The linking of plate tectonics and evolutionary divergences (Reply to Bauzà-Ribot *et al*.)

Matthew J. Phillips[1,*], Timothy J. Page[2,*], Mark de Bruyn[3], Joel A. Huey[2], William F. Humphreys[4], Jane M. Hughes[2], Scott R. Santos[5], Daniel J. Schmidt[2], and Jonathan M. Waters[6]

It is exciting to be living at a time when some of the big questions in biology can be investigated using great advances in genetics and computing [1]. Bauzà-Ribot *et al*. [2] take on one of the fundamental drivers of biodiversity on our planet, the effect of continental drift in the formation of the world's biota [3, 4], employing a combination of whole mitochondrial genomes from next-generation sequencing and modern Bayesian relaxed molecular clock analysis. Bauzà-Ribot *et al*. [2] conclude that vicariance via plate tectonics best explains the genetic divergence between subterranean metacrangonyctid amphipods currently found on islands separated by the Atlantic Ocean. This finding is a big deal in biogeography, and science generally [3], since many other presumed biotic tectonic divergences have been explained as probably due to more recent transoceanic dispersal events [4]. However, molecular clocks can be problematic at the best of times [5, 6] and we have identified three issues with the analyses of Bauzà-Ribot *et al*. [2], that cast serious doubts on their results and subsequent conclusions. When we reanalyzed their mitochondrial data and attempted to account for problems with calibration [5, 6], modeling rates across branches [5, 7] and substitution saturation [5], we inferred a much younger date for their key node. This implies either a later trans-Atlantic dispersal of these crustaceans, or more likely a series of later invasions of freshwaters from a common marine ancestor, but either way probably not the result of ancient tectonic plate movements, thus highlighting pertinent issues for all practitioners of evolutionary dating studies.

Bauzà-Ribot *et al*. [2] use up-to-date molecular dating methods, with calibrations from two paleogeographic events derived from presumed vicariant splits (in the Moroccan High-Atlas 37.2–25.0 mya [million years ago] and the Mediterranean 16–5.5 mya) to make an association between evolutionary divergence and real time. Because rates of molecular evolution can vary greatly between lineages and over time, multiple calibrations from fossils and tectonic events in different

parts of the tree may reduce this error [5], however it is no panacea [6]. We have several concerns with their dating inference. Firstly, they estimate deep node ages from far younger calibrated nodes, without also placing bounds deeper in the tree. This kind of extrapolation can act as a rate-error multiplier for deep nodes and led Thorne and Kishino [8] to require a root prior, so dates are instead interpolated between calibrations. Bauzà-Ribot *et al*.'s [2] High-Atlas calibration largely drives the divergence estimates, which in their various analyses closely converged with or without the Mediterranean calibration. However, the lack of non-metacrangonyctid outgroups in their molecular clock analyses may preclude accurate rate estimation across the root (between the High-Atlas calibration and the trans-Atlantic clade of interest). A similar problem caused a two-fold age overestimation in monotremes [9]. To address this concern, we have added various outgroups (amphipods and deeper-diverging malacostracans) which allowed us to place a fossil calibration prior on the root of the tree (Supplemental Information) while also retaining the younger biogeographic calibrations from Bauzà-Ribot *et al*. [2].

Secondly, we note that the mitochondrial 3rd codon positions are highly saturated, averaging >8 superimposed substitutions per site along some ingroup branches and far more among outgroup branches. Bauzà-Ribot *et al*. [2] test only for saturation extinguishing phylogenetic signal and not its impact on branch length estimation, which is directly relevant to molecular dating. We show that available substitution models under-correct for 3rd codon position saturation in Bauzà-Ribot *et al*.'s [2] original dataset by ~15% (Supplemental Information), so 3rd codon positions were excluded in our analyses. A third concern, that is exacerbated by the need to include outgroups, is that the distribution of rates across the tree is not lognormally distributed (Supplemental Information), as assumed in Bauzà-Ribot *et al*.'s [2] analyses by their choice of model. Rates among their metacrangonyctids are distributed at least bimodally, with outgroups adding an additional rate region (Fig. 1C). Instead of the lognormal distribution model, we use the more flexible random local clocks model [10], but otherwise maintain the same substitution models and tree priors to reanalyze the data.

