

## LETTER

# The origins of global biodiversity on land, sea and freshwater

Cristian Román-Palacios<sup>1</sup>  | Daniela Moraga-López<sup>2</sup> | John J. Wiens<sup>1</sup> <sup>1</sup>Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona, USA<sup>2</sup>Departamento de Ecología, Pontificia Universidad Católica de Chile, Santiago, Chile**Correspondence**

John J. Wiens, Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721-0088, USA.

Email: [wiensj@email.arizona.edu](mailto:wiensj@email.arizona.edu)**Funding information**

U.S. National Science Foundation, Grant/Award Number: DEB 1655690

**Editor:** Jonathan Chase**Abstract**

Many biodiversity studies focus on explaining high tropical species richness, but an equally dramatic yet understudied pattern involves the divergent richness of land, sea and freshwater. Here, we reveal the origins of these richness differences among habitats across animals and plants. Most plant and animal species are terrestrial, although these habitats cover only ~28% of Earth's surface. Marine habitats have fewer species over a larger area (~70%). Freshwater habitats have relatively high richness and exceptional phylogenetic diversity given their tiny area (2%). The relative richness of habitats is related to variation in diversification rates. Based on ancestral reconstructions of habitat, we find that most marine species are descended from marine ancestors and most terrestrial species from freshwater ancestors. Yet, most extant animal richness in freshwater is derived from terrestrial ancestors. Overall, our results reveal the origins of fundamental but neglected biodiversity patterns, and highlight the conservation importance of freshwater habitats.

**KEYWORDS**

animals, diversification, freshwater, marine, plants, species richness, terrestrial

## INTRODUCTION

Understanding patterns of biodiversity is a central challenge in ecology and evolutionary biology. Yet, studies of this topic have often focused on one particular biodiversity pattern: the higher species richness of the tropics relative to the temperate zone. The differences in richness among the three major habitat types on Earth (marine, freshwater and terrestrial) are no less dramatic but are far less understood. Several authors (e.g. Grosberg et al., 2012; May, 1994; Mora et al., 2011) have noted that most of Earth's surface is covered by ocean (~70%) but terrestrial and freshwater habitats together have a far higher percentage of Earth's macroscopic species richness (~80%).

Few studies have tested the causes of these patterns, especially in terms of the processes that directly cause richness gradients: speciation, extinction and dispersal (Ricklefs, 1987). For example, a previous study found faster rates of diversification (speciation minus extinction) in predominately terrestrial animal phyla than in largely marine ones (Wiens, 2015a), but patterns in

plants and freshwater were not addressed. There have also been important reviews of when each habitat was colonised by different groups of plants and animals, and related topics (Benton, 2010; Labandeira, 2005; Martín-Closas, 2003; Miller & Labandeira, 2002; Vermeij, 2017; Vermeij & Dudley, 2000; Vermeij & Grosberg, 2010). However, no studies have quantitatively examined the relative roles of diversification and dispersal among habitats in explaining richness patterns in all three habitats (marine, terrestrial and freshwater) across both animals and plants.

Freshwater diversity is especially intriguing and understudied. Freshwater habitats cover only ~2% of Earth's surface but have similar species richness to marine habitats, at least for animals (Grosberg et al., 2012; Reid et al., 2019; Wiens, 2015a). Furthermore, although animals and plants likely arose in marine environments (e.g. Grosberg et al., 2012; Wiens, 2015a), the origins of freshwater diversity are far less clear. For example, it remains unclear whether most freshwater species are descended from marine or from terrestrial ancestors. Surprisingly, the origins of most terrestrial plant and

animal diversity (i.e. predominantly from freshwater vs. marine ancestors) are also unresolved.

Here, we use phylogenetic approaches to analyse the origins of extant species richness of plants and animals in marine, terrestrial and freshwater habitats. We first quantify richness patterns among habitats for plants and animals. We also examine patterns of phylogenetic diversity, and how they are shaped by different clades in different habitats. Next, we test whether the relative richness of habitats is explained by differences in diversification rates (i.e. more species in habitats in which species proliferate more rapidly), colonisation frequency (more species in habitats that were colonised more frequently) or colonisation times (more species in habitats that were colonised earlier). Finally, we quantify the sources of richness in each habitat. Specifically, we address whether most terrestrial species are derived from freshwater or marine ancestors, and whether most freshwater species descended from marine or terrestrial ancestors. Our analyses are necessarily coarse-grained (e.g. complete species-level trees for plants and animals are lacking), but this need not overturn our conclusions.

## MATERIAL AND METHODS

### Habitat definitions

We initially considered freshwater species to be those with all or part of their life cycle in freshwater (<0.05% dissolved salt; Balian et al., 2008). Marine species live all or part of their lives in the ocean or brackish environments (>0.05% dissolved salt; WoRMS, 2020). Terrestrial species spend their entire life cycle on land. Parasitic species were classified based on host habitat. We also explored how our results were impacted by treating semi-aquatic species as terrestrial (see below).

### Quantifying richness within habitats

We estimated richness in freshwater, marine and terrestrial habitats for animals and plants (Table S1). First, we estimated the richness of each animal phylum (Data S1) and plant phylum in each habitat (Data S2). For animals, we used previously compiled data (Wiens, 2015a) based on the same habitat definitions and sources used here. These sources included marine richness from the World Register of Marine Species (WoRMS, 2020), and freshwater richness from Balian et al. (2008). The remaining species were assigned to terrestrial environments. Plants (Plantae) included land plants, glaucophytes and green and red algae (Adl et al., 2005; Roskov et al., 2020). Taxonomy followed Roskov et al. (2020). We estimated plant richness in freshwater and marine habitats using several sources (Chambers et al., 2008; Dawes, 1998; Guiry et al., 2014). All other species were

considered terrestrial. Data were obtained from the WoRMS database (2020) using the R packages *taxize* version 0.9.9 (Chamberlain & Szöcs, 2013) and *worms* version 0.4.2 (Chamberlain, 2019). We retrieved all valid genera in WoRMS using *taxize* and all species records for each genus using *worms*. We then summed numbers across phyla to estimate overall richness in each habitat. We also identified the most species-rich phyla in each habitat.

