

## LETTER

# Faster diversification on land than sea helps explain global biodiversity patterns among habitats and animal phyla

John J. Wiens\*

Department of Ecology and  
Evolutionary Biology, University of  
Arizona, Tucson, AZ 85721-0081,  
USA

\*Correspondence: E-mail: [wiensj@email.arizona.edu](mailto:wiensj@email.arizona.edu)

### Abstract

Terrestrial environments occupy ~ 30% of the Earth's surface yet contain ~ 80% of all species. The causes of this dramatic biodiversity gradient have remained relatively unstudied. Here, I test the fundamental prediction that predominantly non-marine clades have more rapid rates of diversification than marine clades, using a time-calibrated phylogeny of animal phyla. The results strongly support this hypothesis. This pattern helps explain the higher richness of terrestrial environments and the dramatic variation in species richness among animal phyla. The results show the importance of ecology in explaining large-scale patterns of clade richness and of diversification rates in explaining Earth's largest biodiversity patterns. The results also demonstrate remarkable niche conservatism in habitats, in some cases lasting > 800 million years. Finally, the results highlight the surprisingly high species richness of freshwater habitats, which are nearly equal to marine environments despite their much smaller area (~ 2% of Earth's surface vs. 70% for marine habitats).

### Keywords

Animal phyla, diversification, freshwater, marine, niche conservatism, phylogeny, species richness, terrestrial.

*Ecology Letters* (2015) **18**: 1234–1241

## INTRODUCTION

The disparity in species richness between marine and non-marine environments is arguably the most dramatic global biodiversity gradient, but one that remains poorly understood. Estimates suggest that ~ 75–85% of eukaryotic species occur in non-marine (terrestrial and freshwater) environments, although these habitats occupy only ~ 30% of Earth's surface (May 1994; Mora *et al.* 2011; Costello *et al.* 2012). Many authors have commented on this disparity, and have proposed several ecological factors that may be involved (May 1994; Benton 2001; Vermeij & Grosberg 2010; Grosberg *et al.* 2012). These include greater net primary productivity (or greater area of productive habitats) in terrestrial environments, different patterns of primary productivity and herbivory (typically large plants relative to small herbivores on land vs. mostly microscopic primary producers in oceans), more difficult movement in water than air, and greater habitat heterogeneity and geographic isolation on land (reviews in May 1994; Benton 2001; Vermeij & Grosberg 2010).

A more basic hypothesis than any specific ecological factor is that higher terrestrial richness is caused by higher diversification rates of predominantly terrestrial clades (e.g. Benton 2001; Vermeij & Grosberg 2010). Speciation, extinction and dispersal are the processes that directly change species numbers in a given habitat (e.g. Ricklefs 1987), and ecological factors must therefore act through these processes to drive richness patterns. Differences in species richness between habitats should therefore depend on (Wiens 2011): (1) how long each habitat is inhabited (more time for speciation in habitats colonised earlier), (2) rates of dispersal between habitats and (3) rates of net diversification in each habitat, where net diversification is the balance of speciation and extinction over time

(Magallón & Sanderson 2001). Given that animals likely originated in oceans long before colonising land (e.g. Jeram *et al.* 1990; Narbonne 2005; Kenrick *et al.* 2012), higher terrestrial (or non-marine) richness might be explained by faster net diversification rates in terrestrial clades rather than more time in terrestrial habitats. However, this hypothesis has not yet been tested, at least not at a broad phylogenetic scale. Furthermore, some authors have stated that net diversification rates are uninformative in explaining large-scale diversity patterns (Rabosky *et al.* 2012). Thus, it remains unclear whether higher diversification rates in non-marine clades actually explain higher terrestrial richness. Demonstrating such differences in rates is necessary before testing more specific hypotheses as to why they differ.

Here, I test whether occurrence in marine vs. non-marine environments drives higher diversification rates among major clades of animals. In doing so, I simultaneously address another basic question about global biodiversity: what explains the disparity in richness of animal phyla? Animal phyla vary from a single described species (Placozoa) to > 1 million (Arthropoda; Zhang 2013), but few convincing hypotheses have been proposed to explain these differences. For example, Orme *et al.* (2002) found that body size did not explain richness patterns among animal phyla. Given the richness among habitats described above, another hypothesis to consider is that ecology (i.e. marine vs. non-marine habitat) explains this disparity, but this has not been explicitly tested.

To address the hypothesis that non-marine clades have faster diversification rates, time-calibrated trees including most animal phyla were estimated. Next, the number of described species in each phylum in each habitat was summarised. Net diversification rates were then estimated for each phylum, given their age and richness. The hypothesis was then tested

that clades with higher proportions of non-marine species have higher diversification rates. Major changes in habitat among animal phyla were also reconstructed to help address how long animals have been present in each habitat.

## MATERIALS AND METHODS

### Time-calibrated phylogeny

Estimating a new animal phylogeny was not the goal of this study. Instead, existing data were used to estimate a time-calibrated tree, and the topology was constrained to reflect well-supported relationships found in previous studies. To do this, an expanded version of the matrix of Philippe *et al.* (2011) was generated. This matrix is modified from that of Dunn *et al.* (2008), which is relatively comprehensive in its sampling of animal phyla (including 27). This matrix included 150 genes from 77 taxa (71 metazoans, 6 outgroups) mostly from ribosomal proteins and ESTs (expressed sequence tags). Philippe *et al.* (2011) improved upon this dataset by correcting alignment errors and adding new sequences. To this data matrix, two more putative phyla (Acanthocephala, Placozoa) were added here (Tables S1 and S2). Full methods are provided in Appendix S1.

