

Round 2

**## First of all, we would like to warmly thank the editor for all the work he has realized on our manuscript, including his corrections/suggestions on the text. This is exactly what is expected from an editor and his work has greatly helped us to improve our study. We realize our study is not an easy one to evaluate, as we present a hypothesis consisting of 4 main postulates (in comparison with most other published hypotheses having a single postulate), tested with three different groups using many different analyses and datasets, which come (sometimes) with contradictory results (e.g. those based on extant data vs. fossil data). We have modified the text following all your suggestions that has resulted in a simplified and clearer message. In particular, we have re-written all the discussion and clarified the introduction, which brings a better placement of the current study on the existing literature. We hope this new version of the manuscript is clearer and more concise.**

I apologize for the slowness of this review; as I intimated in my first report, I find this large and ambitious study difficult to evaluate. The "AGE hypothesis" that this paper offers seems to be a formalization of ideas that have been around for a long time, making it important, but its placement in the literature problematic.

**## We agree with the editor, the placement of our study in the literature was unclear in previous versions. Any component of the AGE model (e.g. Holarctic extinction, into the equator dispersal, etc...) is novel on its own. But the novelty is that the AGE model combines and reconciles previous competing hypotheses on the origin of the LDG by placing them in a temporal scenario (the time-variable framework). For instance, there is controversial support around the tropics being 'cradle' or 'museum of diversity', and dispersal prevailing 'out of' or 'into the tropics'. Alternatively the AGE hypothesis invokes the 'museum of diversity' regarding the equatorial tropics as refuge during coldhouse transitions, but also the 'cradle of diversity' during greenhouse periods. Similarly, this hypothesis invokes dispersal 'out of the equator' during greenhouse transitions and dispersals 'into the equator' during coldhouse transitions. The AGE hypothesis also assumes tropical niche conservatism due to physiological limits, as we postulate that most of the tropical-adapted taxa at high latitudes were unable to adapt and thus went extinct or suffered restrictions of their distributions and/or dispersed southward when the tropical biome retreated toward the equator.**

**The detection of diversity loss in the Holarctic is also novel in our study. Many studies have proposed hypotheses related to extinction before. For example, Hawkins et al. (2006) suggested that the avian LDG resulted from the differential extirpation of older warm-adapted clades from the temperate regions newly formed in the Neogene. Pyron (2014) suggested that higher temperate extinction represents a dominant force for the origin and maintenance of latitudinal gradients. More recently, Pulido-Santacruz and Weir (2016) proposed (not found or tested this properly) the terrestrial latitudinal diversity gradient is largely the effect of a post-Eocene increase in extinction rates at high latitudes. However, the problem is that, in our opinion, some of the literature on the LDG is equivocal, mixing diversification terms and concepts, and in many cases proposing hypotheses that go beyond their results. For example, Hawkins et al. (2006) studied the distribution of branch**

lengths across regions and found that the tropics mostly harbor 'basal' (old) species while temperate regions mostly harbor 'derived' (recent) clades. They interpreted this pattern as an evidence of extinction. However, they did not infer diversification rates, and in addition, other processes such as recent dispersal, or high turnover could also explain this pattern. Similarly, despite his claims, the study of Pyron (2014) only detected high regional turnover (*i.e.* high extinction accompanied by high speciation; one fauna is replaced by another), but high turnover can account only for a slow accumulation of lineages, not for diversity declines. Diversity declines (a reduction in the net number of species) only occur when extinction exceeds speciation, resulting in negative net diversification rates ( $r = \lambda - \mu$ ;  $r < 0$ ). The same applies for Pulido-Santacruz and Weir (2016), and other studies on the LDG (e.g., Cardillo et al., 2005; Weir & Schluter, 2007; Leslie et al., 2012; Pyron & Wiens, 2013; Rolland et al., 2015): these studies only proved high regional turnover but not diversity loss. In this sense, we think that 'diversity loss' hypotheses need to be differentiated from 'high turnover' scenarios. The lack of evidence supporting diversity loss could be explained by the lack of phylogenetic methods allowing detecting negative diversification rates until the recent years (Morlon et al., 2011; Stadler, 2011). In addition, the power of time-constant BiSSE models for detecting negative diversification rates is very questionable: inferring negative diversification under a time-constant model (and thus for the entire history of lineages) conflicts with the fact that these groups are still extant. This suggests that the negative diversification rates detected for some clades in the study of Pulido-Santacruz and Weir (2016) using time-constant BiSSE models might be an artifact of the model.