### Phylogenetic trees

We assembled three types of time-calibrated phylogenies (Table S2; details in Appendix S1). First, we assembled phylogenies for phyla of animals (28 of 32 phyla; 88%; Data S1, S3) and plants (100% of 8 phyla; Data S2, S4). These phyla encompass >99% of extant animal species (99.98% of 1,515,954 described species) and 100% of described plant species (387,489 species). Species numbers followed Roskov et al. (2020). Phylum-level trees have one tip per phylum. Next, we assembled family-level phylogenies for animals (1697/5700 families, 30%; Data S5–S6) and plants (714/958 families, 75%; Data S7–S8), with one tip per family. These families include 72% and 95% of described animal and plant species. Third, we used trees in which tips (mostly families) were sampled in proportion to the richness of phyla and major groups within phyla (animals: 1,074 tips, Table S3, Data S9–S10; plants: 729 tips, Table S4, Data S11–S12). These sampled families include 71% of animal species and 95% of plant species. These proportionally sampled trees should be largely equivalent to using 1074 or 729 randomly sampled species as tips (i.e. similar branch lengths, with more species from larger phyla). We recognise that phyla and families are arbitrary. We only used them in analyses of diversification rates (which allowed us to incorporate all species in each taxon, without a complete species-level tree). We also performed rank-free diversification analyses (see below).

### Habitat coding

For each tip, we estimated species richness in each habitat using the sources above. Next, tips were coded based on the most frequent habitat among species within that phylum (phylum-level tree) or family (family-level, proportionally sampled). For example, if the majority (or plurality) of species in a family were freshwater, the family was coded as freshwater. If multiple taxa were sampled within a family, these taxa were coded based on the habitats of those terminal taxa (e.g. genus, species). For most analyses, this should yield results equivalent to sampling the most common state among species in each clade. Furthermore, simulations suggest that the most common state in a clade should generally be the ancestral state (Wiens, 1998).

In the proportionally sampled trees, phyla represented by a single family were generally coded based on the most frequent habitat among species in the phylum. For two phyla (Platyhelminthes and Rotifera), habitat states were based on the sampled families. However, these two phyla consist mostly of internal parasites, and estimated frequencies were broadly similar across habitats (Platyhelminthes: marine = 40%, freshwater = 30% and terrestrial=30%; Rotifera: marine = 19%, freshwater = 39% and terrestrial = 42%; Data S1).

Discrete habitat states were necessary for phylogenetic methods that infer habitat-specific rates (speciation, extinction and transitions) and ancestral states. Habitat data for the phylum-level, family-level and proportionally sampled trees for animals are provided in Data S1, S5 and S9 for animals and Data S2, S7 and S11 for plants.

### Estimating phylogenetic diversity within habitats

We expect phylogenetic and species diversity to generally be related to each other, and that both will therefore be related to habitat area (all else being equal; MacArthur & Wilson, 1967). Habitats with high phylogenetic diversity per unit area are potentially high priorities for conservation (e.g. Faith, 1992). We estimated phylogenetic diversity within habitats based on the proportionally sampled trees. For each tree, we constructed three subtrees using the function “drop.tip” in the R package *ape* version 5.5 (Paradis & Schliep, 2018). Each subtree contained only the tips assigned to that habitat. We then used the R package *picante* version 1.8.2 (Kembel et al., 2010) to calculate the sum of phylogenetic differences (branch lengths) among these taxa as a metric of phylogenetic diversity (equivalent to Faith's [1992] metric). Phylogenetic diversity within each habitat was then divided by the estimated surface area of each habitat (Gleick, 1996; Grosberg et al., 2012). Since the trees used are proportionally sampled, they should yield patterns of *relative* phylogenetic diversity among habitats similar to completely sampled, time-calibrated trees that included all described, extant animals and plants.

### Estimating the timing of habitat colonisation

We performed ancestral-state reconstructions to test whether the timing of habitat colonisation can explain patterns of richness among habitats. Reconstructions were performed on the proportionally sampled phylogenies. Each terminal taxon was assigned one state (marine, freshwater or terrestrial), based on the most common state among its species.

We performed maximum-likelihood reconstructions using the function “ace” (type = “discrete”) in *ape* (Paradis & Schliep, 2018). We first identified the best-fitting model for each tree. We compared three models.

The all-rates-different model (ARD) allowed for different transition rates to and from each pair of states, yielding six rate parameters among the three habitats. The symmetrical-rates model (SYM) had different transition rates between each pair of habitats (three rate parameters), but the same rate in either direction (i.e. terrestrial-to-freshwater equals freshwater-to-terrestrial). The equal-rates model (ER) assumed a single rate for all transitions between all states. When using an ER model, different colonisation patterns between habitats can still be inferred from ancestral-state reconstructions (e.g. more transitions to than from freshwater). For each model, we estimated the Akaike Information Criterion (AIC; Akaike, 1974). The best-fitting model had the lowest AIC. For AIC differences <2 between best-fitting models, we used the next-best model with fewer parameters (Burnham & Anderson, 2002).

For each tree, we estimated the ancestral habitat for each node as the habitat with the highest marginal probability. We then determined the age of the oldest node reconstructed for each habitat. Finally, we examined whether the habitat colonised first had the highest present-day richness. No statistical testing was conducted given that there is only a single datapoint for the oldest colonisation time for each group (plants, animals) in each habitat.

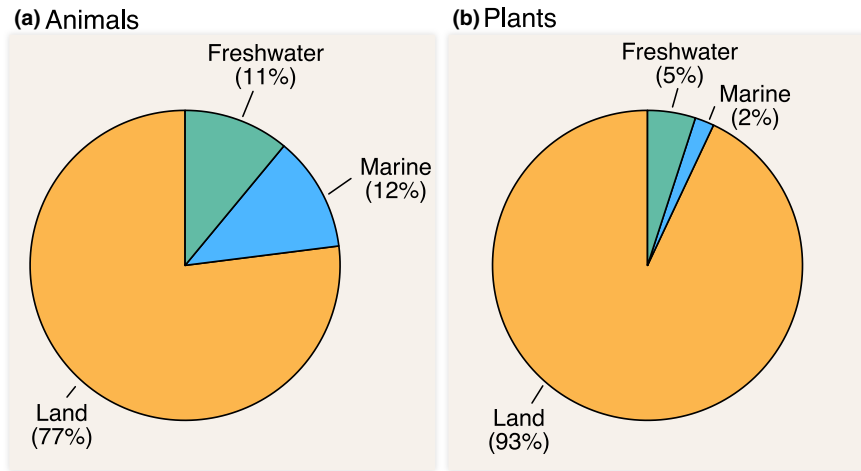
### Paleontological estimates of colonisation times

We also reviewed paleontological evidence for the oldest occurrence of plants and animals in each habitat. We describe the methods in Appendix S2.