A time-calibrated phylogeny was generated using the Bayesian relaxed lognormal approach in BEAST, version 2.1.3 (Bouckaert *et al.* 2014). Given that analysing the large number of genes and taxa simultaneously would have been computationally challenging, the dataset was reduced to 16 relatively complete genes. The six non-metazoan outgroups were removed. PartitionFinder v1.1.1 (Lanfear *et al.* 2014) was used to simultaneously find the best-fitting set of partitions and best-fitting model for each partition (Table S3).

Two sets of topological constraints were used. The first set was based primarily on the tree of Dunn *et al.* (2014), with other relationships constrained based on the trees of Dunn *et al.* (2008) and Philippe *et al.* (2011). The tree of Dunn *et al.* (2014) provides a summary of animal phylogeny based on many recent studies. The second set was based primarily on Philippe *et al.* (2011). These constraints are described in Appendix S1.

To estimate clade ages across the tree, 20 fossil calibration points were used (Appendix S1). For the tree based on Dunn *et al.* (2014), two sets of calibration points were used, one forcing a younger age for the clade uniting Porifera (sponges) with other animals (Appendix S1). Therefore, a total of three phylogenies were generated from these analyses. Furthermore, an additional tree was also considered, based on Erwin *et al.* (2011), which included fewer animal phyla (18). Most analyses were repeated across all four trees.

Importantly, even though there is still uncertainty in the relationships among animal phyla and their ages, the results show that the relationships between diversification rate and habitat remain significant across different trees and clade ages.

### Species richness of clades across habitats

At present there is no comprehensive summary of species richness of each animal phylum in marine vs. non-marine habi-

tats, although overall richness patterns have been estimated by several authors (e.g. May 1994; Mora *et al.* 2011; Costello *et al.* 2012). Grosberg *et al.* (2012) summarised information on species richness in each habitat, but included only 10 animal phyla, and did not provide supporting references. Here, data were assembled on the numbers of extant, described species in each habitat for all 32 currently recognised extant animal phyla, using literature sources and databases (Table S4). Details are described in Appendix S2.

Importantly, all estimates of species richness are underestimates of actual richness. However, these analyses merely assume that the proportions of species in different habitats will remain roughly similar (e.g. more species in terrestrial than marine habitats), even as more species are discovered and described. This assumption is broadly supported by the concordance between my estimates based on described species with those incorporating undescribed species (see Results).

### Diversification rates

The net rate of diversification for each phylum-level clade was estimated using the method-of-moments estimator for stem group ages (Magallón & Sanderson 2001). The stem-group estimator was used (in part) because the phylogeny included too few species for some phyla to represent their crown group age (e.g. some species-poor phyla are represented by single species). More importantly, the stem group incorporates the entire history of the group (from the first split from its sister taxon to the present day), whereas the crown group might represent only a very recent diversification of extant species (e.g. a group with a stem age of 200 Ma might have a crown age of only 20 Ma, especially if richness was greatly reduced by extinction). Therefore, the stem-group diversification rate better represents the overall net diversification of the clade. Estimating the diversification rate of a clade using the method-of-moments estimator requires both clade age and richness as inputs, and an assumed relative extinction rate ( $\epsilon$ , or  $\epsilon$ ). The relative extinction rate is the extinction rate divided by the speciation rate (Magallón & Sanderson 2001). This ratio is typically assumed given that these rates are difficult to estimate separately, especially without relatively complete phylogenies within clades. To address the robustness of the results to different values of epsilon, two extreme values ( $\epsilon = 0$  and 0.9) and an intermediate value ( $\epsilon = 0.50$ ) were used. However, 0.9 generally had the best fit (e.g. lowest AIC, see below). This specific value (0.9) was chosen following Magallón & Sanderson (2001).

Some authors have criticised the method-of-moments estimator because of the potential variability in diversification rates within clades over time (Rabosky 2009). However, this estimator should nevertheless reflect net diversification rates (Wiens 2011): older clades with few extant species will have lower rates, and younger clades with many species will have higher rates, regardless of variation in richness or diversification within clades over time. Furthermore, this estimator remains widely used, even by its critics (e.g. Rabosky & Matute 2013). Some authors have suggested using ln-transformed richness instead of net diversification rates to deal with the problem of variability in rates within clades over time (e.g.

Rabosky 2009). Although these arguments are controversial, a relationship between the proportion of non-marine species and ln-transformed richness of clades was also tested. Furthermore, variation in diversification rates within clades over time could potentially uncouple diversification rates from richness patterns (e.g. if younger clades have higher rates, but their high rates are not sustained over time and so do not lead to high richness; Wiens 2011). Therefore, the relationship between the diversification rates of clades and their richness was also tested for all analyses (Wiens 2011).

In general, many approaches are available for estimating diversification rates (recent review in Morlon 2014). However, most alternate approaches would have required more detailed time-calibrated phylogenies within each group with more complete taxon sampling, which would have been difficult for many phyla. Furthermore, alternate methods would not necessarily give very different estimates of net diversification rates for these clades, even if more detailed phylogenies were available.

