).

6 Page, R. D. M. Parallel phylogenies: reconstructing the history of host-parasite assemblages. *Cladistics* **10**, 155-173 (1994).

7 Weaver, S. C. & Vasilakis, N. Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infection, Genetics and Evolution* **9**, 523-540 (2009).

8 Tassy, P. Trees before and after Darwin. *Journal of Zoological Systematics and Evolutionary Research* **49**, 89-101 (2011).

9 Heather, J. M. & Chain, B. The sequence of sequencers: The history of sequencing DNA. *Genomics* **107**, 1-8 (2016).

10 Pyron, R. A. Post-molecular systematics and the future of phylogenetics. *Trends in Ecology & Evolution* **30**, 384-389 (2015).

11 Wake, D. B., Wake, M. H. & Specht, C. D. Homoplasy: from detecting pattern to determining process and mechanism of evolution. *Science* **331**, 1032-1035 (2011).

12 Sansom, R. S. & Wills, M. A. Differences between hard and soft phylogenetic data. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20172150 (2017).

13 Fernandez, R., Edgecombe, G. D. & Giribet, G. Phylogenomics illuminates the backbone of the Myriapoda Tree of Life and reconciles morphological and molecular phylogenies. *Scientific Reports* **8**, 1-7 (2018).

14 Eme, L., Spang, A., Lombard, J., Stairs, C. W. & Ettema, T. J. Archaea and the origin of eukaryotes. *Nature Reviews Microbiology* **15**, 711 (2017).

15 Asher, R. J., Bennett, N. & Lehmann, T. The new framework for understanding placental mammal evolution. *Bioessays* **31**, 853-864 (2009).

16 Shoshani, J. & McKenna, M. C. Higher taxonomic relationships among extant mammals based on morphology, with selected comparisons of results from molecular data. *Molecular Phylogenetics and Evolution* **9**, 572-584 (1998).

17 Beck, R. M. & Baillie, C. Improvements in the fossil record may largely resolve current conflicts between morphological and molecular estimates of mammal phylogeny. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20181632 (2018).

18 Scotland, R. W., Olmstead, R. G. & Bennett, J. R. Phylogeny reconstruction: The role of morphology. *Systematic Biology* **52**, 539-548 (2003).

19 Regier, J. C. *et al.* Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**, 1079–1083 (2010).

20 Callender-Crowe, L. M. & Sansom, R. S. Osteological characters of birds and reptiles are more congruent with molecular phylogenies than soft characters are. *Zoological Journal of the Linnean Society* **XX**, 1-13, doi:10.1093/zoolinnean/zlaa136 (2021).

21 Wahlberg, N. *et al.* Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. *Proceedings of the Royal Society B: Biological Sciences* **272**, 1577-1586 (2005).

22 He, L. *et al.* A molecular phylogeny of selligueoid ferns (Polypodiaceae): Implications for a natural delimitation despite homoplasy and rapid radiation. *Taxon* **67**, 237-249 (2018).

23 Zou, Z. & Zhang, J. Morphological and molecular convergences in mammalian phylogenetics. *Nature Communications* **7**, 12758, doi:10.1038/ncomms12758 (2016).

24 Hillis, D. Approaches for assessing phylogenetic accuracy. *Systematic Biology* **44**, 3-16 (1995).

25 Thompson, N. *Alfred Russell Wallace Contributions to the theory of Natural Selection, 1870, and Charles Darwin and Alfred Wallace,'On the Tendency of Species to form Varieties'(Papers presented to the Linnean Society 30th June 1858)*. (Routledge, 2004).

26 Croizat, L. Panbiogeography; or an Introductory Synthesis of Zoogeography, Phytogeography, and Geology, Vols. 1, 2a, 2b. (private publication, 1958).

27 Means, J. C. & Marek, P. E. Is geography an accurate predictor of evolutionary history in the millipede family Xystodesmidae? *Peerj* **5**, e3854 (2017).

28 Wills, M. A., Barrett, P. M. & Heathcote, J. F. The modified Gap Excess Ratio (GER*) and the stratigraphic congruence of dinosaur phylogenies. *Systematic Biology* **57**, 891-904 (2008).

29 Fisher, D. C. Stratocladistics: integrating temporal data and character data in phylogenetic inference. *Annual Review of Ecology, Evolution, and Systematics* **39**, 365-385 (2008).

30 Lazarus, D. B. & Prothero, D. R. The role of stratigraphic and morphologic data in phylogeny. *Journal of Paleontology* **58**, 163-172 (1984).