### Diversification rates among habitats

We used two general approaches to estimate the relationships between habitats and diversification rates. As a first approach, we estimated net diversification rates of clades (families, phyla) using the method-of-moments estimator for stem-group ages (MS estimator; Magallón & Sanderson, 2001). We estimated the mean diversification rate across all families and phyla assigned to a given habitat. We then evaluated whether differences in the mean diversification rates among clades in each habitat paralleled the relative richness of habitats (e.g. highest mean rates in predominantly terrestrial phyla). We also tested whether habitat predicted diversification rates among clades using phylogenetic regression and ANOVA (details in Appendix S3).

We used the MS estimator because it can estimate diversification rates for clades without detailed, time-calibrated, species-level phylogenies within each clade. These were unavailable for many clades but are necessary for most alternative methods. Based on simulations, the MS estimator is accurate under many conditions (Kozak



**FIGURE 1** Distribution of species richness among habitats for animals and plants. In animals (a), the 1.516 million extant, described species occur mostly on land (1,164,269 species), with fewer in marine habitats (187,890 species), and freshwater (163,795 species). In plants (b), most of the 387,489 extant, described species occur on land (361,210 species), with far fewer in freshwater (19,197 species), and marine habitats (7082 species). Habitat richness per phylum is summarised in Data S1 (animals) and Data S2 (plants)

& Wiens, 2016; Meyer et al., 2018; Meyer & Wiens, 2018), including when rates vary over time within clades, when rates vary between subclades within each clade, and when rates are faster in younger clades. Thus, it does not require constant rates within or between clades. Indeed, it allows for more rate variation than most alternative methods. We discuss method selection and criticisms of this method in Appendix S3. Positive relationships between diversification rates of clades and their richness are not inevitable or artefactual (Kozak & Wiens, 2016; Scholl & Wiens, 2016), and the overall richness of habitats and of clades are obviously not the same.

We calculated stem-based net diversification rates for phyla and families using the “bd.ms” function in the R package *geiger* version 2.0.7 (Harmon et al., 2008; Pennell et al., 2014). We used three standard relative extinction fractions ( $\epsilon$ ), the assumed ratio of extinction to speciation ( $\epsilon = 0, 0.5$  and  $0.9$ ). Note that  $\epsilon$  is a correction for unsampled clades due to extinction, and does not require constant extinction rates within or among extant clades (the MS estimator can be accurate when one  $\epsilon$  value is used but extinction rates vary among clades; Meyer & Wiens, 2018).

The MS estimator uses the richness and stem-group age for each clade. Richness for animal and plant phyla are in Data S1–S2, and their diversification rates and ages are in Data S13–S14. Richness for animal and plant families are in Data S5 and S7 and their diversification rates and ages are in Data S15–S16. Additional details on MS-based diversification rates are in Appendix S3.

As a second general approach, we estimated habitat-specific rates of diversification using MuHiSSE (Beaulieu & O’Meara, 2016; Caetano et al., 2018). We then used these rates to inform estimation of ancestral habitat states and transition rates among habitats. Unfortunately, some of the results were problematic, since SSE methods can give misleading results when

some states are rare (Davis et al., 2013), as freshwater and marine habitats are for plants and animals. We describe the methodology and results in Appendix S4.

### Estimating habitat transitions and origins of extant diversity in each habitat

We used ancestral-habitat reconstructions from the proportionally sampled trees to estimate the relative number of transitions between all pairs of habitats (six total). We also used these reconstructions to estimate how much of the extant richness in each habitat was descended from transitions from other habitats, and what those habitats were. For example, we estimated whether most extant terrestrial species were descended most recently from ancestors that were marine versus freshwater. We used the proportionally sampled trees, treating each tip as equivalent to a species (given that the number of tips in each phylum is proportional to its richness). We used the ancestral-state reconstructions for each node described above, and extracted all nodes between the root and each tip using the function “nodepath” in *ape* (Paradis & Schliep, 2018). For each tip, we determined the habitat transition that it descended from most recently based on the most likely habitat state prior to its current habitat. These analyses also allowed us to address whether the most species-rich habitats were those that were colonised most frequently.

### Examining the effects of alternative habitat coding

We examined whether coding semi-aquatic species as terrestrial changed our conclusions. This greatly reduced the number of freshwater animal species, but our overall results remained similar (Appendix S5, Data S17–S20).

## Undescribed species diversity

We address the potentially vast numbers of undescribed species in Appendix S6. We explain why including these projected species is unlikely to overturn our conclusions.

## RESULTS

In animals (Figure 1a), we estimated 77% of described, extant species inhabit land (1,164,269 species; Table S1), 12% marine habitats (187,890 species) and 11% freshwater (163,795 species; Data S1). In plants (Figure 1b), 93% of extant, described species inhabit terrestrial habitats (361,210 species), 5% freshwater (19,197 species) and 2% marine (7082 species; Data S2). We considered species inhabiting both terrestrial and aquatic habitats as freshwater or marine. Analyses considering these semi-aquatic species as exclusively terrestrial yielded similar results (Appendix S5).

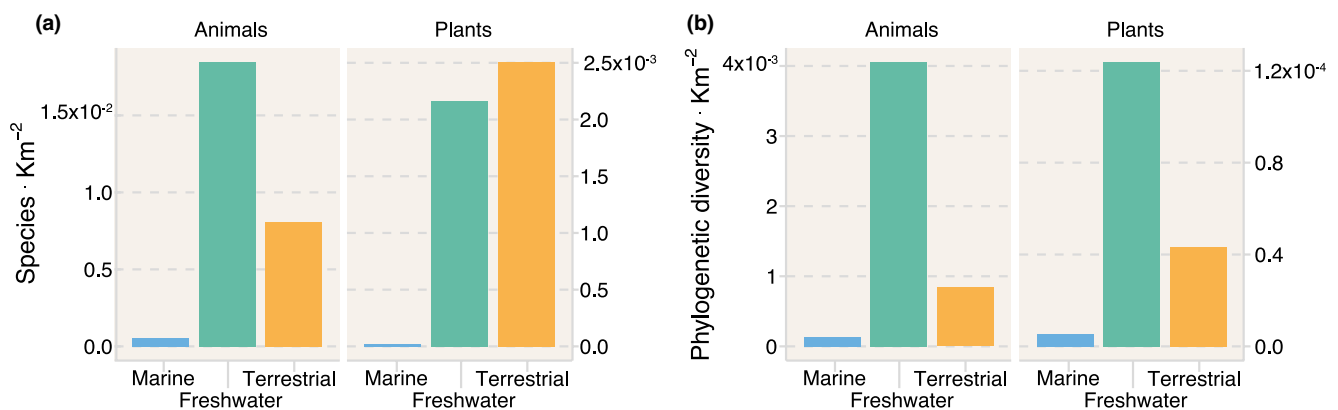
Next, we examined diversity patterns among habitats relative to their surface area. At a coarse scale, ~70% of Earth's surface is marine, ~28% is terrestrial and ~2% is freshwater (Gleick, 1996). Terrestrial habitats contained much higher richness per unit area than marine habitats for both plants and animals (Figure 2a; Table S1). The richness of freshwater habitats per unit area was similar to terrestrial habitats in plants, but was much higher in animals (Figure 2a). For both plants and animals, phylogenetic diversity per unit area in freshwater was much higher than in other habitats (Figure 2b).