Finally, it is important to note that these analyses are based on numbers of described species, and the actual numbers of species in these clades are doubtless much greater. However, the analyses here do not require that the actual numbers are known. Instead, they merely assume that the relative richness of clades will remain similar as more species are described (e.g. more arthropod than cnidarian species), not that the actual richness is identical to the number of currently described species.

#### Testing the relationship between habitat and diversification

The relationship between occurrence in non-marine environments and rates of diversification among phyla was tested using phylogenetic generalised least-squares regression (PGLS; Martins & Hansen 1997). This approach accounts for the potential statistical non-independence of phyla due to phylogeny. For each phylum, the proportion of species in non-marine environments (freshwater + terrestrial) relative to the total extant, described richness was calculated. For PGLS, the R package CAPER version 0.5.2 (Orme 2013) was used, with the maximum likelihood transformation of branch lengths optimised for the data ('lambda = ML'). Estimated values of lambda (Pagel 1999) were used, and kappa and delta were fixed at 1. The relationship between the proportion of non-marine species and either diversification rates or richness were tested (diversification rates:  $\epsilon = 0, 0.5$  and  $0.9$ , and ln-transformed richness; Table S14). These proportions were not arcsine-transformed, since this transformation can be problematic (Warton & Hui 2011). However, preliminary analyses using this transformation gave similar results to those using untransformed data.

The relationships between the proportion of non-marine species in each clade and the diversification rate of that clade were estimated among the 28 phyla using the three phylum-level topologies. The robustness of these relationships was then tested with several additional analyses.

First, many phyla were subdivided to yield a total of 49 higher-level clades (see Appendix S3). The species richness of each clade in each habitat was then estimated (Appendix S3;

Tables S9 and S10) as were the ages and diversification rates of these clades for each of the three trees (Tables S11–S13). The relationships between the proportions of non-marine species and diversification rates were then tested, as described earlier. These analyses were intended to ensure that the main results were not an artefact of the division of animal species into phyla. Also, many subclades within phyla differ dramatically in their richness and proportions of marine and non-marine species. The subclades are also younger (mean  $\sim 500$  Myr old vs.  $\sim 600$  Myr old for phyla). However, this analysis of 49 clades was not the primary analysis, since they do not include as much of animal richness as phyla. Furthermore, it may be difficult to simultaneously explain most variation both within and between phyla using a single ecological variable.

Similar analyses were also conducted on the time-calibrated phylogeny of animals from Erwin *et al.* (2011). The species included in that tree were subdivided into 41 subclades (see Appendix S4; Table S14). The richness of these subclades in each habitat was then estimated (Appendix S4; Table S15), the stem-group ages of clades were obtained from their tree (Table S16), diversification rates were estimated (Table S16), and the relationship between occurrence in non-marine habitats and diversification rates was tested as described earlier. However, this analysis was not used as the primary focus of this study because Erwin *et al.* (2011) only included 18 animal phyla. All trees used in the PGLS analyses are given in nexus format in Appendices S6–S15.

A set of analyses was also conducted to test the relationship between habitats and diversification within three species-rich phyla (Annelida, Arthropoda and Mollusca), using similar methods (Appendix S17). These phyla have the largest number of subclades among the 49, but sample sizes were still small ( $n = 5, 6$  and  $5$  respectively). Analyses were repeated for all three trees.

The robustness of the results was also tested after placing freshwater species in the same habitat category with marine species. Specifically, these analyses tested whether diversification rates were higher in terrestrial clades vs. aquatic clades, considering both freshwater and marine species to be aquatic. Assigning species to freshwater vs. terrestrial environments was difficult in some cases, but for a relatively small number of species overall (Appendix S2). After re-assigning species to each category, the relationship was tested between the proportion of aquatic species (marine + freshwater) and diversification rates and richness (diversification rates:  $\epsilon = 0, 0.5$  and  $0.9$ , and ln-transformed richness), for all three topologies and for 28 phyla and 49 subclades. These analyses were also repeated using the 41 clades from Erwin *et al.* (2011).

#### Reconstructing ancestral habitats

The evolution of habitats among the sampled metazoans was estimated and visualised using maximum likelihood in Mesquite 2.75 (Maddison & Maddison 2011). For this analysis, the three full trees of 73 taxa were utilised. A data matrix was constructed by coding the habitat of each sampled species as primarily 0 (marine), 1 (freshwater) or 2 (terrestrial), based on literature sources (Table S2). The analyses used a single

rate for all changes between states (Mk1 model), the most straightforward model given three or more states. A node was considered ambiguously reconstructed (i.e. poorly supported) if the proportional likelihood of the most likely state was 0.87 or less, following the standard threshold utilised in Mesquite 2.75. Species sampling was clearly limited relative to the overall number of species across these clades, which could lead to errors (especially within phyla). However, many clades are entirely or mostly invariant for habitat, and the reconstructed state for each phylum generally matched the most common state among the species in that phylum (Fig. 1). I did not attempt to account for the potential influence of habitat-specific diversification rates on these ancestral-state reconstructions (e.g. with BiSSE; Maddison *et al.* 2007), because species-level sampling was limited and not proportional to richness of phyla (FitzJohn *et al.* 2009). However, it seems unlikely that habitat-specific diversification differences would overturn the overall reconstruction results here (e.g. a bias towards higher non-marine diversification rates should presumably make it more difficult to reconstruct a marine ancestor for animals, even though this is what was found).