31 Camerini, J. R. Evolution, biogeography, and maps: An early history of Wallace's Line. *Isis* **84**, 700-727 (1993).

32 Upchurch, P., Hunn, C. A. & Norman, D. B. An analysis of dinosaurian biogeography: evidence for the existence of vicariance and dispersal patterns caused by geological events. *Proceedings of the Royal Society B: Biological Sciences* **269**, 613-621 (2002).

33 Ferreira, G. S., Bronzati, M., Langer, M. C. & Sterli, J. Phylogeny, biogeography and diversification patterns of side-necked turtles (Testudines: Pleurodira). *Royal Society Open Science* **5**, 1-17 (2018).

34 IUCN (2019). *The IUCN Red List of Threatened Species*, Version 2019-2. https://www.iucnredlist.org. Downloaded on [15/11/2019]. (2014).

35 GBIF.org (2019). *GBIF Home Page.* https://www.gbif.org. Downloaded on [29/12/2019]. (2016).

36 Uetz, P., Freed, P., Aguilar, R. & Hošek, J. (Eds.) (2021). *The Reptile Database*. http://www.reptile-database.org, Downloaded on [24/12/2019]. (2012).

37 Archie, J. W. Homoplasy excess ratios: new indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index. *Systematic Zoology* **38**, 253-269 (1989).

38 O'Connor, A. & Wills, M. A. Measuring stratigraphic congruence across trees, higher taxa, and time. *Systematic Biology* **65**, 792-811 (2016).

39 Colless, D. H. Review of phylogenetics: the theory and practice of phylogenetic systematics. *Systematic Zoology* **31**, 100-104 (1982).

40 Lartillot, N. & Philippe, H. Improvement of molecular phylogenetic inference and the phylogeny of Bilateria. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 1463-1472 (2008).

41 Beck, R. M. D. & Baillie, C. Improvements in the fossil record may largely resolve current conflicts between morphological and molecular estimates of mammal phylogeny. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20181632 (2018).

42 Sansom, R. S., Choate, P. G., Keating, J. N. & Randle, E. Parsimony, not Bayesian analysis, recovers more stratigraphically congruent phylogenetic trees. *Biology Letters* **14**, 20180263 (2018).

43 Wiens, J. J. The role of morphological data in phylogeny reconstruction. *Systematic Biology* **53**, 653-661 (2004).

44 O'Leary, M. A. & Kaufman, S. G. (2012). *MorphoBank 3.0: Web application for morphological phylogenetics and taxonomy.* http://www.morphobank.org.

45 de Queiroz, A. & Gatesy, J. The supermatrix approach to systematics. *Trends in Ecology & Evolution* **22**, 34-41 (2007).

46 Wilkinson, M. A comparison of two methods of character construction. *Cladistics* **11**, 297-308 (1995).

47 Brazeau, M. D. Problematic character coding methods in morphology and their effects. *Biological Journal of the Linnean Society* **104**, 489-498 (2011).

48 Drummond, A. J., Ho, S. Y., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**, e88, https://doi.org/10.1371/journal.pbio.0040088 (2006).

49 O'Reilly, J. E., Puttick, M. N., Pisani, D. & Donoghue, P. C. Probabilistic methods surpass parsimony when assessing clade support in phylogenetic analyses of discrete morphological data. *Palaeontology* **61**, 105-118 (2018).

50 Keating, J. N., Sansom, R. S., Sutton, M. D., Knight, C. G. & Garwood, R. J. Morphological phylogenetics evaluated using novel evolutionary simulations. *Systematic Biology* **69**, 897-912 (2020).

51 Makarenkov, V. *et al.* Weighted bootstrapping: a correction method for assessing the robustness of phylogenetic trees. *BMC Evolutionary Biology* **10**, 1-16 (2010).

52 Stayton, C. T. The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. *Evolution* **69**, 2140-2153 (2015).

53 Sattler, R. Homology - a continuing challenge. *Systematic Botany* **9**, 382-394 (1984).

54 Jenner, R. A. & Schram, F. R. The grand game of metazoan phylogeny: rules and strategies. *Biological Reviews* **74**, 121-142 (1999).

55 Swofford, D. L. *et al.* Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Systematic Biology* **50**, 525-539 (2001).

