Time and space in biogeography: response to Parenti & Ebach (2013)

ABSTRACT
A recent Guest Editorial by Parenti & Ebach (2013, Journal of Biogeography, 40, 813–820) disagrees with the methods or interpretations in two of our recent papers. In addition, the authors open a debate on biogeographical concepts, and present an alternative philosophy for biogeographical research in the context of their recently described biogeographical subregion called ‘Pandora’. We disagree with their approach and conclusions, and comment on several issues related to our differing conceptual approaches for biogeographical research; namely, our use of molecular phylogenetic analyses, including time estimates; and Parenti & Ebach’s reliance on taxon/general area cladograms. Finally, we re-examine their ‘tests’ supporting the existence of ‘Pandora’.

Keywords Biogeography, cladistics, dispersal, diversification, general area cladograms, historical biogeography, molecular clock, Pandora, philosophy of science, vicariance.

Many biogeographical explanations are hampered by invoking simple notions of mechanism or process – dispersal and vicariance – or constraints, such as dispersal from a centre of origin, and, in so doing, dismiss more complex geological phenomena such as emergent volcanoes within island chains or composite areas as irrelevant. Moreover, they do not search for, therefore never discover, biogeographical patterns that may better explain the distribution of biota through time.

Parenti & Ebach (2013, p. 813) Documenting biogeographical patterns and testing hypotheses for the processes that generate patterns are challenging, as they require reconstruction of an area’s geological and biotic history. Parenti & Ebach (2013) criticize two recent papers on the origins and biogeography of the biota from the Indo-Australian Archipelago (IAA) and from Sulawesi (respectively: Lohman et al., 2011; Stelbrink et al., 2012). Parenti & Ebach (2013) assert that the answers to why and how species distributions arise are solely hypotheses. Obviously, the answers to these questions result from actual events that took place, but our interpretation of these events is hypothesized because they were not directly observed. This is as trivial as it is true. They further state that ‘treating a hypothesis as if it were evidence or an empirical observation … favour[s] one explanation over another simply because no other explanation was ever considered’ (p. 813). This is a problem for all biogeographical analyses (Beaumont et al., 2010), but we believe that palaeogeographical reconstructions by geological experts (e.g. Hall, 2009) for the IAA and Sulawesi are the most suitable framework within which to test biogeographical hypotheses, as opposed to the use of hypothetical (see below) relationships of taxa (cladograms), by biologists, to reconstruct palaeogeography. Spurious ‘patterns’ can also result from failing to sample intermediate lineages between focal taxa (Turner et al., 2009), as we show here to be the case for one of the two data sets used by Parenti & Ebach (2013) to justify the recognition of Pandora (see Appendix S1 in Supporting Information). Thus, cladograms are also hypotheses (of species relationships), which are constantly revised as additional data are collected.

In our opinion, cladistic biogeographical analyses using ‘taxon area cladograms’, or more generally, ‘general area cladograms’, are weak tests of biogeographical hypotheses. Cladograms neglect temporal information (divergence times) preserved in biological data (molecular or other), and thus ignore linkages between distribution patterns, their potential causes, and their timing (e.g. geological events: separation of a tectonic fragment from its adjacent landmass, opening of a water strait, emergence of oceanic islands, etc.). Inconsistencies in timeframes result in so-called ‘pseudo-congruence’ (see Donoghue & Moore, 2003), and ignore potentially deviating biogeographical histories across lineages/clades due to potentially different causes (e.g. geological, biological, climatic). For example, geological changes on the island of Sulawesi of particular interest to biogeographers include the opening of the Makassar Strait and separation of the Sulawesi Spur, which can be attributed to major geological changes during the middle Eocene and middle Miocene, respectively (Figure 1 in Stelbrink et al., 2012). Different organismal lineages may have different temporal histories associated with these geological events, and – while a molecular clock framework is a (working) hypothesis, as Parenti & Ebach (2013) correctly point out – the relative importance of each of these events can be segregated only if the timing of lineage evolution through space is understood, thus providing more predictive power than lumping all taxa – with potentially very different diversification histories – into a single analysis of area relationships. We therefore strongly emphasize that temporal information (estimates) of both geological events and divergence times are essential to test biogeographical patterns and their causes.