We then examined which clades shape diversity most strongly in each habitat (Figure 3; Data S1–S2). Terrestrial richness is dominated by a single phylum in both plants and animals (animals: arthropods = 93% of terrestrial richness; plants: tracheophytes = 94% of terrestrial richness). In marine habitats, animal diversity is more evenly distributed among phyla, with eight phyla that together make up >90% of marine richness (arthropods = 29%, mollusks = 23%, chordates = 10%,

annelids = 7%, platyhelminths = 6%, cnidarians = 6%, sponges = 4%, nematodes = 4%). Marine plants consist mostly of rhodophytes (69% of species) and chlorophytes (27%). In freshwater, the most species-rich animal phyla are arthropods (72% of freshwater animal species, with 85% of these being semi-aquatic; Data S1; Appendix S5), chordates (15%) and platyhelminths (5%). The dominant plant phyla in freshwater are chlorophytes (31% of species), charophytes (31%), tracheophytes (26%) and rhodophytes (10%). Thus, the distribution of richness among clades is strikingly different between the ocean and land, but the clade composition of freshwater is intermediate between these habitats. Importantly, the high phylogenetic diversity of freshwater per unit area (Figure 2b) is explained by the presence of a moderate number of diverse clades in a small area, as opposed to many diverse clades over a large area (marine animals) or very few diverse clades in an intermediate area (terrestrial).

Next, we used phylogenetic methods to examine the drivers of richness patterns among habitats (Figure 4). We estimated mean diversification rates of habitats by averaging rates across all families and phyla that occurred predominantly in each habitat. Results here are based on one animal tree (Tree II) and one relative extinction fraction ( $\epsilon = 0.5$ ), but alternatives yielded similar results (Table S5). Overall, freshwater animals and plants diversified faster than marine ones (Figure 4), but slower than terrestrial ones, paralleling richness patterns among habitats (animal phyla: freshwater = 0.013 species/million-years, marine = 0.009, terrestrial = 0.016; plant phyla: freshwater = 0.007, marine = 0.005, terrestrial = 0.017; family-level results were similar: Table S5). We also found strong effects of habitat on diversification rates among phyla for both plants and animals, but not families (Tables S6–S23, Figs S1–S2). Habitat did not significantly predict ages of these taxa (Table S24). These results are explained in detail in Appendix S3.

The richness of habitats might also be explained by their timing of colonisation, with habitats colonised earlier potentially having higher richness. We used

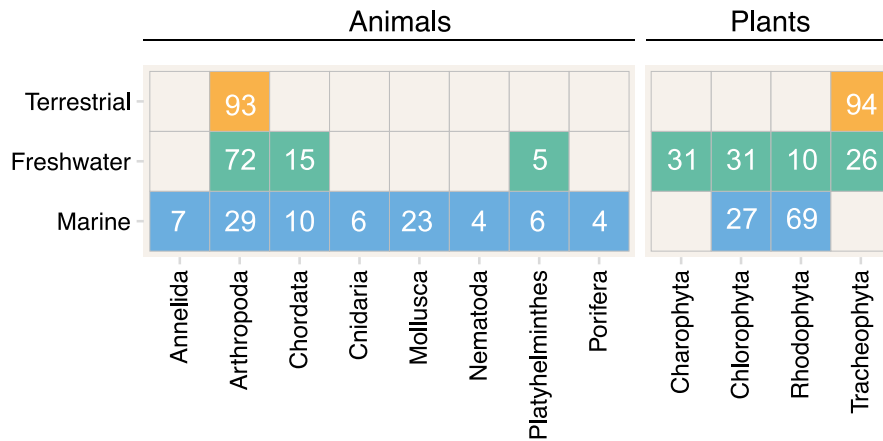


**FIGURE 2** Diversity per unit area in marine, freshwater and terrestrial habitats for animals and plants. For each habitat, we indicate the values of species richness (a) and phylogenetic diversity (b) divided by the area of these habitats. Phylogenetic diversity within habitats is measured as the sum of branch lengths (in millions of years) among sampled species within that habitat. Note that scales differ among panels

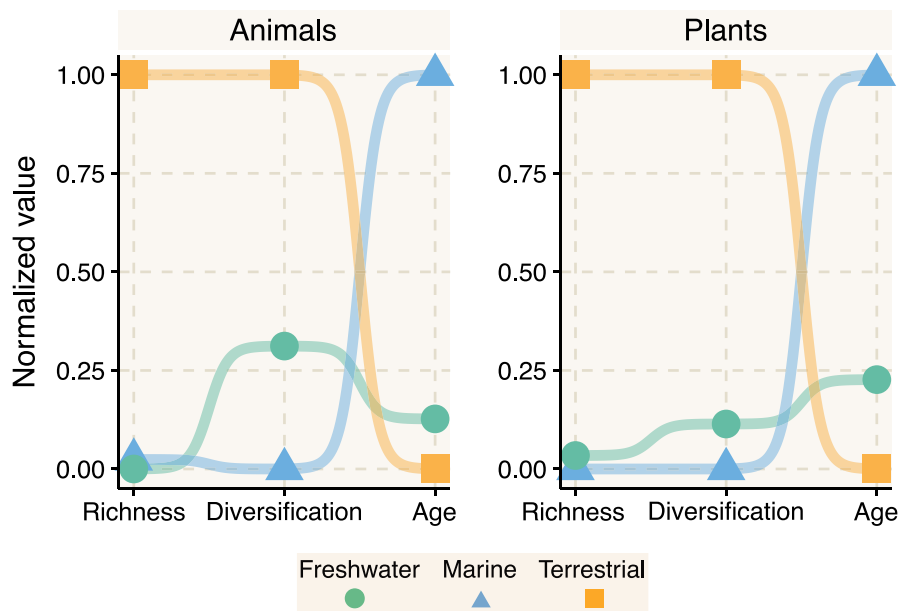
ancestral reconstructions to infer the timing of colonisation of each habitat (Tables S25–S26), using the proportionally sampled phylogenies (Tree II for animals). These analyses inferred that both animals and plants arose first in marine habitats (animals = 1030.6 Million years ago [Ma]; plants = 1550.6 Ma; Table S27). This was followed by colonisation of freshwater (animals = 651.7 Ma; plants = 937.3), with subsequent colonisation of land (animals = 591.5 Ma; plants = 701.1).