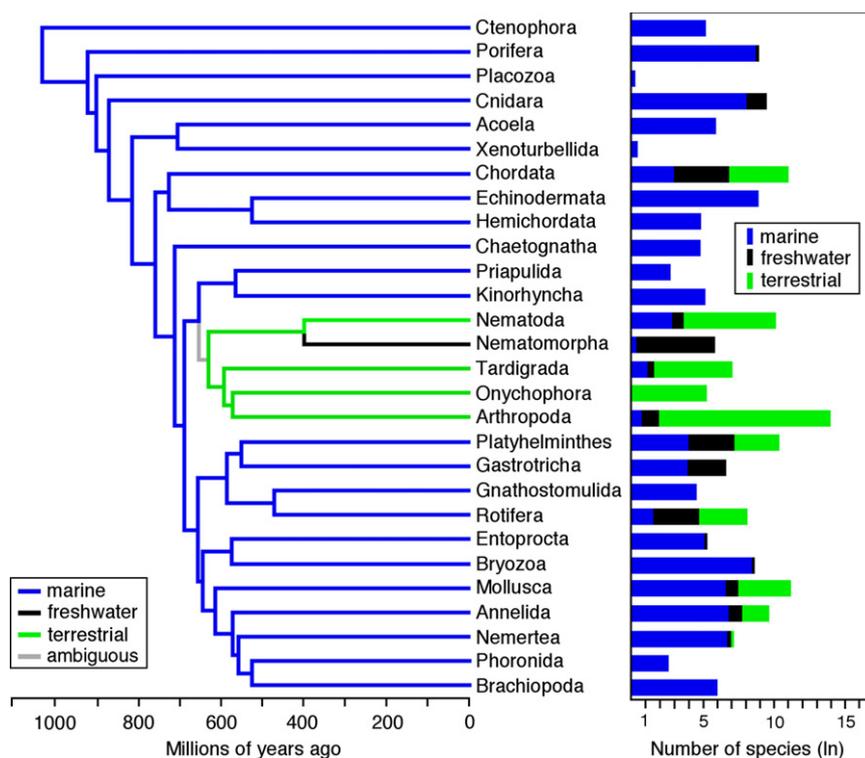
## RESULTS

Patterns of species richness among habitats, among phyla and across the phylogeny are summarised in Fig. 1. Based on the estimates here, 76.8% of all described animal species occur in terrestrial environments, whereas 12.4% occur in marine envi-

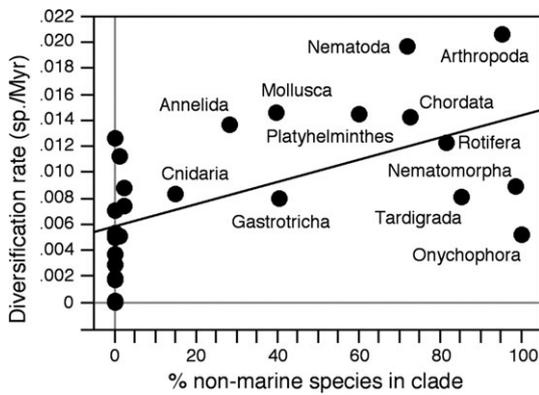
ronments (Table S4). Approximately 10.9% occur in freshwater (either partially or exclusively), which is surprising given that these habitats occupy only ~2% of Earth's surface (May 1994). Overall, these estimates of marine vs. non-marine richness for animals are similar to those for all eukaryotic species, including all described species and all estimated undescribed species (25.3% marine vs. 74.7% non-marine; Mora *et al.* 2011). These numbers are doubtless underestimates of actual richness, but similar proportions of undescribed species have been estimated for each habitat (91% of species are estimated to be undescribed in marine habitats vs. 86% in non-marine habitats; Mora *et al.* 2011). However, some authors have suggested that Mora *et al.* (2011) overestimated marine richness (e.g. Appeltans *et al.* 2012), and that only ~15% of living species are marine (Costello *et al.* 2012).

Phylogeny-based regression analyses demonstrate that clades with more non-marine species have significantly higher diversification rates (Fig. 2; Tables S17–S19). Clades with a majority of non-marine species have diversification rates roughly twice that of those with a majority of marine species (Tables S17–S19). Habitat explains ~30–37% of the variation in diversification rates among animal phyla (note: here and throughout the results, the range of values summarises results across the three trees and across the three relative extinction fractions for each tree).

In turn, rates of diversification are strongly related to patterns of richness, with variation in diversification rates explaining ~85–89% of the variation in species richness



**Figure 1** Time-calibrated phylogeny of 28 animal phyla (topology based on Dunn *et al.* 2014) showing maximum likelihood reconstructions of ancestral habitats and richness of phyla in different habitats. Ancestral reconstructions were done on the full tree of 73 taxa (Fig. S1), and are merely summarised here. Reconstructions using two alternative trees are shown in Figs S2 and S3. Note that the number of origins of terrestriality have likely been underestimated here (e.g. in ecdysozoans; Kenrick *et al.* 2012; Rota-Stabelli *et al.* 2013), but these results are ambiguous (Fig. S2) and have little consequence for the main conclusions.