Parenti & Ebach (2013) believe it possible for organisms to colonize a volcanic island in situ and cite the example of Hawaiian gobies: ‘We hypothesize that Hawaii’s endemic freshwater fish lineages ranged throughout the region where the Hawaiian Islands formed and, that the lineages of part of Hawaii’s terrestrial biota ranged throughout that region, not necessarily as marine organisms, but on once emergent lands’ (p. 814). However, while this hypothesis is theoretically falsifiable, practically, it is not (see falsifiability and pseudo-science discussion in e.g. Popper, 1994), as we are unable to sample these unidentified hypothetical islands (geologically) or their ancestral biota (biologically) to test these ideas. Further, to the best of our knowledge, there have been extended periods (millions of years) of time when
the Hawaiian Islands were completely submerged below sea level (Clague, 1996), although some panbiogeographers dismiss this evidence entirely (e.g. Heads, 2011). A counter-argument may be made that amphidromous gobies were able to disperse great distances across the sea (de Queiroz, 2005), a possibility not articulated by Parenti & Ebach (2013). This idea has been explicitly advanced by other ichthyologists (McDowall, 2007; Lindstrom et al., 2012).

The assumption of a vicariant distribution due to formation of a volcanic chain cannot be applied to the biota of Sulawesi based on our current state of geological knowledge. Two major vicariance events were revealed by the geological history of the composite island of Sulawesi: first, the opening of the Makassar Strait separating West Sulawesi from Borneo (Sunda Shelf); and secondly, the extension and westward movement of the Sula Spur from its former adjacent New Guinean/Australian landmass (Sahul Shelf; see discussion in Stelbrink et al., 2012). These represent ideal geological constraints with which to test whether vicariant scenarios can be excluded for terrestrial and freshwater taxa, through analysis comparing geological timing with estimated divergence times (see e.g. Figure 1 in Crisp et al., 2011). Parenti & Ebach (2013) note that ‘rejecting particular vicariance events to explain a distribution does not demonstrate dispersal as it ignores other vicariance events that may be invoked as an explanation.’ There might have been other vicariance events, but at the present time we cannot provide alternative explanations for Sulawesi’s fauna, and neither do Parenti & Ebach (2013).

Parenti & Ebach (2013) consider Sulawesi to form part of a biogeographical subregion they call ‘Pandora’ (Parenti & Ebach, 2010), spanning an area including parts of Australia, Melanesia, Madagascar, Africa, Samoa and the Hawaiian Islands. Their evidence to erect this new subregion is based on uninterrupted distributions ‘across a large portion of the ancient southern continent, Gondwana’ (Parenti & Ebach, 2010, p. 311). The map they provide clearly shows these areas were highly disjunct even 30 Ma; however, these continents are known to have split apart far earlier (115 Ma according to Chakrabarty et al., 2012; one of the studies cited in Parenti & Ebach, 2013). Parenti & Ebach (2013) select two studies as tests for their newly proposed subregion: Austin (2000) and Chakrabarty et al. (2012). These are both molecular phylogenetic studies that use molecular clocks to seek a simple older than/younger than result for their conclusions, a practice criticized by Parenti & Ebach (2013). We re-examined these studies in detail. Austin (2000) found an ancient sister relationship between Malagasy and Melanesian boine snakes. However, Austin’s (2000) study was made redundant by the more comprehensive study of Noonan & Chippindale (2006), which places the Malagasy boines with African taxa, rejecting the clade’s former relationship with Melanesian taxa and overturning the sole basis for including these snakes as evidence for Pandora (Appendix S1). We re-analysed Chakrabarty et al.’s (2012) data set on troglobite fishes but excluded one of their four calibrations. This fossil was placed in the wrong part of the tree, was incompatible with the other calibration points, and significantly influenced the overall result (see Appendix S2). When these analyses were re-run, otherwise identically, the age estimates are significantly younger (see Figs S1 & S2 in Appendix S2), suggesting neither a Pandoran nor a Gondwanan explanation. Neither molecular study (Austin, 2000; Chakrabarty et al., 2012) provides support for this putative biogeographical subregion, casting doubt on Pandora’s validity. We hope that the other studies used as raw data to derive Parenti & Ebach’s (2010) Pandora areagrams are more robust.