56 Jaeger, J. J. & Martin, M. African marsupials - vicariance or dispersion? *Nature* **312**, 379-379 (1984).

57 Smith, B. T. *et al.* The drivers of tropical speciation. *Nature* **515**, 406-409 (2014).

58 Raxworthy, C. J., Forstner, M. R. J. & Nussbaum, R. A. Chameleon radiation by oceanic dispersal. *Nature* **415**, 784-787 (2002).

59 Simkanin, C. *et al.* Exploring potential establishment of marine rafting species after transoceanic long-distance dispersal. *Global Ecology and Biogeography* **28**, 588-600 (2019).

60 Stehli, F. G. & Webb, S. D. *The Great American Biotic Interchange*. Vol. 4 (Springer Science & Business Media, Berlin/Heidelberg, 2013).

61 Ronquist, F. Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Systematic Biology* **46**, 195-203 (1997).

62 Ricklefs, R. E. & Bermingham, E. The concept of the taxon cycle in biogeography. *Global Ecology and Biogeography* **11**, 353-361 (2002).

63 Ma, H. An analysis of the equilibrium of migration models for biogeography-based optimization. *Information Sciences* **180**, 3444-3464 (2010).

64 Yiming, L., Niemelä, J. & Dianmo, L. Nested distribution of amphibians in the Zhoushan archipelago, China: can selective extinction cause nested subsets of species? *Oecologia* **113**, 557-564 (1998).

65 Crisci, J., Katinas, L., Posadas, P. & Crisci, J. V. *Historical Biogeography: an Introduction*. (Harvard University Press, 2009).

66 Dillman, C. B. & Hilton, E. J. The cause and effect of polarization: thoughts on the "morphological vs. molecular Debate" in systematics, with examples from the study of sturgeons (Actinopterygii: Acipenseridae). *Zootaxa*, 79-117 (2011).

67 Smith, A. B. Echinoderm phylogeny: morphology and molecules approach accord. *Trends in Ecology & Evolution* **7**, 224-229 (1992).

68 Bateman, R. M., Hilton, J. & Rudall, P. J. Morphological and molecular phylogenetic context of the angiosperms: contrasting the 'top-down' and 'bottom-up' approaches used to infer the likely characteristics of the first flowers. *Journal of Experimental Botany* **57**, 3471-3503 (2006).

69 Morris, J. L. *et al.* The timescale of early land plant evolution. *Proceedings of the National Academy of Sciences* **115**, E2274-E2283, doi:10.1073/pnas.1719588115 (2018).

70 Newman, M. E. J. A model of mass extinction. *Journal of Theoretical Biology* **189**, 235-252 (1997).

71 Cobbett, A., Wilkinson, M. & Wills, M. A. Fossils impact as hard as living taxa in parsimony analyses of morphology. *Systematic Biology* **56**, 753-766 (2007).

72 Ruta, M., Krieger, J., Angielczyk, K. & Wills, M. A. The evolution of the tetrapod humerus: morphometrics, disparity, and evolutionary rates. *Earth and Environmental Sciences Transactions of the Royal Society of Edinburgh* **109**, 351-369 (2018).

73 Puttick, M. N., Thomas, G. H. & Benton, M. J. High rates of evolution preceded the origins of birds. *Evolution* **68**, 1497-1510 (2014).

74 Sansom, R. S. & Wills, M. A. Fossilization causes organisms to appear erroneously primitive by distorting evolutionary trees. *Scientific Reports* **3**, 1-5 (2013).

75 Brinkworth, A., Sansom, R. & Wills, M. A. Phylogenetic incongruence and homoplasy in the appendages and bodies of arthropods: why broad character sampling is best. *Zoological Journal of the Linnean Society* **187**, 100-116 (2019).

76 Brown, J. W. & Smith, S. A. The Past Sure is Tense: On Interpreting Phylogenetic Divergence Time Estimates. *Systematic Biology* **67**, 340-353 (2018).

77 Barba-Montoya, J., dos Reis, M. & Yang, Z. H. Comparison of different strategies for using fossil calibrations to generate the time prior in Bayesian molecular clock dating. *Molecular Phylogenetics and Evolution* **114**, 386-400 (2017).

78 Wilkinson, M. Characters, congruence and quality: a study of neuroanatomical and traditional data in caecilian phylogeny. *Biological Reviews* **72**, 423-470 (1997).

79 Sanderson, M. J. & Donoghue, M. J. Patterns of variation in levels of homoplasy. *Evolution* **43**, 1781-1795 (1989).