Our result for the divergence linking both sides of the Atlantic was 39.9 mya (47.5–34.3 mya 95% highest posterior distribution, HPD) (Fig. 1A). The posterior distributions for the two biogeographic calibrations strongly conflict with the fossil calibration, and are tightly pressed to

their minima when enforced and fall much younger when free (Fig. 1B), implying these events may not be associated with the chosen divergences. Upon excluding the two biogeographic calibrations, the trans-Atlantic divergence becomes even more recent at 20.3 mya (24.9–15.8 mya 95% HPD) (Fig. S1A).

Bauzà-Ribot *et al*. [2] lay out a clear biogeographic hypothesis that the widening and deepening of the Tethys Sea around 110–95 mya explains the trans-Atlantic divergence, and adopt this vicariant conclusion based on their 79 mya (108–60 mya 95% HPD) dating of the trans-Atlantic divergence, which, while not a perfect match, at least partially overlaps the relevant range. Bauzà-Ribot *et al*. [2] suggest that younger inferred divergence times would lend credence to a dispersal scenario from the old world to the new, which fits with our results better. This should come as no surprise since Bauzà-Ribot *et al*. [2] say that the ancestral population of these freshwater taxa was a wide-ranging marine species ("thalassoid"), and therefore must have independently colonized caves in each location later (common in subterranean fauna [3]), presumably at different times in different lineages. Some might suggest that the island home of every member of the relevant trans-Atlantic clade (Hispaniola, Fuerteventura, Mallorca, Menorca, Elba) would actually imply that this lineage was an active and successful disperser at times, instead of being only a passive passenger on tectonic plates.