Meanwhile, numerous fossil investigations have detected signatures of extinction and diversity loss in the Northern Hemisphere, including birds (Mayr et al., 2017), invertebrates (Wilf et al., 2005; Archibald et al., 2010, 2013), mammals (Blois & Hadly, 2009; Rose et al., 2011; Marcot et al., 2016) and plants (Frederiksen 1988; Peralta-Medina & Falcon-Lang, 2012; Smith et al., 2012; Xing et al., 2014). Unfortunately, fossil studies are generally restricted at the geographic and temporal scale, which make difficult the extrapolation of local inferences of extinction to the context of the LDG.

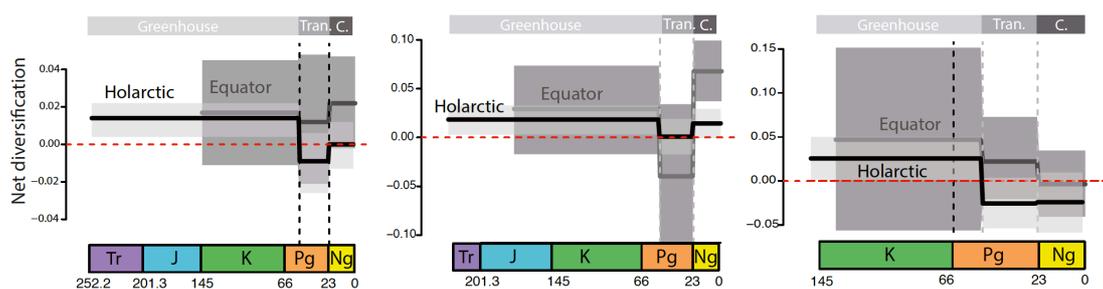
**In summary, high extinction does not necessarily mean diversity loss, if accompanied by high speciation. In this context, 'diversity loss' hypotheses have being repeatedly proposed, but to our knowledge never demonstrated. We have now clarified this and be more explicit in the introduction/discussion.**

There is no doubt that the combination of reconstructed phylogenetic and fossil tests is novel, and offers a deep time perspective on the Latitudinal Diversity Gradient (LDG).

The authors make several claims about the process leading to the present-day LDG using several different data streams for several different taxa, sometimes keeping the data streams separate, and sometimes combining them. The main red thread seems to be that reconstructed phylogenies can capture neither (i) high extinction (leading to "negative diversification") in Holarctic regions, nor (ii) "into the tropics" dispersal and range contractions, both occurring during the transition from Greenhouse to Coldhouse conditions between ~50mya to ~23 mya. Because of this, these two processes are potentially underappreciated drivers of the LDG.

The authors present PyRate analyses of fossil genera for the three groups, considering origination and extinction of lineages in Holarctic and Equatorial realms for Greenhouse, Transition, and Coldhouse periods (so, two rates for three times for three clades, or 18 separate rates, Figure 4. To summarize, turtles show the predicted decrease in origination and increase in extinctions during the Transition phase in the Holarctic leading to negative diversification there. Crocodiles show a decrease in origination rates in both realms in the Transition phase, but an increase in extinction in *equatorial* realm during the transition phase, and *no change* in the Holarctic; overall, there seems to be negative diversification in the Holarctic since the onset of the Transition phase. Squamates show net decreases in origination in both realms with the onset of the Transition phase, though more extreme in the *equatorial realm*, and *no change* in extinction rates in either realm during the Transition phase, and, finally, a decrease in extinction rates in the *Coldhouse* phase. These major results suggest that there are potentially strong global climate – origination and extinction links in all three groups, but only the Turtles show an overall pattern consistent with AGE. (I have italicized the contrary-to-expectation patterns in this paragraph.)