80 O'Leary, M. A. *et al.* The placental mammal ancestor and the post–K-Pg radiation of placentals. *Science* **339**, 662-667 (2013).

81 Kluge, A. G. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Biology* **38**, 7-25 (1989).

82 Tolson, P. J. Phylogenetics of the boid snake genus Epicrates and Caribbean vicariance theory. *Occasional Papers of the Museum of Zoology: the University of Michigan* **715**, 1-68 (1987).

83 Clopper, C. J. & Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**, 404-413 (1934).

559 **Acknowledgements**

560 We thank Tim Astrop for useful discussions and suggestions related to plotting the data as
561 well as Tamás Székely, Polly Russell and Catherine Klein for useful discussions. This
562 research was supported by the John Templeton Foundation grants 43915 and 61408.

563 **Author Information**

564 Affiliations

565 **Milner Centre for Evolution, University of Bath, Bath, UK**
566 Jack W. Oyston & Matthew A. Wills

567 **Vertebrates Division, Life Sciences, Natural History Museum, London, UK**
568 Mark Wilkinson

569 **School of Life Sciences, College of Science, University of Lincoln**
570 Marcello Ruta

571 **Contributions**

572 J.W.O. and M.A.W. conceived the study, devised tests of biogeographical congruence,
573 developed the methods and theory, and wrote the paper. J.W.O. compiled and analysed the
574 data. M.W. carried out the compatibility tests and conducted the experiments, analysed the
575 data, and performed the simulations. M.W. and M.R. analysed data and contributed text.

576 **Corresponding authors**

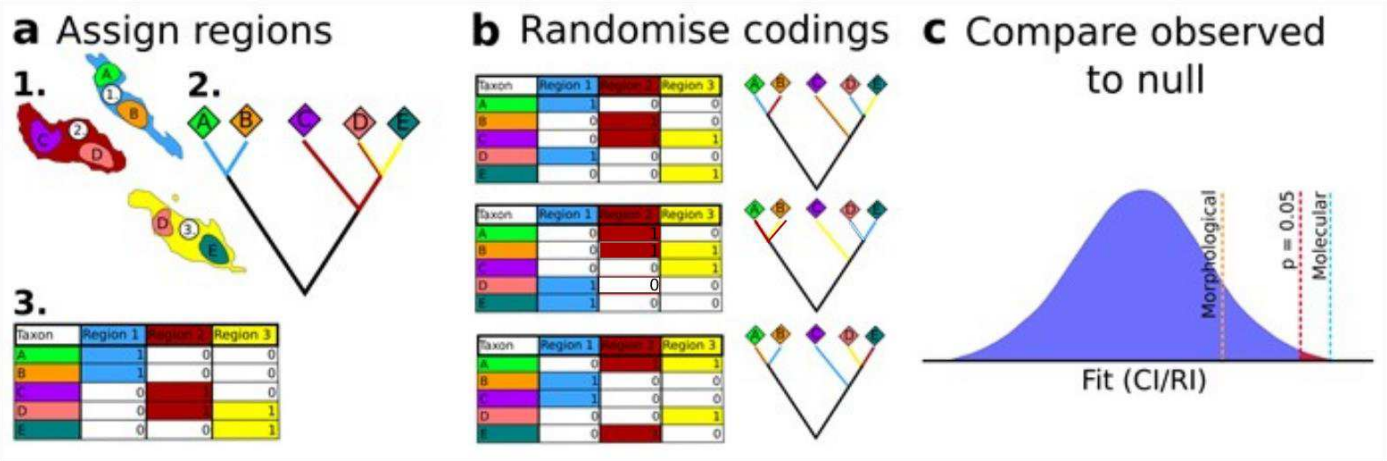
577 Correspondence to Jack W. Oyston or Matthew A. Wills.