We also reviewed the oldest occurrence of each group in each habitat based on paleontological evidence (Appendix S2). This review also showed that animals and

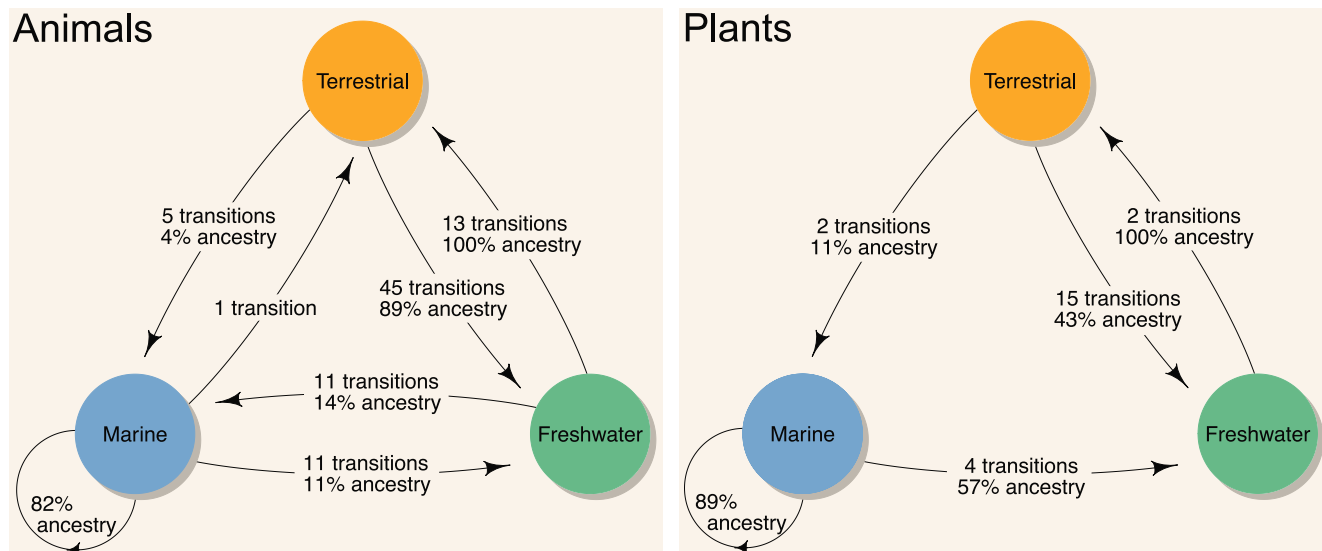
plants occurred first in the ocean, then in freshwater, and then on land, but with younger colonisation times inferred for each habitat than estimates based on molecular phylogenies (Table S28). This overall pattern is broadly concordant with other paleontological reviews (Becker & Marin, 2009; Benton, 2010; Labandeira, 2005; Miller & Labandeira, 2002). Importantly, older colonisation times cannot explain the high richness of terrestrial habitats relative to marine and freshwater. In summary, our results suggest that richness patterns among habitats are explained by variation in diversification rates and not colonisation times (Figure 4).



**FIGURE 3** The predominant animal and plant phyla in terrestrial, freshwater and marine habitats. Numbers in boxes are the percentage of extant, described species in each habitat belonging to that phylum. For each group and habitat, we indicate the most species-rich phyla in that habitat, including those that together comprise >90% of the species richness in each habitat



**FIGURE 4** Differences among habitats in species richness, diversification rates and the age of the oldest colonisation of each habitat. We present these values for species richness (Table S1), diversification rates (Table S5) and times of first colonisation (age; Table S27) on a scale from 0 to 1 (to facilitate comparison). Habitats with the highest species richness, fastest rates of diversification and oldest colonisation age have a standardised value of 1. Those with the lowest richness, slowest diversification and youngest colonisation age have a value of 0. Intermediate habitats have values normalised relative to the habitat with the lowest value for that variable. Diversification rates are based on phylum-level estimates using ( $\epsilon = 0.5$ ) and Tree II for animals. Alternative values are given in Table S5



**FIGURE 5** Transitions among habitats and sources of species richness in each habitat for animals and plants. We show the estimated numbers of transition events between each pair of habitats (but note that our focus is on the relative number of transitions, since the absolute numbers will be underestimated). We also show the percentage of extant richness within each habitat that is descended from colonisation from another habitat (“% ancestry”). For example, 57% of the extant plant richness in freshwater is descended from four transitions from marine to freshwater habitats. For each tip in each tree (with tips equivalent to species in the proportionally sampled trees), we determined the habitat transition that it descended from most recently, based on the most likely habitat state prior to its current habitat (e.g. for a terrestrial species, we determined whether it was descended most recently from freshwater or marine ancestors). Many marine species were descended from the inferred marine ancestor of plants and of animals, with no intervening habitat transitions. We estimated transitions among habitats based on maximum likelihood ancestral-state reconstructions using proportionally sampled phylogenies (using Tree II for animals)

We then used the proportionally sampled trees to estimate the relative numbers of transitions among habitats (Figure 5; Table S29; instantaneous transition rates in Table S30). These numbers generally paralleled the relative timing of colonisation of each habitat: most transitions from the ocean to other habitats were to freshwater (animals = 92% of 12 transitions [Tree II]; plants = 100% of 4) and most transitions from freshwater were to land (animals = 54% of 24; plants = 100% of 2). However, the largest numbers of transitions overall were from land back to freshwater (animals = 52% of all 86; plants = 65% of all 23). Our trees are not comprehensive at the species level and so underestimate absolute numbers of transitions. Nevertheless, they can potentially reveal patterns in relative numbers. These analyses also show that the relative numbers of colonisation events do not explain richness patterns (i.e. the most species-rich habitat has been colonised relatively infrequently).