In conclusion, we echo a recent call (Waters et al., 2013) to move beyond the ‘ancient vicariance’ argument to explain each and every biogeographical scenario. While area cladograms may provide useful (although potentially flawed – see Appendix S1) hypotheses for biogeography, as theory and methods have evolved, approaches that integrate geography and biology within a temporal framework (e.g. molecular dating) provide far more power for inferring biogeographical patterns and processes.

ACKNOWLEDGEMENTS

We thank Alan de Queiroz and two anonymous referees for their useful and constructive comments on the manuscript. D.J.L. is supported by grant DEB-1120380 from the US National Science Foundation.

REFERENCES


The explanatory power of biogeographical patterns: a reply to de Bruyn et al.

ABSTRACT

Confusion between evidence and hypothesis in biogeographical studies was the focus of our recent Guest Editorial (Parenti & Ebach, 2013, Journal of Biogeography, 40, 813–820). That editorial was critiqued by de Bruyn et al. (2013, Journal of Biogeography, doi: 10.1111/jbi.12166) to whom we reply briefly here. Despite our shared goals – to understand what lives where and why – we argue from different philosophical premises. Although we may have little common ground, such debate encourages the good health of the field of biogeography.

Keywords Biogeography, cladistics, dispersal, diversification, general area cladograms, historical biogeography, molecular clock, Pandora, philosophy of science, vicariance.

Modern historical biogeographical studies fall into one of two kinds: those that generate explanations and those that discover patterns (Ebach & Humphries, 2002). We argue that discovery of patterns, such as that demonstrated through the shared components of areagrams, is primary in historical biogeography and that the generation of explanations logically follows the discovery of the patterns. Our approach to biogeography is presented in our recent book (Parenti & Ebach, 2009), to which we refer the interested reader. We expanded one topic in that book on the relationship between evidence and hypothesis in biogeography in a Guest Editorial (Parenti & Ebach, 2013), to which de Bruyn et al. (2013) have responded. We welcome this response as it encourages discussion and debate of some of the principles of historical biogeography and further exposes the assumptions made when proposing biogeographical hypotheses. We choose part of one sentence from the response of de Bruyn et al. (2013) to begin our comments. De Bruyn et al. (2013) “…believe that palaeo-geographical reconstructions by geological experts…for the IAA [Indo-Australian Archipelago] and Sulawesi are the most suitable framework within which to test biogeographical hypotheses, as opposed to the use of hypothetical…relationships of taxa (cladograms), by biologists, to reconstruct palaeogeography’. We disagree with this statement for several reasons.

First, biological and geological patterns are independent. They provide reciprocal illumination (Rosen, 1978); they shed light on each other, but do not test, and therefore cannot reject, each other. The information contained in biogeographical distributions should not be discounted: it tipped the scale in favour of continental drift away from the long-held notion of a stable Earth (e.g. McCarthy, 2009), not to mention that it played a pivotal role in the formulation of a theory of biological evolution (e.g. Wallace, 1863).

Second, biogeographical regions are defined by the taxa that live in them (as opposed to biomes or ecoregions which are defined by climatic and ecological factors). We recognize two regions, Indo-Malayan (Wallace, 1863) and Pandora (Parenti & Ebach, 2010), which together range from east Africa/Madagascar to the central Pacific. They were allopatric in the Oligocene, but today overlap, in part, notably in the centre of the Indo-Australian Archipelago. We drew them on a map of the Oligocene to demonstrate that they were separate regions at that time and also to illustrate that they include both marine and terrestrial areas. Pandora covers a portion of what was formerly Gondwana, as well as the seas between those areas. On
Appendix S1 Discussion of Austin (2000).