578 **Ethics declarations**

579 Competing interests

580 The authors declare no competing interests.

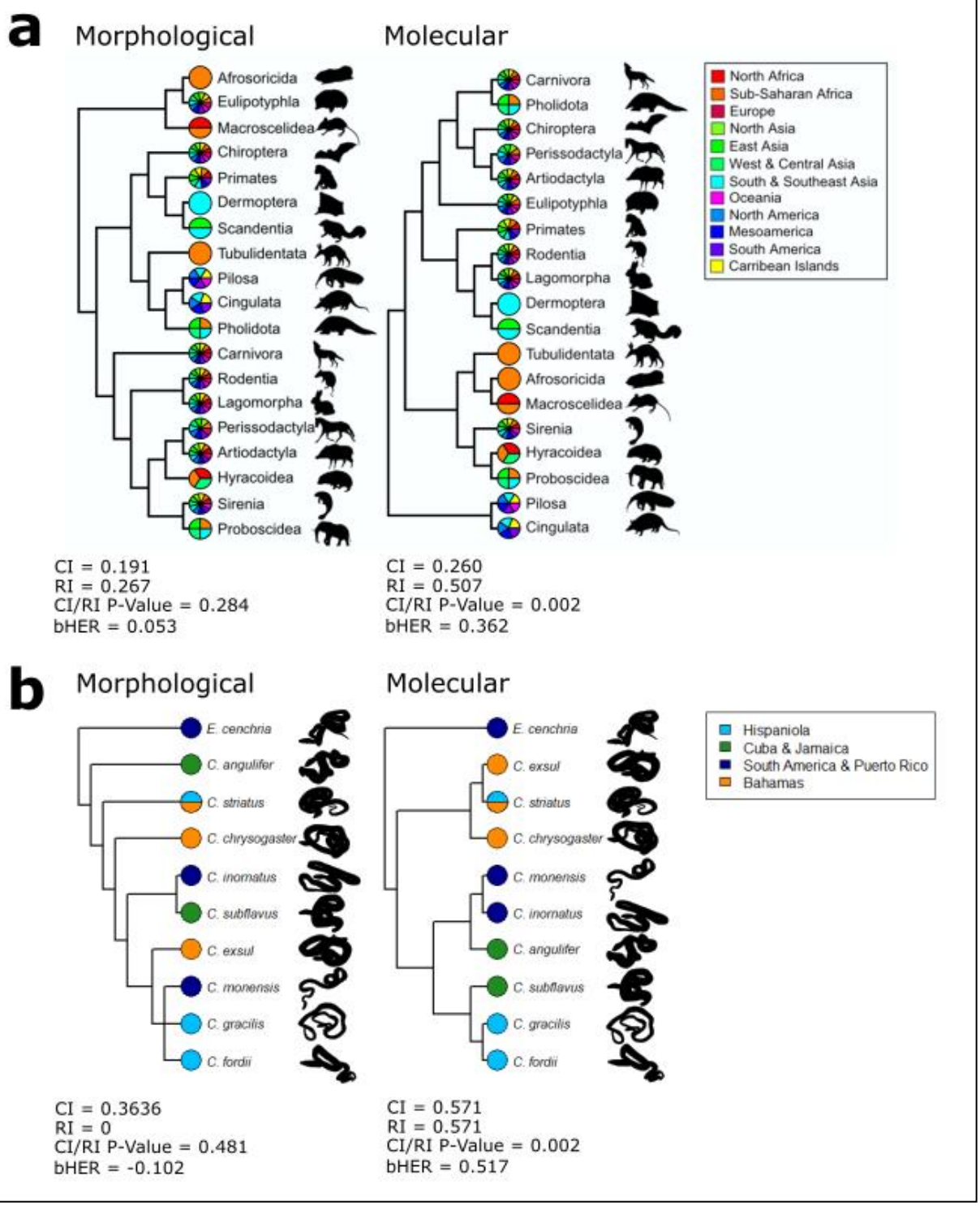
Fig. 1: Testing the biogeographical congruence of phylogenetic trees.



a Defining biogeographical regions and coding taxon presence/absences. **1.** Occurrence data on the distribution of extant species is used to produce a list of biogeographical regions for the clade and to summarise ranges for the taxa in the published phylogenies. **2.** This distributional information is converted into a matrix of binary characters representing taxa in biogeographical regions, where 0 indicates the taxon is absent and 1 indicates the taxon is present. **3.** Characters in the occurrence matrix are mapped onto the morphological and molecular phylogeny selected for each clade, allowing standard measures of character fit (CI, RI) to be calculated for each tree. **b** Presence and absence codings in each matrix are randomly reassigned to taxa, keeping the presence and absence codings fixed for each row. Characters from the new randomised matrix are mapped onto the original trees and both CI and RI are recalculated. The entire randomisation process is performed 10,000 times. **c** The 10,000 CI and RI values from matrices' biogeographical region reassignments form a null distribution of expected congruence values if taxa in the clade were randomly distributed in biogeographical regions. The observed CI and RI of region characters for a given tree is compared to the null distribution for that same tree in order to determine whether biogeographical congruence is significantly higher than expected by chance.