**Figure 2** Relationship between diversification rates of 28 animal phyla and their proportion of non-marine species ( $r^2 = 0.351$ ;  $P = 0.0009$ ), based on the tree in Fig. 1. Results shown are based on the raw data for easier visualisation, but phylogenetically corrected results are very similar (Table S17). Diversification rates are estimated assuming high extinction relative to speciation ( $\epsilon = 0.9$ ), but relationships are similar assuming lower fractions (Table S17). Results from alternative trees are shown in Tables S18 and S19.

among clades (Tables S20–S22). The results here also show significant positive relationships between species richness of clades and their occurrence in non-marine environments ( $r^2 = 0.24$ ,  $P = 0.0019$ ; Tables S17–S19).

The overall pattern of higher non-marine diversification is driven partly by arthropods, which are predominantly terrestrial and include most animal species (Fig. 2; Table S4). However, the relationship is robust to excluding this phylum (Appendix S16). Furthermore, other largely terrestrial or non-marine clades have similar diversification rates (e.g. Chordata, Nematoda, Platyhelminthes), as do some clades with substantial diversity in both marine and non-marine environments (e.g. Annelida, Mollusca; Fig. 2; Table S4). Conversely, several exclusively marine phyla with low diversification rates also contribute to the pattern (e.g. Acoela, Chaetognatha, Ctenophora, Gnathostomulida, Hemichordata, Kinorhyncha, Phoronida, Priapulida and Xenoturbellida; Table S4). Importantly, even though four phyla were not included in the phylogeny, they nevertheless follow this general pattern: three are low-diversity marine clades (Cyclophora: 2 species, Locifera: 26 and Mesozoa: 147), whereas the fourth is species poor and occurs in freshwater (Micrognathozoa: 1 species; Table S4).

The relationship between habitat and diversification is robust to subdividing the 28 sampled phyla into a total of 49 subclades (Tables S23–S25). However, the relationship between diversification rates and non-marine environments is weaker ( $r^2 = 0.13$ – $0.18$  under the best-fitting relative extinction fraction), as expected given the problem of simultaneously explaining diversification patterns both within and between phyla. The relationship between non-marine habitats and species richness remains significant among the 49 clades ( $r^2 = 0.22$ ;  $P < 0.0001$ ; Tables S23–S25), as does the relationship between diversification rates and richness ( $r^2 = 0.56$ – $0.86$ ;  $P < 0.0001$ ; Tables S26–S28). The relationship between non-marine habitats and diversification is also significant

using an additional alternative phylogeny (Erwin *et al.* 2011), containing fewer phyla but including 41 higher level clades ( $r^2 = 0.30$ ,  $P < 0.0001$ ; under the best-fitting relative extinction fraction; Table S29), as is the relationship between diversification and richness (Table S30).

A positive relationship between non-marine habitat and diversification is also present within Annelida, Arthropoda and Mollusca (Appendix S17). In Annelida, this relationship is heavily influenced by an outlier (Errantia), but is otherwise strong. In Arthropoda, the relationship merely approaches significance ( $P = 0.08$ – $0.09$ ), but habitat nevertheless explains substantial variation in diversification rates ( $r^2 \geq 0.50$ ). In Mollusca, the relationship is consistently strong and significant ( $r^2 = 0.86$ – $0.91$ ;  $P = 0.01$ – $0.02$ ).

The relationship between habitat and diversification is also robust to repartitioning species into aquatic and terrestrial categories, instead of marine and non-marine categories. This relationship is relatively strong among the 28 phyla ( $r^2 = 0.25$ – $0.28$ ;  $P < 0.0001$ ; Tables S31–S33), the 49 clades ( $r^2 = 0.18$ – $0.25$ ;  $P < 0.0001$  for the best-fitting relative extinction fraction;  $r^2 = 0.10$ – $0.19$ ;  $P = 0.0001$ – $0.0093$  for the others; Tables S34–S36) and the 41 clades in the Erwin *et al.* (2011) tree ( $r^2 = 0.35$ – $0.36$ ;  $P < 0.0001$ ; Table S37). Interestingly, across phyla, variation in diversification rates is explained more strongly by non-marine habitats than terrestrial habitats ( $r^2 = 0.30$ – $0.37$  vs.  $0.25$ – $0.28$ ), but terrestrial habitats are more strongly related to diversification when phyla are subdivided (using 49 clades: relationship with non-marine habitats:  $r^2 = 0.06$ – $0.18$  and relationship with terrestrial habitats:  $r^2 = 0.10$ – $0.25$ ; for 41 clades: relationship with non-marine habitats:  $r^2 = 0.19$ – $0.30$  and relationship with terrestrial habitats:  $r^2 = 0.35$ – $0.36$ ). Concordantly, a recent study showed a strong relationship between terrestrial habitats and diversification rates among major vertebrate clades, but no significant relationship with non-marine habitats (Wiens 2015).

Maximum likelihood reconstruction of habitats across the phylogeny confirms the pattern expected from the fossil record, with marine habitats being the ancestral state for the deepest clades within animals (Fig. 1). There were then many independent invasions of terrestrial and freshwater environments in various clades, including annelids, arthropods, chordates, molluscs and platyhelminths (Figs S1–S3). However, the number of terrestrial invasions is most likely underestimated on this tree (e.g. in ecdysozoans; Kenrick *et al.* 2012; Rota-Stabelli *et al.* 2013).

## DISCUSSION

The results of this study show that predominantly non-marine animal clades have higher rates of diversification than predominantly marine clades. This finding has two major implications. First, it helps explain why non-marine and terrestrial environments have greater species richness, a dramatic gradient in global biodiversity. Second, the results help explain the striking variation in species richness among animal phyla, variation which has been largely unexplained. The results also have broader implications for studies of diversification and niche conservatism.