Austin (2000) found a deep sister relationship between boine snakes from Melanesia and Madagascar, and while he does indeed make the statement quoted in Parenti & Ebach (2013) that this is not a ‘recent dispersal event’, he also points out that differentiating between ancient dispersal and ancient vicariance is not possible here (a point also made by Parenti & Ebach, 2013, ‘…they assume that vicariance is old and dispersal is new’ p. 816). Austin (2000) concludes with the statement ‘…whereas the African–Papuan–American boinae radiation/distribution as having its origins in the Gondwanan supercontinent, neither the fossil record, nor robust estimates of phylogeny calibrated by fossil dates, support these hypotheses’ p. 348. These caveats are not mentioned by Parenti & Ebach (2013). Moreover, the ‘ancient’ sister relationship between Madagascan and Melanesian boines is based on analysis of a short 270 base pair (bp) mitochondrial fragment from a small number of species, which Austin (2000) acknowledges is weak evidence. A more recent analysis (Noonan & Chippindale, 2006), based on 4307 bp from five nuclear and one mitochondrial gene of all boine genera and related groups, contradicts Austin’s (2000) result by demonstrating that the Malagasy genus Sanzinia is not sister to the Melanesian genus Candoia, as inferred by Austin (2000), but instead belongs to a well supported clade with other Malagasy and African genera. This Africa/Madagascar clade is no more closely related to the Melanesian genus than it is to genera from South America, the Middle East, India and North America. The strongly supported hypotheses of area relationships from Noonan & Chippindale (2006) imply that the biogeographical patterns of boid snakes are not consistent with the existence of ‘Pandora’.
Appendix S2 Discussion and re-analyses of Chakrabarty et al. (2012).

A more recent study (Chakrabarty et al., 2012) is also proffered by Parenti & Ebach (2013) as support for Pandora, and as an ‘extraordinary case of Gondwanan vicariance’ (Chakrabarty et al., 2012; quoted by Parenti & Ebach, 2013, p. 815). Unlike Austin’s (2000) study, in which the Madagascar–Melanesia relationship was weakly supported, the cavefish from Australia and Madagascar form a stronger clade, with 73% bootstrap support and ‘> 100% posterior probability’. Although posterior probabilities are known to grossly overestimate confidence when the chosen substitution model is incorrect (Suzuki et al., 2002), posterior probability > 100% is an analytical impossibility. Chakrabarty et al. (2012) conclude that this sister-relationship is consistent with Gondwanan vicariance. This is based not on the topology (although that is a vital starting point) but on the age of their ancestral node, which is calculated using fairly standard relaxed molecular clock methods and fossil calibrations. The node linking Australia (Mileryinga) and Madagascar (Typhleotris) is dated by Chakrabarty et al. (2012) as 77 million years old (Ma; 116–44 Ma 95% higher posterior densities). Interestingly, the authors also conclude that the last terrestrial link between Indo-Madagascar and Antarctica–Australia was 115 Ma, and thus a vicariant explanation would require a complex series of geological and biological events, and this geological date is actually more in line with their own reported age for the entire worldwide eleotrid/Milyeringa/Typhleotris clade (109 Ma: 150–74 Ma). We note that this logic of using an inferred divergence date from a molecular phylogeny to infer either vicariance or dispersal by comparison with a geological divergence is critiqued heavily by Parenti & Ebach (2013), yet this study (Chakrabarty et al., 2012) was cited by these authors as support for ‘Pandora’.