581

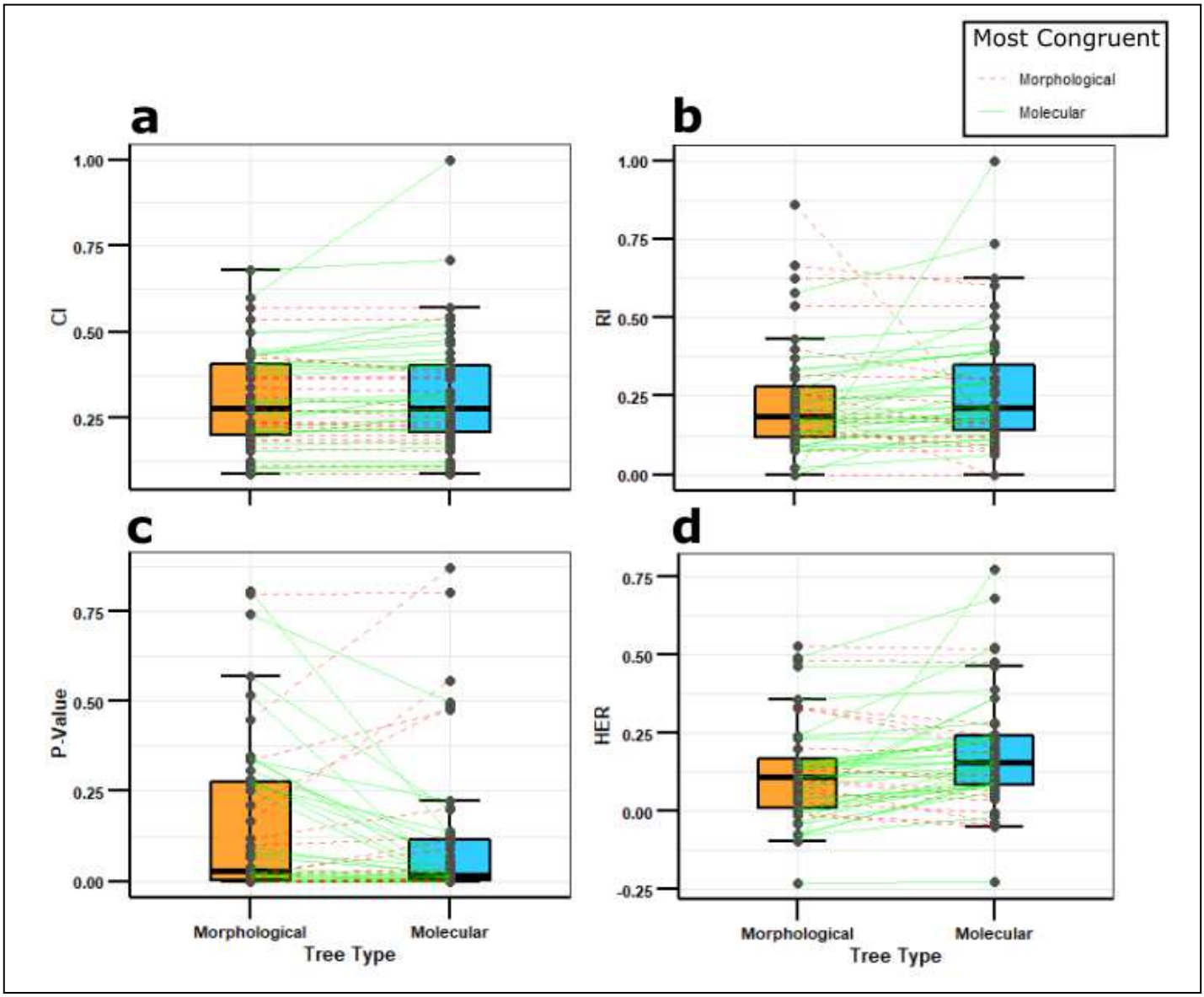
Fig. 2: Biogeographical congruence in morphological and molecular phylogenies.



Binary region characters mapped onto morphological and molecular phylogenies of **a:** placental mammals (Eutheria) from O'Leary et al. 2013[80] and **b:** Caribbean boas (*Chilabothrus*/*Epicrates*), with the morphological tree taken from Kluge 1989[81] and the molecular tree taken from Tolson 1987[82]. Regions for which the terminal taxon is coded present are represented as coloured pie slices. Consistency index (CI), retention index (RI) and biogeographical HER (bHER) values given are for the matrix of biogeographical region presences and absences, while CI and RI p-value is calculated using 10,000 randomised region matrices.