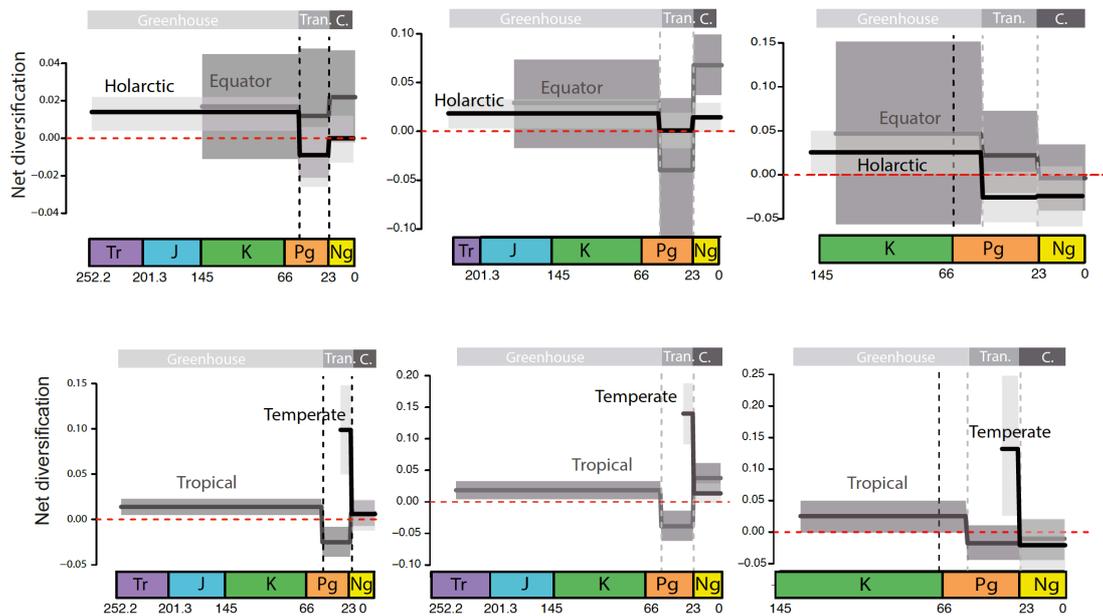
**## We would like to clarify that the results must be interpreted here based on the net diversification rates, instead of dissecting the pattern for speciation/extinction dynamics. A change in net diversification rate is the only mechanism (together with dispersal) that could promote loss or gain of diversity in a region (i.e. changes in the net number of species), while, by their own, speciation/extinction dynamics could vary without impacting net number of species, as explained above. When focused on the net diversification rate, we can see in the figure below that diversity losses are detected for both turtles and crocodiles during the Paleogene in the Holarctic (= AGE prediction 1). In the case of turtles (on the left) the diversity loss in the Holarctic ( $r < 0$ ) is the product of decreases in speciation and increases on extinction. For crocodiles (on the right) the decline is due to decreases in speciation and constant extinction. For squamates (in the middle) we only detected high Holarctic turnover ( $r = 0$ ).**



My reading of the methods suggests that Figure 5 is just a reworking of Figure 4 that combines place and time in the Greenhouse phase to produce 12 rather than 18 rates.

**## We would like to clarify that Figure 5 is not a reworking of Figure 4 but a new and independent analysis. There are four different datasets: the ‘tropical’, the ‘equatorial’, the ‘Holarctic’ and the ‘temperate’ taxa for each group. And data included in each dataset is different from the others. For example: the ‘tropical’ dataset includes all the fossils living under tropical conditions, that is to say, all the fossils appearing in the equator through time, plus all the fossils**

appearing in the Holarctic until the Oligocene (30 Ma), when paratropical conditions disappeared from the region and macroclimatic conditions shifted to temperate ones. The ‘equatorial’ dataset only includes the fossils appearing in the equator. Meanwhile, the ‘Holarctic’ dataset includes all fossils living in the Holarctic, under both tropical and temperate conditions. Doing so, the goal is to evaluate whether diversification dynamics changed between taxa living under tropical or temperate conditions, independently of whether they were distributed in the Holarctic or in the equator. Altogether, these different analyses allowed us to infer that, among the Holarctic extinctions detected in the Paleogene, the most impacted clades were the tropical-adapted ones and not the temperate clades, which actually had high diversification rates during this period. We hope this clarifies the purpose of Figures 4 and 5.



The authors then consider range contractions and dispersals. Here I think there is a very useful message – the authors offer great examples where fossil-based constraints on ancestral lineage area of occupancy leads to very different scenarios of dispersal. As the authors succinctly put it, there is no need to formally compare model fits of ancestral occupancy and movement, since the fossils are clear evidence that a particular lineage was in fact where the fossil was. This, and the comparison of BiSSE and PyRate estimates of diversification, highlight clearly how geographically biased extinction can severely bias any sort of reconstruction on a tree – we "knew" that nonhomogeneity of process can cause all sorts of problems, but these empirical comparisons that make use of fossil data hit home. These two sets of comparisons, if presented clearly, will be read with care and interest by comparative biologists.

With regard to the fossil-based reconstructions (Figure 7), all three groups show evidence for rates of dispersal into the equator to be generally increasing through time and greater than those out of the equator, though with no discernable effect of the various temperature phases. Range extinctions show idiosyncratic patterns, again with no clear effect of the Transition or the Cold phase.