The results here help shed light on the causes of a major gradient in global biodiversity. Specifically, they suggest that non-marine richness dramatically exceeds marine richness because the many clades that have invaded non-marine habitats have accelerated rates of diversification. The amount of variation in diversification rates among phyla that is explained by habitat is substantial but not overwhelming (~30–37%; Tables S17–S9). The unexplained variation in this relationship may be related to predominantly non-marine phyla that have moderate diversification rates (e.g. Onychophora and Tardigrada) and marine clades that have high rates relative to other marine clades (e.g. Bryozoa and Echinodermata; Tables S6–S8). Nevertheless, the significant effect of habitat on diversification rates is striking, given the vast timescale (> 800 million years; Fig. 1) and the diverse morphologies and ecologies of the organisms involved. Specifically, the clades included here range from microscopic to macroscopic size, from obligate endoparasites to free-living organisms and from sessile filter feeders to mobile predators (Hickman *et al.* 2012).

This result is also notable in that some authors have suggested that variation in diversification rates does not explain any large-scale patterns of species richness (Rabosky *et al.* 2012). In fact, variation in diversification rate helps explain two of Earth's most important biodiversity gradients: the greater species richness of non-marine habitats (this study) and the higher richness of tropical regions (e.g. Pyron & Wiens 2013; Rolland *et al.* 2014). Although not all studies have supported higher diversification rates in tropical lineages (e.g. within families of amphibian and mammals; Wiens *et al.* 2006, 2009; Soria-Carrasco & Castresana 2012), studies at broader phylogenetic scales in these same groups have (e.g. Pyron & Wiens 2013; Rolland *et al.* 2014). The results here also show that variation in diversification rates is closely tied to variation in species richness at the largest phylogenetic scales across animals (i.e. among phyla).

The results here also help explain why animal phyla differ dramatically in species richness. Specifically, the results show that ecology (marine vs. non-marine habitat) helps explain why some clades have tens of thousands of species (or more) and others have far fewer. In short, most species-poor clades are marine, whereas the most species-rich clades are dominated by non-marine species.

In theory, the invasion of non-marine habitats might be associated with clades that have a higher intrinsic propensity to diversify, rather than non-marine habitats driving higher diversification rates. However, higher diversification rates occur in clades with higher *proportions* of non-marine species, not greater absolute numbers of non-marine species. Furthermore, the relationship between diversification and non-marine environments remains significant after subdividing many sampled phyla into smaller clades. Thus, it seems most likely that habitat predominantly drives diversification rather than vice versa. Nevertheless, other factors will be critically important in explaining variation in diversification rates that is not explained by habitat (e.g. other ecological or phenotypic traits of clades that influence diversification).

Given the finding here that non-marine environments increase diversification, a key challenge for future research

will be to identify the specific ecological factors in non-marine environments that underlie this pattern. However, some explanations already seem unlikely. For example, explanations based on the physical medium of water alone may not fully explain higher non-marine diversity, given that ray-finned fish have higher richness in freshwater than marine environments and are all aquatic (Carrete-Vega & Wiens 2012). Furthermore, differences in productivity may not be the primary explanation, since net primary productivity is similar in marine and terrestrial environments (e.g. Field *et al.* 1998). Perhaps the most compelling hypothesis so far is that non-marine environments offer more effective barriers to dispersal, which may promote speciation, endemism and geographic turnover (e.g. May 1994; Benton 2001; Vermeij & Grosberg 2010). These barriers include oceans, rivers, mountains and valleys for terrestrial species and terrestrial habitats isolating rivers, streams and lakes for freshwater species. A possible role for higher extinction rates in marine environments should also be investigated, especially given the many ancient, species-poor clades in the oceans.

Some readers may be dissatisfied that this study does not resolve why marine habitats have lower diversification rates. However, it is difficult to test a mechanistic explanation for this pattern in a group including > 1 million species. Most importantly, this study shows that the pattern of higher non-marine diversification applies across vast timescales and numbers of species (even though these scales are not ideal for testing detailed mechanistic explanations). The results here set the stage for smaller-scale ecological studies within these clades to address why diversification rates differ.

An alternative to diversification rates in explaining richness patterns is the idea that habitats colonised earlier have greater time for richness to accumulate through *in situ* speciation (Wiens 2011). Time explains many richness patterns across habitats, especially over shorter time scales (e.g. Kozak & Wiens 2010; Hutter *et al.* 2013; Wiens *et al.* 2013). However, time does not explain higher animal richness in terrestrial environments, given paleontological evidence that animals occurred in marine environments before terrestrial habitats (e.g. Jeram *et al.* 1990; Narbonne 2005; Kenrick *et al.* 2012). Furthermore, reconstructions of ancestral habitats on the phylogeny here are broadly consistent with this pattern (Fig. 1; Figs S1–S3). This result contrasts with the intuition that ancestral reconstructions of ecological traits will necessarily become misleading or ambiguous at the deepest timescales. Here, these deep reconstructions are unambiguous and broadly concordant with paleontological evidence over > 800 million years (Fig. 1).