583

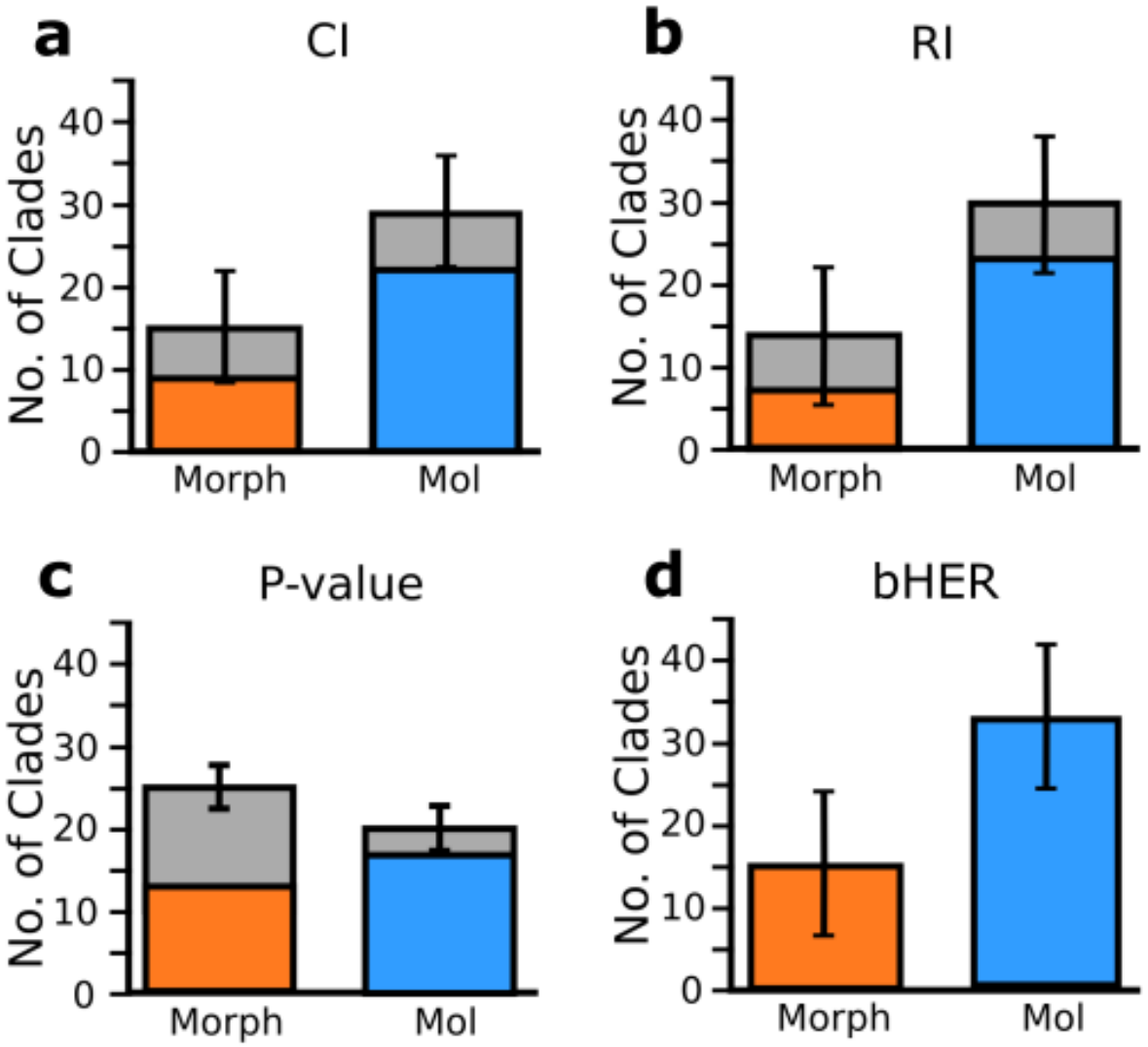
Fig. 3: Differences in biogeographical congruence between morphological and molecular trees.