Finally, we quantified the sources of richness in each habitat (Figure 5). First, we calculated the number of tips descended from the most recent colonisation of each habitat (Table S31), using the proportionally sampled trees. For both plants and animals, we estimated that >82% of currently marine species are descended from the initial origin of the group in marine habitats, not colonisations from other habitats. Second, we inferred that nearly all extant richness in terrestrial habitats is descended (most recently) from freshwater ancestors (animals: ~100%; plants: ~100%). Thus,

tracing the ancestral habitat of each sampled terrestrial species back to its most recent non-terrestrial ancestor that ancestor was almost always freshwater. Third, in animals, most extant richness in freshwater is derived from terrestrial habitat (89%). In plants, freshwater richness is derived in similar proportions from terrestrial (43%) and marine habitats (57%). Overall, these latter analyses provide possibly the first quantitative estimates of the sources of extant species diversity in each habitat for both animals and plants.

We also explored the impacts of using MuHiSSE and alternative coding of semi-aquatic species. The main results for animals were generally similar using MuHiSSE, but results for plants were very different (Appendix S4; Tables S32–S38). Most importantly, diversification rates from MuHiSSE in animals showed the same relative rankings among habitats (terrestrial>freshwater>marine) as those from the MS estimators. However, some MuHiSSE analyses were compromised by the relative rarity of freshwater and marine species, especially in plants (Appendix S4). Repeating the main analyses after treating semi-aquatic species as fully terrestrial also gave similar results for animals, and nearly identical results for plants (Appendix S5; Tables S39–S40).

## DISCUSSION

In this study, we address the origins of richness patterns among marine, freshwater and terrestrial habitats

in plants and animals. We show that most richness is in terrestrial environments (Figure 1), despite the smaller area these habitats occupy, and that freshwater habitats have exceptional richness per area (in animals) and exceptional phylogenetic diversity per area (in animals and plants; Figure 2). We find that differences in richness among habitats are related to differences in diversification rates (Figure 4). We also find that the overall sequence of colonisation among habitats is from marine to freshwater to terrestrial, but that most freshwater animal richness is derived from terrestrial ancestors (Figure 5). Below, we discuss how our findings relate to other studies, the processes that may underlie these patterns, and areas for future research.

Our study may be the first to simultaneously address richness patterns among all three habitats and their potential causes, in both plants and animals. Nevertheless, some of our results do have precedents in earlier studies. A previous study (Wiens, 2015a) found faster diversification rates in animal phyla with higher proportions of terrestrial species, but did not address freshwater richness or plants. Other studies (e.g. Grosberg et al., 2012; May, 1994) discussed richness patterns among habitats (similar to those found here), but did not test the causes of these patterns. Several studies have compared diversification rates between marine and freshwater taxa. For example, some studies in fish have found similar rates between these habitats (Carrete Vega & Wiens, 2012; Miller et al., 2018), whereas others found faster rates in marine species (Betancur-R et al., 2015) or freshwater species (Bloom et al., 2013; Tedesco et al., 2017). The latter is consistent with our overall results, as is a study in diatoms (Nakov et al., 2019).

Several studies have discussed the timing and number of colonisation of habitats based on the fossil record. Labandeira (2005) and Benton (2010) reviewed major colonisation events of terrestrial habitats by plants and animals. However, these studies were not comprehensive (i.e. including all three habitats for both plants and animals) and did not focus on explaining richness patterns. Vermeij and Dudley (2000) discussed transitions among habitats in plants and animals, but did not provide overall quantitative estimates of these transitions. Nevertheless, they concluded that transitions from land to sea were more common than from sea to land. This is supported by our results (Figure 5). Intriguingly, we found that the most frequent transitions among habitats were from land to freshwater. These transitions make a crucial contribution to freshwater richness: we estimate that ~89% of freshwater animal species are descended from terrestrial ancestors. Furthermore, the majority of freshwater animal species are actually semi-aquatic (66.7%; 109,267/163,795; Appendix S5). There has also been discussion about whether plants are ancestrally marine or freshwater based on ancestral-state reconstructions on molecular phylogenies (Nakov et al., 2017; Sánchez-Baracaldo, Bianchini, et al., 2017; Sánchez-Baracaldo, Raven, et al., 2017): our results favour a marine origin.

Our results suggest that the relative richness of habitats is explained by differences in diversification rates. Thus, a crucial question is: what explains these differences in diversification rates? One potential explanation is that high terrestrial richness is primarily explained by just two exceptional clades and their interactions: arthropods (i.e. insects) and angiosperms (Vermeij & Grosberg, 2010). These two clades each include >90% of plant and animal species in terrestrial habitats (Figure 3). Furthermore, recent large-scale analyses confirm that plant–insect interactions (e.g. herbivory, pollination) do contribute to increased diversification rates in each clade (Hernández-Hernández & Wiens, 2020; Wiens et al., 2015). However, plant–insect interactions may not be the sole explanation for high terrestrial diversification rates and richness. First, we obtained similar effects of habitat on diversification rates after excluding arthropods and tracheophytes (Tables S41–S48). Second, terrestrial habitat use is strongly related to increased diversification rates among vertebrate clades (Wiens, 2015b). Similarly, non-marine habitats (freshwater+terrestrial) strongly increase diversification rates among molluscs (Wiens, 2015a). Molluscs and chordates are the second and third largest animal phyla (after arthropods; Wiens, 2015a). A compelling explanation for variation in diversification rates among habitats should apply to all these groups.

One of the strongest hypotheses to explain higher diversification rates in non-marine habitats is that freshwater and terrestrial habitats have more effective barriers to dispersal than marine habitats (e.g. May, 1994; Vermeij & Grosberg, 2010; Wiens, 2015a). These barriers can potentially increase speciation, endemism and geographic turnover. These barriers include oceans, rivers and mountains for terrestrial species and any terrestrial habitat for freshwater species. By contrast, marine habitats are generally thought to have fewer barriers to dispersal. Barriers are crucial because allopatric speciation appears to be the predominant geographic mode in both animals and plants and in all three habitats (Hernández-Hernández et al., 2021). There is also support for this hypothesis (relating higher diversification rates to more limited dispersal) from large-scale analyses in fishes (Tedesco et al., 2017). Other explanations seem less likely. For example, net primary productivity is broadly similar in marine and terrestrial environments (Field et al., 1998), strongly suggesting that it does not explain higher richness or diversification in terrestrial environments. Likewise, physical differences between land and water (Vermeij & Grosberg, 2010) cannot explain the higher richness or diversification rates in freshwater relative to marine habitats.

In addition to speciation, extinction may also contribute to these patterns. In vertebrates, many extant marine taxa may be descendants of recolonisations from land (for amniotes; e.g. Kelley & Pyenson, 2015; Miller & Wiens, 2017; Pyenson et al., 2014) and from freshwater (for ray-finned fishes; Carrete Vega & Wiens, 2012), following



extinctions of other members of these groups in the ocean. Extinction may contribute to the lower diversification rates of clades in marine environments for higher level taxa (e.g. phyla) even if these differences in rates are less apparent among extant, lower level taxa (Miller & Wiens, 2017). Extinction might also contribute to lower diversification rates in freshwater than on land (Wiens, 2015a).