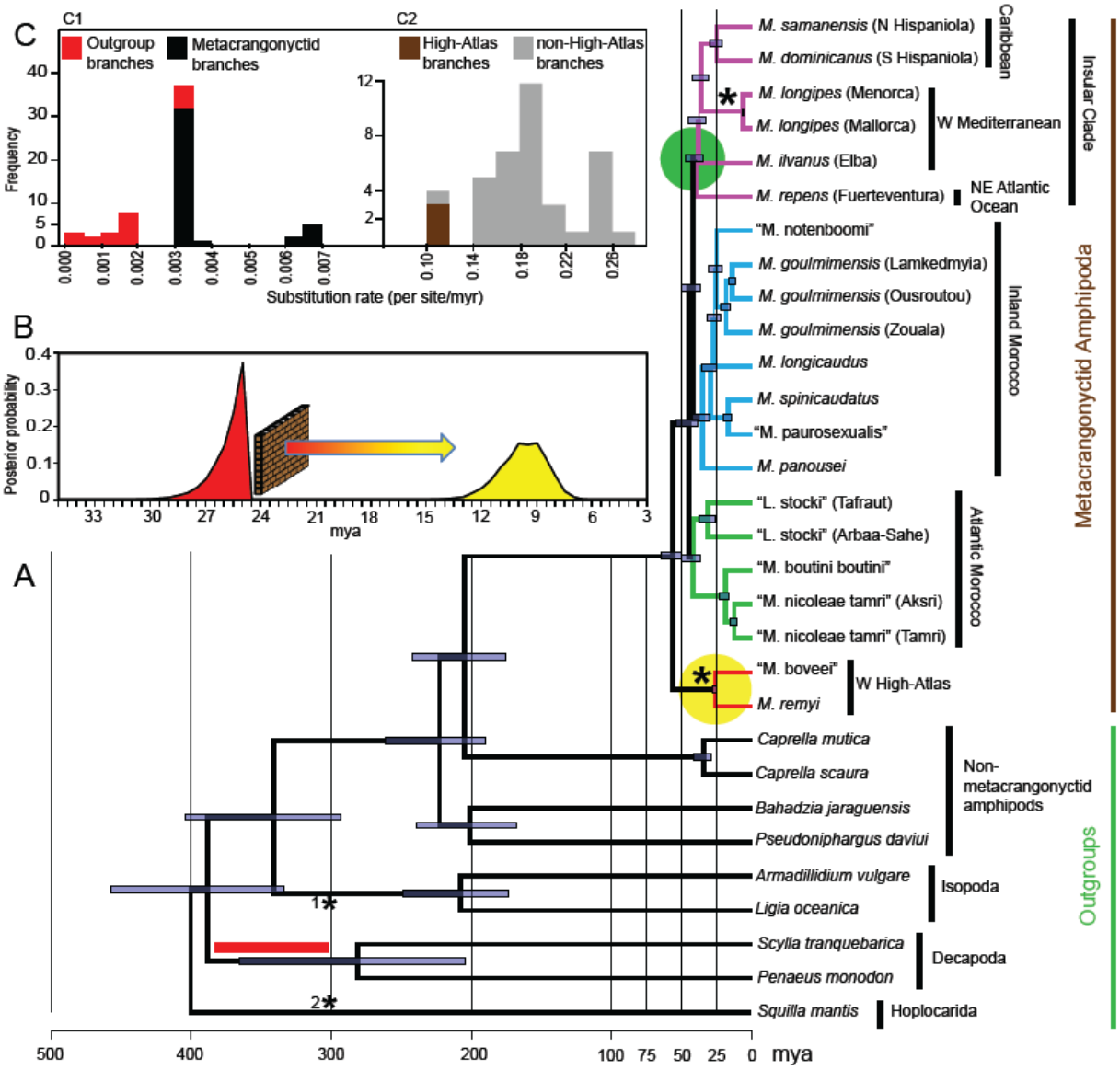
Rather than providing a definitive "answer", our differing results and conclusions highlight the difficult nature of some of biology's big questions. Given the rapid substitution rates (in both our and Bauzà-Ribot *et al*. [2] analyses) and the great age of the question being considered, slower evolving nuclear sequences [1] may be better suited to this particular biogeographic question. The higher substitution rates of mitochondrial genes can give extra resolution [2], but can also amplify model misspecification [7]. The alluring nature of next-gen sequencing and relaxed molecular clocks can be a cruel mistress. To paraphrase Voltaire (or perhaps Spider-Man's Uncle Ben): "with great power comes great responsibility". This big question will need closer calibrations, and perhaps the addition of independent nuclear loci, before we can deal with it appropriately.

[1]School of Earth, Environmental and Biological Sciences, Queensland University of Technology, Brisbane, QLD, 4001, Australia. [2]Australian Rivers Institute, Griffith University, Nathan, QLD, 4111, Australia. [3]Molecular Ecology & Fisheries Genetics Laboratory, School of Biological Sciences, Bangor University, Bangor, LL57 2UW, UK. [4]Western Australian Museum, Locked Bag 49, Welshpool DC, WA, 6986, Australia. [5]Department of Biological Sciences, Auburn University, Auburn, AL 36849, USA. [6]Department of Zoology, University of Otago, PO Box 56,Dunedin, 9054, New Zealand.
* These authors contributed equally. E-mails: m9.phillips@qut.edu.au, t.page@griffith.edu.au