The main problem with Chakrabarty et al. (2012) is their usage of molecular clocks to substantiate vicariance rather than dispersal. Heads (2005) warns against inappropriate usage, as do Parenti & Ebach (2013). Chakrabarty et al. (2012) use relatively short markers (1793 base pairs) of a single, variable locus (mitochondrial protein coding fragments) to date this potentially ancient node. Multiple, independent nuclear loci would have provided a more accurate estimate of divergence on this time-scale. The argument put forward in Chakrabarty et al. (2012) that only mitochondrial data were employed because existing publically available data are all mitochondrial is not a strong argument. Nuclear sequences would only have been required from a small subset (the two target taxa and a few outgroups) once the general structure had been established with the larger mitochondrial data set.

We re-analysed Chakrabarty et al.’s (2012) data set to scrutinize the conclusions they made. We found that one of their four calibration priors (C2: Beryciformes, 150–94 Ma) was incompatible with the others and largely drove the overall result. Graphing the posterior distribution for the age of C2 (Fig. S1) revealed that it was tightly pressed to the 94 Ma minimum bound, and strongly pushed the majority of the posterior distribution of the root age (C1) to be older than its respective soft maximum bound. This pattern is typical when one calibration is incorrectly assigned (Sanders & Lee, 2007) or when the model of rate evolution across the tree is poorly specified (Ho & Phillips, 2009).

In the present case, the middle Cretaceous fossil, Hoplopteryx is assigned as a stem-trachichthyid. More likely, the affinities of Hoplopteryx are far deeper, invalidating the minimum bound that Chakrabarty et al. (2012) used for the divergence of two Beryciformes, Hoplostethus (Trachichthyidae) and Anoplogaster (Anoplogasteridae). It has long been thought that trachichthyids are plesiomorphic among Beryciformes (e.g. Patterson, 1964) and thus the family has potential ‘waste basket status’ for falsely including plesiomorphic fossil taxa, such as Hoplopteryx. Moreover, the only matrix-based phylogenetic analysis we are aware of that infers the placement of Hoplopteryx among Beryciformes (Moore, 1993) finds that this fossil taxon falls outside of the clade that includes Trachichthyidae and Anoplogasteridae. To examine statistical support for Moore’s (1993) result we ran a bootstrap analysis on his data set. Random-addition (10 replicate) maximum parsimony heuristic
searches were run in PAUP* 4.0b10 (Swofford, 2002) for 500 bootstrap pseudoreplicates. Trachichthyidae and Anoplogasteridae formed a clade with Diretmidae, Monocentridae and Anomalopidae that gained 97% support, and Trachichthyidae and Anoplogasteridae grouped together (with or without these additional families) to the exclusion of Hoplopteryx with 100% bootstrap support, thus showing it had been incorrectly placed in the calibration scheme adopted by Chakrabarty et al. (2012).

When calibration C2 was removed from the Chakrabarty et al. (2012) data set and the analyses rerun, otherwise identically, the age estimate for the root node changes to fall within the fossil bounds and, significantly, the date of the key (Madagascar/Australia) node dropped from 77 Ma to 40 Ma (Fig. S2). Similarly, another recent paper that proclaimed ancient vicariance based on molecular clock analyses (Bauzà–Ribot et al., 2012), was shown to have used an inappropriate calibration scheme and when re-analysed (Phillips et al., in press) gave much younger dates that do not fit the presumed tectonic explanation.

Heads (2012, p. 1754) argues that assigning prior probability curves around fossil calibrations can seriously bias the age estimate and should be made with care, particularly if the researchers have ‘strong prior beliefs’. However, even had fossils and calibration curves been appropriately assigned, we a priori have reason to expect that the beryciform calibration could lead to spurious date estimates. These are among the longest-lived fish (de Magalhaes & Costa, 2009) and correspondingly have very slow substitution rates (Miya et al., 2003) relative to almost all other acanthomorph fish (including members of Gobiioidei, which are the focus of the study). Longer generation times are correlated with low substitution rates, leading to short branches in a phylogeny (Welch et al., 2008). Pinpointing where rate-slowdowns associated with longevity-associated substitution rates occur along stem lineages or within clades (such as Beryciformes) presents an identifiability problem, even for relaxed-clock models (Welch et al., 2008). Substantial rate-slowdowns typically result in stem rates of long-lived clades being underestimated, with clades outside the calibrated node being pulled older as a corresponding artefact (Phillips, in press).