We acknowledge that some readers might dismiss our conclusions because the trees used are not comprehensive and because our study is not based primarily on fossils. Although our trees do not directly include every plant and animal species as tips, our sampling of higher-level clades encompasses most of these species. Our ancestral-state reconstructions are based on proportionally sampled trees. These latter results should generally reflect patterns that would be obtained with a complete species-level tree for all plant and animal species. An important exception is the absolute number of transitions among habitats (which are underestimated), but this need not overturn our conclusions about the *relative* frequency of these transitions and their impact on diversity patterns. Finally, our results are not based primarily on fossil evidence, but our overall goal is to understand present-day richness patterns among habitats, not those at each timepoint in the past. Furthermore, many large-scale analyses we perform here would be difficult with paleontological data. Nevertheless, paleontological analyses can provide complementary insights into these patterns, as we show here (Appendix S2).

In summary, we have shown the patterns of richness and phylogenetic diversity among habitats in plants and animals, and the likely causes of these patterns. Our results demonstrate the remarkable species richness and phylogenetic diversity of freshwater habitats per unit area, and highlight the need for global conservation efforts focused on these habitats (e.g. Reid et al., 2019). These results also have important implications for broader questions about what factors can be truly general explanations for global biodiversity patterns. For example, productivity and species–area relationships fail to explain these richness patterns among habitats. Time-integrated area (e.g. Fine, 2015) also seems unlikely to explain them, since the relative areas of these habitats appear relatively stable for the relevant time interval (e.g. Miller et al., 2005; Stephens et al., 2020). Instead, these results highlight the importance of diversification rates for driving large-scale biodiversity patterns. Therefore, a crucial next step for understanding these patterns will be to identify the specific factors underlying variation in diversification among habitats. We also reveal the origins of richness in each habitat, showing that most extant marine species are descended from an initial origin in these habitats, terrestrial richness is derived from ancient colonisation from freshwater and freshwater richness is largely dominated by more recent recolonisation from land.

## ACKNOWLEDGEMENTS

We thank J. Beaulieu for extensive help with SSE models, and M. Benton, R. Gillespie, M. S. Barker, M. J. Sanderson and anonymous reviewers for comments on the manuscript. DML thanks Maria Fernanda Perez for support. Analyses were conducted using High-Performance Computing (HPC) resources supported by the University of Arizona TRIF, UITS and RDI. J.J.W. acknowledges support from U.S. National Science Foundation grant DEB 1655690.

## AUTHORSHIP

C.R.P. and J.J.W. designed the study and drafted the manuscript. C.R.P. and D.M.L. collected data. C.R.P. performed analyses. All authors contributed to revisions.

## PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/ele.13999>.

## DATA AVAILABILITY STATEMENT

All data are included with this submission as Supporting Information and Data S1–S20, and are available on Dryad (<https://doi.org/10.5061/dryad.stjq2c50>).

## ORCID

Cristian Román-Palacios  <https://orcid.org/0000-0003-1696-4886>

John J. Wiens  <https://orcid.org/0000-0003-4243-1127>

## REFERENCES

- Adl, S.M., Simpson, A.G.B., Farmer, M.A., Andersen, R.A., Anderson, O.R., Barta, J.R. et al. (2005) The new higher level classification of Eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology*, 52, 399–451.
- Akaike, H.A. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.
- Balian, E.V., Segers, H., Martens, K. & Lévêque, C. (2008). The freshwater animal diversity assessment: an overview of the results. In: Balian, E.V., Lévêque, C., Segers, H. & Martens, K. (Eds.) *Freshwater animal diversity assessment*. Dordrecht: Springer, pp. 627–637.
- Beaulieu, J.M. & O'Meara, B.C. (2016) Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology*, 65, 583–601.
- Becker, B. & Marin, B. (2009) Streptophyte algae and the origin of embryophytes. *Annals of Botany*, 103, 999–1004.
- Benton, M.J. (2010) The origins of modern biodiversity on land. *Philosophical Transactions of the Royal Society of London*, 365, 3667–3679.
- Betancur-R, R., Orti, G. & Pyron, R.A. (2015) Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecology Letters*, 18, 441–450.
- Bloom, D.D., Weir, J.T., Piller, K.R. & Lovejoy, N.R. (2013) Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of New World silversides (Atherinopsidae). *Evolution*, 67, 2040–2057.
- Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multi-model inference: a practical information-theoretic approach*, 2nd edition. New York: Springer-Verlag.
- Caetano, S., O'Meara, B.C. & Beaulieu, J.M. (2018) Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution*, 72, 2308–2324.