**## This is a very good point that we did not consider before. Indeed, the ‘range contractions’ and ‘into the equator’ dispersals detected in our study increase**

from the Cretaceous until the present (Figure 4), and are not restricted to the time interval considered here to represent the transition phase to cold (between 51 and 23 Ma, after the early Eocene Climatic Optimum 'EECO'), as postulated in our AGE prediction 2. This finding could suggest that other processes different to climate change mediated the extinction and range contraction of Holarctic lineages in the Cretaceous. Alternatively, this result could suggest that a transition phase to cold started before the relatively short interval considered here. We consider this interval between 51 and 23 Ma, based on paleontological evidence showing that paratropical conditions and the associated warm-adapted taxa definitely disappeared from high latitudes between the mid-late Eocene and the Neogene (Wolfe 1975, Morley 2007). However, some studies considered the Eocene warming of the EECO only represented a transient temperature peak within an otherwise cooling trend that started in the Cretaceous (Linnert et al. 2014; Zachos et al., 2008). This trend was intensified by the Cretaceous-Paleogene (K-Pg) mass extinction, and the drop in temperatures caused by the impact-associated winter. Indeed, we can observe a negative relationship between temperature and the number of dispersals 'into the tropics' (Figure 4). In our study, lineage extinctions, range extinctions and southward dispersals increased as temperature decreased between the K-Pg and Neogene (Figures 4 and 7), suggesting an additive effect of K-Pg and Neogene cooling on the species depopulation in the Holarctic. We have included these points in the discussion section.

As you can see, my reading and re-reading of the paper does not have me seeing how the data are so consistent with AGE as the authors. Given PCI is a nascent and voluntary endeavor, it is likely best that this go to a journal that has expert reviewers in paleontology and comparative methods at hand to offer another opinion.

**## We hope the clarifications above have convinced you about the support of our results for the AGE model, at least for turtles and crocodiles. The general pattern that could be extracted from our study is that the AGE model was poorly supported for all groups by analyses based purely on data for extant species, including biogeographic and time-constant BiSSE and time-variable diversification models (although both BiSSE analyses support AGE P2). Conversely, this evolutionary scenario was supported for crocodiles and turtles using fossil and fossil-informed phylogenetic investigations (AGE postulates P1–P2).**

Support for the AGE model is mixed in squamates. On the one hand the detected Holarctic range contractions are in agreement with "higher Holarctic diversity loss" scenarios (AGE P2). On the other hand, the evidence for high Holarctic turnover does not support the AGE P1 and are more in line with "slower Holarctic diversity accumulation" hypotheses. Nonetheless, we think that a better description of the squamate fossil record could reveal more high-latitude extinctions and range contractions not detected in our study. Please notice that in our study we have analyzed 4,798 squamate fossil occurrences on the diversification analyses, of which 30 fossils were used as constraints in the biogeographic analyses, for a group with 10,000 species (i.e. 0.47 fossils per species, and 0.003 fossil biogeographic constraints per species). In comparison, we had 4,084 fossil occurrences, of which we used 23 biogeographic constraints for turtles, in a group composed by 300 species (i.e. 13.61 fossils per species, 0.08 fossil biogeographic constraints per species).

**This shows that proportionally, in our study we considered very few fossils of Squamata in comparison with the fossils considered in other groups. We think a better understanding of fossils may change our vision of the squamate LDG pattern through time.**

I have, however, made extensive comments and suggestions on the manuscript itself (as notes using Preview), which I hope will help the authors as they prepare the paper for submission to a major journal. I will send this file to PCI directly to pass on.

**## Thank you very much for the editing work on the manuscript. We have included all the suggested corrections in the revised version.**

Overall, the multiple comparisons among constant rate BiSSE, time-dependent BiSSE, Pyrate using both time and space and time+space = "environment" sampling, and constrained, semi-constrained and unconstrained DEC analyses for each of the three clades make for a study that risks getting lost in its many, contradictory tests. That would be a shame, as the general idea of being able to test deep time climate-driven negative diversification using some sort of combined test with trees and PyRate would be a major advance in our thinking about this "oldest pattern in ecology."

Sincerely, Arne Mooers

**## We have copied below one comment from the text that required a proper answer.**

This is one main issue I still have with this study - do we believe the BiSSE results or not - one cannot say things are sometimes "consistent" with BiSSE results if one is also claiming that BiSSE is problematic because there are no fossils

**## We would said that our fossil-based analyses suggest that BiSSE is providing biased estimates of diversification rates. However, these methods are still today a reference for the study of diversification, and up to our knowledge no study has yet evaluated their potential to infer negative diversification rate in a given time period. Therefore, we think they need to be presented but interpreted with utmost caution in the context of the LDG. In addition, we think it is important, if not crucial, to perform several independent analyses to address the same question given the many criticisms on the different diversification models (e.g. SSE models, BAMM). One needs to rely on more than one method to have solid results (see Legendre & Condamine 2018 – Syst. Biol. for an example on insect eusociality and diversification).**