New data on a recently discovered species of cavefish from the same Australian genus, Milyeringa (Larson et al., 2013), casts further doubt on the supposed Gondwanan affinities of these Australian and Malagasy cavefish. Larson et al. (2013) found that genetic divergence between the two Australian congeneric species is roughly the same as between either Australian species and the Malagasy species, implying a nearly equally ancient divergence within Australia itself. This implies that the neat and tidy ancient Australia/Madagascar divergence that Chakrabarty et al. (2012) prefer is unlikely. The magnitude of the divergences among species in this taxon is not supportive of either the influence of Gondwana or Pandora. Hence, neither of the two studies (Austin, 2000; Chakrabarty et al., 2012) that Parenti & Ebach (2013) select as ‘tests’ for Pandora offers support for the existence of such a biogeographical subregion.

Detailed methods of re-analyses
The data set from Chakrabarty et al. (2012) was downloaded from Dryad at http://datadryad.org/resource/doi:10.5061/dryad.6305t. In accordance with Chakrabarty et al. (2012) our re-analysis was run in BEAST 1.6.1 (Drummond & Rambaut, 2007), with the separate substitution models they employed for each of the four mitochondrial genes (COI; HKY+G, NADH1; GTR+G, NADH2 and Cytb; GTR+I+G) (there are no NADH2 sequences for the two target genera, Milyeringa and Typhleotris). An uncorrelated relaxed clock model was used with rates among branches distributed according to a lognormal distribution. Two independent runs totalling 50,000,000 Markov chain Monte Carlo (MCMC) generations ensured estimated sample size values > 100 (as estimated in TRACER 1.5) (Drummond & Rambaut, 2007) for all node height, prior, posterior, −lnL, tree, and substitution parameters. Chains were sampled every 5000th generation after burn-ins of 5,000,000 generations.
Initially the molecular clock analyses were run including all four of Chakrabarty et al.’s (2012) calibration priors and the resulting BEAST posterior distributions for the age of Acanthomorpha (the C1 calibration, which is the root of the tree) and Beryciformes (the C2 calibration) graphed (see Fig. S1). Chakrabarty et al. (2012) calibrated both Beryciformes and the root as lognormal prior distributions with hard minimum bounds at 94 Ma and soft maximum bounds at 150 Ma. The interaction of the data and these calibrations results in the Beryciformes posterior being pushed up against the 94 Ma hard minimum bound, and the majority of the root age distribution being pushed older than the soft maximum. Given that lognormal priors heavily penalise dates very close to their minimum, a uniform prior would press the Beryciformes posterior even flatter against the hard minimum.

Analyses were then re-run including three of the calibration points (C1, C3, C4), but excluding C2 (Beryciformes). The new estimate from the resulting BEAST relaxed-clock timetree (Fig. S2) for the age of the key Milyeringa (Australia) / Typhleotris (Madagascar) divergence is 40.0 million years ago (53.4–26.8 Ma 95% higher posterior densities), much less than the 77 (116–44) Ma result when C2 is included (as per Chakrabarty et al., 2012), and even more removed from the 115 Ma geological date of the last terrestrial link between Indo-Madagascar and Antarctica–Australia referenced by Chakrabarty et al. (2012).

Figure S1 BEAST posterior distributions for the age of Acanthomorpha (the C1 calibration, which is the root of the tree) and Beryciformes (the C2 calibration), with all four of Chakrabarty et al.’s (2012) calibration priors included in the analysis.
**Figure S2** BEAST relaxed-clock timetree, with the C2 (Beryciformes) calibration excluded, but otherwise analysed as per Chakrabarty *et al.* (2012), including calibrations C1, C3 and C4. Blue bars are 95% highest posterior distributions. The mean estimate for the *Milyeringa* (Australia) / *Typhleotris* (Madagascar) divergence is 40.0 million years ago.
REFERENCES