1. Emerson, B.C., and Hewitt, G.M. (2005). Phylogeography. Curr. Biol. *15*, R367-R371.
2. Bauzà -Ribot, Maria M., Juan, C., Nardi, F., Oromí, P., Pons, J., and Jaume, D. (2012). Mitogenomic Phylogenetic Analysis Supports Continental-Scale Vicariance in Subterranean Thalassoid Crustaceans. Curr. Biol. *22*, 2069-2074.
3. Maderspacher, F. (2012). Evolution: Drift Will Tear Us Apart. Curr. Biol. *22*, R909-R912.
4. Crisp, M.D., Trewick, S.A., and Cook, L.G. (2011). Hypothesis testing in biogeography. Trends Ecol. Evol. *26*, 66-72.
5. Pulquério, M.J.F., and Nichols, R.A. (2007). Dates from the molecular clocks: how wrong can we be? *22*, 180-184.
6. Heads, M. (2012). Bayesian transmogrification of clade divergence dates: a critique. J. Biogeogr. *39*, 1749-1756.
7. Phillips, M.J. (2009). Branch-length estimation bias misleads molecular dating for a vertebrate mitochondrial phylogeny. *441*, 132-140.
8. Thorne, J.L., and Kishino, H. (2002). Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. *51*, 689-702.
9. Phillips, M.J., Bennett, T., and Lee, M.S. (2009). Molecules, morphology, and ecology indicate a recent, amphibious ancestry for echidnas. *106*, 17089-17094.
10. Drummond, A.J., and Suchard, M.A. (2010). Bayesian random local clocks, or one rate to rule them all. BMC Biol. *8*, 114.

Figure Caption

Figure 1. (A) BEAST random local clock timetree employing paleogeographic calibration bounds (asterisks represent calibrations) and fossil calibration at the root, with the 95% soft bounds prior indicated by a red bar and the associated minimum bound fossils indicated (asterisk 1 for the isopod, *Hesslerella* and asterisk 2 for the hoplocarid, *Gorgonophontes*). Blue bars are 95% highest posterior distributions (HPDs). Green circle highlights divergence of Trans-Atlantic clades and dates to 39.9 mya (47.5–34.3 mya 95% HPD). Yellow circle highlights the High-Atlas calibration node. (B) Posterior age distributions for the highlighted High-Atlas calibration node when both fossil and paleogeographic calibrations are used (red on left) with a “hard” lower boundary (i.e. brick wall at 25 mya); and when only fossil calibration is used (in yellow on right). (C) Posterior distributions of substitution rates along branches, inferred under random local clocks for (C1) the full taxon set (3rd codon positions excluded) and for (C2) metacrangonyctids alone (all codon positions). Root calibrations only were employed for inferring these distributions, thus avoiding rate distortion owing to conflict between calibrations (see Supplemental Information).


C
C1
C2
Outgroup branches
Metacrangonyctid branches
High-Atlas branches
non-High-Atlas branches
Frequency
Substitution rate (per site/myr)
B
Posterior probability
mya
A
M. samanensis (N Hispaniola)
M. dominicanus (S Hispaniola)
Caribbean
M. longipes (Menorca)
M. longipes (Mallorca)
W Mediterranean
M. ilvanus (Elba)
M. repens (Fuerteventura)
NE Atlantic Ocean
Insular Clade
"M. notenboomi"
M. goulmimensis (Lamkedmyia)
M. goulmimensis (Ousroutou)
M. goulmimensis (Zouala)
M. longicaudus
M. spinicaudatus
"M. paurosexualis"
M. panousei
Inland Morocco
"L. stocki" (Tafraut)
"L. stocki" (Arbaa-Sahe)
"M. boutini boutini"
"M. nicoleae tamri" (Aksri)
"M. nicoleae tamri" (Tamri)
Atlantic Morocco
Metacrangonyctid Amphipoda
"M. boveei"
M. remyi
W High-Atlas
Caprella mutica
Caprella scaura
Bahadzia jaraguensis
Pseudoniphargus daviui
Non-metacrangonyctid amphipods
Armadillidium vulgare
Ligia oceanica
Isopoda
Scylla tranquebarica
Penaeus monodon
Decapoda
Squilla mantis
Hoplocarida
Outgroups
1*
2*
mya