- Carrete Vega, G. & Wiens, J.J. (2012) Why are there so few fish in the sea? *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 279, 2323–2329.
- Chamberlain, S.A. (2019) worrms: World Register of Marine Species (WoRMS) Client. R package version 0.4.0.9100.
- Chamberlain, S.A. & Szöcs, E. (2013) taxize: taxonomic search and retrieval in R. *F1000Research*, 2, 191–<https://doi.org/10.12688/f1000research.2-191.v1>
- Chambers, P.A., Lacoul, P., Murphy, K.J. & Thomaz, S.M. (2008) Global diversity of aquatic macrophytes in freshwater. *Hydrobiologia*, 595, 9–26.
- Davis, M.P., Midford, P.E. & Maddison, W. (2013) Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology*, 13, 38.
- Dawes, C.J. (1998) *Marine botany*, 2nd edition. New York, NY: John Wiley and Sons.
- Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1–10.
- 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.
- Fine, P.V.A. (2015) Ecological and evolutionary drivers of geographic variation in species diversity. *Annual Review of Ecology, Evolution, and Systematics*, 46, 369–392.
- Gleick, P.H. (1996) Water resources. In: Schneider, S.H. (Ed.) *Encyclopedia of Climate and Weather*. Oxford University Press, pp. 817–823.
- Grosberg, R.K., Vermeij, G.J. & Wainwright, P.C. (2012) Biodiversity in water and on land. *Current Biology*, 22, R900–R903.
- Guiry, M.D., Guiry, G.M., Morrison, L., Rindi, F., Miranda, S.V., Mathieson, A.C. et al. (2014) AlgaeBase: an on-line resource for algae. *Cryptogamie, Algologie*, 35, 105–115.
- Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. (2008) GEIGER: investigating evolutionary radiations. *Bioinformatics*, 24, 129–131.
- Hernández-Hernández, T., Miller, E.C., Román-Palacios, C. & Wiens, J.J. (2021) Speciation across the Tree of Life. *Biological Reviews*, 96, 1205–1242.
- Hernández-Hernández, T. & Wiens, J.J. (2020) Why are there so many flowering plants? A multiscale analysis of plant diversification. *American Naturalist*, 195, 948–963.
- Kelley, N.P. & Pyenson, N.D. (2015) Evolutionary innovation and ecology in marine tetrapods from the Triassic to the Anthropocene. *Science*, 348, 301.
- Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D. et al. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26, 1463–1464.
- Kozak, K.H. & Wiens, J.J. (2016) Testing the relationships between diversification, species richness, and trait evolution. *Systematic Biology*, 65, 975–988.
- Labandeira, C.C. (2005) Invasion of the continents: cyanobacterial crusts to tree-inhabiting arthropods. *Trends in Ecology & Evolution*, 20, 253–262.
- MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton, NJ: Princeton University Press.
- Magallón, S. & Sanderson, M.J. (2001) Absolute diversification rates in angiosperm clades. *Evolution*, 55, 1762–1780.
- Martin-Closas, C. (2003) The fossil record and evolution of freshwater plants: a review. *Geologica Acta*, 1, 315–338.
- May, R.C. (1994) Biological diversity: differences between land and sea. *Philosophical Transactions of the Royal Society of London*, 343, 105–111.
- Meyer, A.L.S., Román-Palacios, C. & Wiens, J.J. (2018) BAMM gives misleading rate estimates in simulated and empirical datasets. *Evolution*, 72, 2257–2266.
- Meyer, A.L.S. & Wiens, J.J. (2018) Estimating diversification rates for higher taxa: BAMM can give problematic estimates of rates and rate shifts. *Evolution*, 72, 39–53.
- Miller, E.C., Hayashi, K.T., Song, D. & Wiens, J.J. (2018) Explaining the ocean's richest biodiversity hotspot and global patterns of fish diversity. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 285, 20181314.
- Miller, E.C. & Wiens, J.J. (2017) Extinction and time help drive the marine-terrestrial biodiversity gradient: is the ocean a death-trap? *Ecology Letters*, 20, 911–921.
- Miller, K.G., Kominz, M.A., Browning, J.V., Wright, J.D., Mountain, G.S., Katz, M.E. et al. (2005) The Phanerozoic record of global sea-level change. *Science*, 310, 1293–1298.
- Miller, M.F. & Labandeira, C.C. (2002) Slow crawl across the salinity divide: delayed colonization of freshwater ecosystems by invertebrates. *GSA Today*, 12, 4–10.
- 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 Biology*, 9, e1001127.
- Nakov, T., Beaulieu, J.M. & Alvenson, A.J. (2019) Diatoms diversify and turn over faster in freshwater than marine environments. *Evolution*, 73, 2497–2511.
- Nakov, T., Boyko, J.D., Alvenson, A.J. & Beaulieu, J.M. (2017) Models with unequal transition rates favor marine origins of cyanobacteria and photosynthetic eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, 114, E10606–E10607.
- Paradis, E. & Schliep, K. (2018) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526–528.
- Pennell, M.W., Eastman, J.M., Slater, G.J., Brown, J.W., Uyeda, J.C., FitzJohn, R.G. et al. (2014) geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, 30, 2216–2218.
- Pyenson, N.D., Kelley, N.P. & Parham, J.F. (2014) Marine tetrapod macroevolution: physical and biological drivers on 250 Ma of invasions and evolution in ocean ecosystems. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 400, 1–8.
- Reid, A.J., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T.J. et al. (2019) Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, 94, 849–873.
- Ricklefs, R.E. (1987) Community diversity: relative roles of local and regional processes. *Science*, 235, 167–171.
- Roskov, Y., Kunze, T., Paglinawan, L., Orrell, T., Nicolson, D., Culham, A. et al. (2020) Species 2000 & ITIS Catalogue of Life, 2013 Annual Checklist. Available from <https://www.catalogueoflife.org>
- Sánchez-Baracaldo, P., Bianchini, G., Huelsenbeck, J.P., Raven, J.A., Pisani, D. & Knoll, A.H. (2017) Model choice requires biological insight when studying the ancestral habitat of photosynthetic eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, 114, E10608–E10609.
- Sánchez-Baracaldo, P., Raven, J.A., Pisani, D. & Knoll, A.H. (2017) Early photosynthetic eukaryotes inhabited low-salinity habitats. *Proceedings of the National Academy of Sciences of the United States of America*, 114, E7737–E7745.
- Scholl, J.P. & Wiens, J.J. (2016) Diversification rates and species richness across the Tree of Life. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 283, 20161335.
- Stephens, G.L., Slingo, J.M., Rignot, E., Reager, J.T., Hakuba, M.Z., Durack, P.J. et al. (2020) Earth's water reservoirs in a changing climate. *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences*, 476, 20190458.
- Tedesco, P.A., Paradis, E., Leveque, C. & Hugué, B. (2017) Explaining global-scale diversification patterns in actinopterygian fishes. *Journal of Biogeography*, 44, 773–783.
- Vermeij, G.J. (2017) How the land became the locus of major evolutionary innovations. *Current Biology*, 27, 3178–3182.
- Vermeij, G.J. & Dudley, R. (2000) Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems? *Biological Journal of the Linnean Society*, 70, 541–554.

- Vermeij, G.J. & Grosberg, R.K. (2010) The great divergence: when did diversity on land exceed that in the sea? *Integrative and Comparative Biology*, 50, 675–682.
- Wiens, J.J. (1998) The accuracy of methods for coding and sampling higher-level taxa for phylogenetic analysis: a simulation study. *Systematic Biology*, 47, 381–397.
- Wiens, J.J. (2015a) Faster diversification on land than sea helps explain global biodiversity patterns among habitats and animal phyla. *Ecology Letters*, 18, 1234–1241.
- Wiens, J.J. (2015b) Explaining large-scale patterns of vertebrate diversity. *Biology Letters*, 11, 20150506.
- Wiens, J.J., Lapoint, R.T. & Whiteman, N.K. (2015) Herbivory increases diversification across insect clades. *Nature Communications*, 6, 8370.
- WoRMS, (2020) World register of marine species. Available from <https://www.marinespecies.org>

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Román-Palacios, C., Moraga-López, D. & Wiens, J.J. (2022) The origins of global biodiversity on land, sea and freshwater. *Ecology Letters*, 25, 1376–1386. Available from: <https://doi.org/10.1111/ele.13999>