Boxplots of raw values and differences in values between morphological and molecular trees for **a:** consistency index (CI: V = 305, p =0.027), **b:** retention index (RI: V = 295, p =0.020), **c:** p-values for the CI/RI randomizations (P-Value: V = 662, p =0.104) and **d:** biogeographical HER (bHER: V = 288, p =0.002). Boxes delimit the upper and lower quartiles of the data, while central bars are median values. Whiskers delimit plus or minus 1.5 times the inter-quartile range, from the first and third quartiles. Coloured lines connected pairs of values from the same clade, where red dashed lines indicate the morphological tree is most biogeographically congruent and green solid lines indicate the molecular tree is most biogeographically congruent.

584

Fig. 4: The number of morphological and molecular trees most congruent with biogeography.



Comparison of the number of trees in each sample (morphological or molecular) with a greater biogeographical fit than its counterpart. **a:** consistency index (CI), grey bars show totals for the whole sample, coloured bars indicate totals in the subset significantly different from the expected null (randomized p-value <0.05). **b:** retention index (RI), grey bars show totals for the whole sample, coloured bars indicate totals in the subset significantly different from the expected null (randomized p-value <0.05). **c:** CI/RI randomization p-values (P-value), where grey bars show totals for the whole sample, coloured bars are clades with values <0.05. **d:** biogeographical HER (bHER), counts are for the whole dataset. Bars show the number of clades in each subset, with binomial confidence intervals calculated using the approach of Clopper and Pearson (1934)[83].

Table 1: Biogeographical and stratigraphical congruence of morphological and molecular phylogenies.

586

Metric	Wilcoxon signed-rank test statistic (V)	p-Value
Publication Year	32	0.362
Phylogenetic Characters	2	<0.001
I*m*	547	0.743
CI	305	0.027
RI	295	0.020
CI/RI p-value	662	0.104
bHER	288	0.002
SCI	59.5	0.159
MSM*	102	0.486
GER	121	0.305
GER*	79	0.925

Results of paired Wilcoxon signed-rank tests on the two data partitions (Morphological & Molecular) for the following metrics: publication year, number of phylogenetic characters, Colless's index of tree balance (I*m*), consistency index (CI), retention index (RI), probability of CI & RI values falling within the null distribution (CI/RI p-value), biogeographical homoplasy excess ratio (bHER), stratigraphic consistency index (SCI), the modified Manhattan stratigraphic measure (MSM*), the gap excess ratio (GER) and the modified gap excess ratio (GER*). The sample size is 44 trees for SCI, MSM*, GER, GER* and 96 trees for all other metrics. Statistically significant results are highlighted in green.

587

588

589

590

591

592

593

594

595

596

Table 2: Biogeographical congruence metrics modelled by potential confounding variables.

Model	Linear Regression			Wilcoxon signed-rank test	
	AIC	R^2	p-value	test statistic (V)	p-value
CI ~ Taxa + Regions	-442.95	0.578	<0.001	671	0.017
CI ~ Taxa + Regions + *Im* + Res	-440.23	0.584	<0.001	867	0.001
bHER ~ *Im* + Res	-334.93	0.085	0.018	751	0.048
RI ~ Regions + Res	-332.67	0.109	0.005	529	0.901
bHER ~ Taxa + Regions + *Im* + Res	-332.14	0.097	0.058	740	0.063
RI ~ Taxa + Regions + *Im* + Res	-329.71	0.119	0.022	528	0.710

Results of models predicting measures of consistency index (CI), retention index (RI) and biogeographical homoplasy excess ratio (bHER) of geographic region characters from the log of the number of terminal taxa (Taxa), log of the number of biogeographical regions (Regions), Colless's index of tree balance (I*m*) and the proportion of resolved nodes (Res). Both the model with all explanatory variables and the model with minimal Akaike information criterion (AIC) are given for each congruence measure. Wilcoxon signed-rank tests are between model residuals for morphological and molecular trees in the sample. Statistically significant results are highlighted in green.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- RegionMatrices.zip
- JWOMolecularphylogeniesmaptobiogeographybetterthanmorphologicalonesESMComms.Biol.submission.pdf
- JWOMolecularphylogeniesmaptobiogeographybetterthanmorphologicalonesSupplementaryTable1Comms.Biol.submission.pdf
- JWOMolecularphylogeniesmaptobiogeographybetterthanmorphologicalonesSupplementaryTable3Comms.Biol.submission.pdf
- JWOMolecularphylogeniesmaptobiogeographybetterthanmorphologicalonesSupplementaryTable4Comms.Biol.submission.pdf
- Trees.zip
- StratigraphicAges.zip
- BiogeogPICA.zip