The reconstructions of ancestral habitats here also offer a striking example of niche conservatism. The ancestral reconstructions (Fig. 1) suggest that many (if not most) phyla occurring in oceans today have ancestors that occurred there continuously for > 800 million years (e.g. Brachiopoda, Ctenophora, Echinodermata, Porifera). This result demonstrates a simple but potentially important quantifier of niche conservatism: the absolute amount of time that members of a clade have occurred continuously in a given environment. This quantifier has been largely ignored in the

literature on niche conservatism, which tends to focus on the fit of characters to a phylogeny rather than the absolute timescales over which traits are conserved (e.g. Cooper *et al.* 2010; Wiens *et al.* 2010). Using this index, the results here suggest that ancient marine animal clades may offer some of the most dramatic examples of niche conservatism yet known.

Finally, another major result of this study is the relatively high estimated species richness of freshwater environments. The summary here shows that for animals, freshwater richness is surprisingly similar to marine richness (163 796 freshwater species vs. 188 040 marine; Table S4), even though freshwater environments cover ~2% of Earth's surface while marine environments cover ~70%. Among animal clades (Table S4), freshwater richness is dominated by arthropods (~118 000 freshwater species, including crustaceans and many insect clades) and chordates (~24 000 species, including ray-finned fish). Assigning species to freshwater vs. terrestrial habitats (or freshwater vs. marine) can sometimes be difficult, especially for species that occur in freshwater habitats for only part of their life cycles (e.g. many insect clades with aquatic larvae). Here, species were considered freshwater if they were dependent on these habitats, even if they were not confined to them for their entire lives. For one clade (Platyhelminthes) species were arbitrarily split between freshwater and terrestrial habitats, but even if this partition is grossly inaccurate, it still accounts for a relatively small fraction of freshwater animal richness overall (~9 000 species). Overall, these results highlight the biodiversity of freshwater habitats (and the importance of their conservation) and suggest that explaining high freshwater richness should be another priority for future studies.

In conclusion, the results of this study help shed light on two major biodiversity patterns: the higher species richness of terrestrial and non-marine environments and the dramatic differences in richness of animal phyla. They also highlight the surprising richness of freshwater habitats and a neglected index of niche conservatism (and some striking examples of niche conservatism over unprecedented timescales). Finally, this study underscores the fact that many animal phyla consist of ancient, depauperate marine clades. Thus, much of the phylogenetic diversity of animals might soon be lost if current anthropogenic change (e.g. acidification, climate change and overharvesting) drives widespread marine extinctions (e.g. McCauley *et al.* 2015).

#### ACKNOWLEDGEMENTS

I thank E. Park for assistance compiling data and preliminary analyses, R. Walls and the iPlant Collaborative for computational assistance, R. Walls and three anonymous reviewers for helpful comments on the manuscript and K. Peterson, D. Pisani and D. Erwin for providing their tree.

#### AUTHORSHIP

JJW designed the study, performed analyses and wrote the paper.

#### REFERENCES

- Appeltans, W., Ah Yong, S.T., Anderson, G., Angel, M.V., Artois, T., Bailly, N. *et al.* (2012). The magnitude of global marine species diversity. *Curr. Biol.*, 22, 2189–2202.
- Benton, M.J. (2001). Biodiversity on land and in the sea. *Geol. J.*, 36, 211–230.
- Bouckaert, R., Heled, J., Kuhnert, D., Vaughan, T., Wu, C.-H. & Xie, D. *et al.* (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.*, 10, e1003537.
- Carrete-Vega, G. & Wiens, J.J. (2012). Why are there so few fish in the sea? *Proc. R. Soc. Lond. B*, 279, 2323–2329.
- Cooper, N., Jetz, W. & Freckleton, R.P. (2010). Phylogenetic comparative approaches for studying niche conservatism. *J. Evol. Biol.*, 23, 2529–2539.
- Costello, M.J., Wilson, S. & Houlding, B. (2012). Predicting total global species richness using rates of species description and estimates of taxonomic effort. *Syst. Biol.*, 61, 871–883.
- Dunn, C.W., Hejnol, A., Matus, D.Q., Pang, K., Browne, W.E., Smith, S.A. *et al.* (2008). Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature*, 452, 745–749.
- Dunn, C.W., Giribet, G., Edgecombe, G.D. & Hejnol, A. (2014). Animal phylogeny and its evolutionary implications. *Ann. Rev. Ecol. Evol. Syst.*, 45, 371–395.
- Erwin, D.H., Laflamme, M., Tweedt, S.M., Sperling, E.A., Pisani, D. & Peterson, K.J. (2011). The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science*, 334, 1091–1097.
- Field, C.B., Behrenfeld, M.J., Randerson, J.T. & Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, 281, 237–240.
- FitzJohn, R.G., Maddison, W.P. & Otto, S.P. (2009). Estimating trait dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst. Biol.*, 58, 595–611.
- Grosberg, R.K., Vermeij, G.J. & Wainwright, P.C. (2012). Biodiversity in water and on land. *Curr. Biol.*, 22, R900–R903.
- Hickman, C., Jr., Roberts, L., Keen, S., Larson, A. & Elsenhour, D. (2012). *Animal Diversity*, 6th edn. McGraw-Hill Higher Education, NY.
- Hutter, C.R., Guayasamin, J.M. & Wiens, J.J. (2013). Explaining Andean megadiversity: the evolutionary and ecological causes of glassfrog elevational richness patterns. *Ecol. Lett.*, 16, 1135–1144.
- Jeram, A.J., Selden, P.A. & Edwards, D. (1990). Land animals in the Silurian: arachnids and myriapods from Shropshire, England. *Science*, 250, 658–661.
- Kenrick, P., Wellman, C.H., Schneider, H. & Edgecombe, G.D. (2012). A timeline for terrestrialization: consequences for the carbon cycle in the Palaeozoic. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 367, 519–536.
- Kozak, K.H. & Wiens, J.J. (2010). Niche conservatism drives elevational diversity patterns in Appalachian salamanders. *Am. Nat.*, 176, 40–54.
- Lanfear, R., Calcott, B., Kainer, D., Mayer, C. & Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol. Biol.*, 14, 82.
- Maddison, W.P. & Maddison, D.R. (2011). Mesquite: a modular system for evolutionary analysis. Version 2.75. Available at: <http://mesquiteproject.org>. Last accessed 30 September 2011.
- Maddison, W.P., Midford, P.E. & Otto, S.P. (2007). Estimating a binary character's effect on speciation and extinction. *Syst. Biol.*, 56, 701–710.
- Magallón, S. & Sanderson, M.J. (2001). Absolute diversification rates in angiosperm clades. *Evolution*, 55, 1762–1780.
- Martins, E.P. & Hansen, T.F. (1997). Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am. Nat.*, 149, 646–667.
- May, R.M. (1994). Biological diversity: differences between land and sea. *Phil. Trans. R. Soc. Lond. B*, 343, 105–111.
- McCauley, D.J., Pinsky, M.L., Palumbi, S.R., Estes, J.A., Joyce, F.H. & Warner, R.R. (2015). Marine defaunation: animal loss in the global ocean. *Science*, 347, 1255641.

- Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B. & Worm, B. (2011). How many species are there on Earth and in the Ocean? *PLoS Biol.*, 9, e1001127.
- Morlon, H. (2014). Phylogenetic approaches for studying diversification. *Ecol. Lett.*, 17, 508–525.
- Narbonne, G.M. (2005). The Ediacara biota: Neoproterozoic origin of animals and their ecosystems. *Ann. Rev. Earth Plan. Sci.*, 33, 421–442.
- Orme, D. (2013). The caper package: comparative analysis of phylogenetics and evolution in R. R package version 0.5.2.
- Orme, C.D.L., Quicke, D.L.J., Cook, J.M. & Purvis, A.M. (2002). Body size does not predict species richness among the metazoan phyla. *J. Evol. Biol.*, 15, 235–247.
- Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401, 877–884.
- Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G. *et al.* (2011). Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biol.*, 9, e1000602.
- Pyron, R.A. & Wiens, J.J. (2013). Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proc. R. Soc. B Biol. Sci.*, 280, 20131622.
- Rabosky, D.L. (2009). Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecol. Lett.*, 12, 735–743.
- Rabosky, D.L. & Matute, D.R. (2013). Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds. *Proc. Natl Acad. Sci. USA*, 110, 15354–15359.
- Rabosky, D.L., Slater, G.J. & Alfaro, M.E. (2012). Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS Biol.*, 10, e1001381.
- Ricklefs, R.E. (1987). Community diversity: relative roles of local and regional processes. *Science*, 235, 167–171.
- Rolland, J., Condamine, F.L., Jiguet, F. & Morlon, H. (2014). Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biol.*, 12, e1001775.
- Rota-Stabelli, O., Daley, A.C. & Pisani, D. (2013). Molecular timetrees reveal a Cambrian colonization of land and a new scenario for ecdysozoan evolution. *Curr. Biol.*, 23, 392–398.
- Soria-Carrasco, V. & Castresana, J. (2012). Diversification rates and the latitudinal gradient of diversity in mammals. *Proc. R. Soc. B Biol. Sci.*, 279, 4148–4155.
- Vermeij, G.J. & Grosberg, R.K. (2010). The great divergence: when did diversity on land exceed that in the sea? *Int. Comp. Biol.*, 50, 675–682.
- Warton, D.I. & Hui, F.K. (2011). The arcsine is asinine: the analysis of proportions in ecology. *Ecology*, 92, 3–10.
- Wiens, J.J. (2011). The causes of species richness patterns across space, time, and clades and the role of 'ecological limits'. *Q. Rev. Biol.*, 86, 75–96.
- Wiens, J.J. (2015). Explaining large-scale patterns of vertebrate diversity. *Biol. Lett.*, 11, 20150506.
- Wiens, J.J., Graham, C.H., Moen, D.S., Smith, S.A. & Reeder, T.W. (2006). Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: treefrog trees unearth the roots of high tropical diversity. *Am. Nat.*, 168, 579–596.
- Wiens, J.J., Sukumaran, J., Pyron, R.A. & Brown, R.M. (2009). Evolutionary and biogeographic origins of high tropical diversity in Old World frogs (Ranidae). *Evolution*, 63, 1217–1231.
- Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V. *et al.* (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.*, 13, 1310–1324.
- Wiens, J.J., Kozak, K.H. & Silva, N. (2013). Diversity and niche evolution along aridity gradients in North American lizards (Phrynosomatidae). *Evolution*, 67, 1715–1728.
- Zhang, Z. (2013). Animal biodiversity: an update of classification and diversity in 2013. *Zootaxa*, 3703, 5–11.

#### SUPPORTING INFORMATION

Additional Supporting Information may be downloaded via the online version of this article at Wiley Online Library ([www.ecologyletters.com](http://www.ecologyletters.com)).

Editor, Howard Cornell

Manuscript received 29 June 2015

First decision made 4 August 2015

Manuscript accepted 7 August 2015