

Early Tetrapodomorph Biogeography: Controlling for Fossil Record Bias in Macroevolutionary Analyses

Jacob D. Gardner^{a,1}, Kevin Surya^{b,1}, and Chris L. Organ^{a*}

^a Department of Earth Sciences, Montana State University, Bozeman, MT 59717, USA

^b Honors College, Montana State University, Bozeman, MT 59717, USA

* Corresponding author at: Department of Earth Sciences, Montana State University, Bozeman, MT 59717, USA. E-mail address: organ@montana.edu (C. L. Organ)

¹ Contributed equally to this work

ABSTRACT

The fossil record provides direct empirical data for understanding macroevolutionary patterns and processes. Inherent biases in the fossil record are well known to confound analyses of this data. Sampling bias proxies have been used as covariates in regression models to test for such biases. Proxies, such as formation count, are associated with paleobiodiversity, but are insufficient for explaining species dispersal owing to a lack of geographic context. Here, we develop a sampling bias proxy that incorporates geographic information and test it with a case study on early tetrapodomorph biogeography. We use recently-developed Bayesian phylogeographic models and a new supertree of early tetrapodomorphs to estimate dispersal rates and ancestral habitat locations. We find strong evidence that geographic sampling bias explains supposed radiations in dispersal rate (potential adaptive radiations). Our study highlights the necessity of accounting for geographic sampling bias in macroevolutionary and phylogenetic analyses and provides an approach to test for its effect.

Keywords: sampling bias, fossil record, biogeography, phylogenetics, macroevolution, tetrapod water-land transition

1. Introduction

Our understanding of macroevolutionary patterns and processes are fundamentally based on fossils. The most direct evidence for taxonomic origination and extinction rates come from the rock record, as do evidence for novelty and climate change unseen in data sets gleaned from extant sources. There are no perfect data sets in science; there are inherent limitations and biases in the rock record that must be addressed when we form and test paleobiological hypotheses. For instance, observed stratigraphic ranges of fossils can mislead inferences about diversification and extinction rates (Raup and Boyajian, 1988; Signor and Lipps, 1982). Observed species diversity is also known to increase with time due to the preferential preservation and recovery of fossils in younger geological strata—referred to as "the Pull of the Recent" (Jablonski et al., 2003). Large and long-surviving clades with high rates of early diversification tend to result in an illusionary rate slow-down as diversification rates revert back to a mean value—referred to as “the Push of the Past” (Budd and Mann, 2018). Paleobiologists test and account for these biases when analyzing diversification and extinction at local and global scales (Alroy et al., 2001; Benson et al., 2010; Benson and Butler, 2011; Benson and Upchurch, 2013; Benton et al., 2013; Foote, 2003; Jablonski et al., 2003; Koch, 1978; Lloyd, 2012; Sakamoto et al., 2016a, 2016b). These bias-detection and correction techniques include fossil occurrence subsampling (Alroy et al., 2001; Jablonski et al., 2003; Lloyd, 2012); correcting origination, extinction, and sampling rates using evolutionary predictive models (Foote, 2003); the use of residuals from diversity-sampling models (Benson et al., 2010; Benson and Upchurch, 2013; Sakamoto et al., 2016b); and the incorporation of sampling bias proxies as covariates in regression models (Benson et al., 2010; Benson and Butler, 2011; Benton et al., 2013; Sakamoto et al., 2016a). Benton et al. (2013), studying sampling bias proxies, demonstrated that diversity through time closely tracks formation count (Benton et al., 2013).

However, case studies in England and Wales suggest that proxies for terrestrial sedimentary rock volume (such as formation count) do not accurately explain paleobiodiversity, particularly if the fossil record is patchy (Dunhill et al., 2014a, 2014b, 2013). Marine outcrop area and paleoecological-associated facies changes are, however, associated with shifts in paleobiodiversity (Dunhill et al., 2014b, 2013). Moreover, Benton et al. (2013) argue that the direction of causality between paleobiodiversity and formation count is unclear; there may be a common cause to explain their covariation, such as sea level (Benton et al., 2013). Nonetheless, formation count is a widely-used sampling bias proxy in phylogenetic analyses of macroevolution (O'Donovan et al., 2018; Sakamoto et al., 2016a, 2016b; Tennant et al., 2016a, 2016b). The advent of computational modeling approaches, particularly phylogenetic comparative methods, has made it easier to include proxies, like formation count, into models. Additional sampling bias proxies used in these studies include occurrence count, valid taxon count, and specimen completeness and preservation scores. Absent from these proxies is geographic context, which could confound many types of macroevolutionary analyses.

Despite advancements made in understanding the origin and evolution of early tetrapodomorphs, biogeographical studies are hindered by the incompleteness of the early tetrapodomorph fossil record. For example, “Romer’s Gap” represents a lack of tetrapodomorph fossils from the end-Devonian to mid-Mississippian, a period crucial for understanding early tetrapodomorph diversification. Recent collection efforts recovered tetrapodomorph specimens from “Romer’s Gap”, suggesting that a collection and preservation bias explains this gap (Clack et al., 2017; Marshall et al., 2019). In addition, a trackway site in Poland demonstrates the existence of digit-bearing tetrapodomorphs 10 million years before the earliest elpistostegalian body fossil, showcasing the limitation of body fossils to reveal evolutionary history (Niedźwiedzki et al., 2010).

A recent study by Long et al. (2018) leveraged phylogenetic reconstruction of early tetrapodomorphs to frame hypotheses about the origin of major clades, as well as their dispersal patterns, including the hypothesis that stem-tetrapodomorphs dispersed from Eastern Gondwana to Euramerica. However, this study did not use phylogenetic comparative methods to estimate ancestral geographic locations or to model dispersal patterns.

Here, we present a phylogeographic analysis of early tetrapodomorphs. Our goals are: 1) to construct a phylogenetic supertree of early tetrapodomorphs that synthesizes previous phylogenetic reconstructions; 2) to estimate the paleogeographic locations of major early tetrapodomorph clades using recently-developed phylogeographic models that account for the curvature of the Earth; and 3) to test for the influence of geographic sampling bias on dispersal rates. Our results indicate that geographic sampling bias substantially confounds analyses of dispersal and paleogeography. We conclude with a discussion about the necessity of controlling for fossil record biases in macroevolutionary analyses.

2. Materials and Methods

2.1. Nomenclature

Tetrapoda has been informally defined historically to include all terrestrial vertebrates with limbs and digits (Laurin, 1998). Gauthier et al. (1989) first articulated a phylogenetic definition of Tetrapoda as the clade including the last common ancestor of amniotes and lissamphibians. This definition excludes stem-tetrapodomorphs, like *Acanthostega* and *Ichthyostega*. Stegocephalia was coined by E.D. Cope in 1868 (Cope, 1868), but was more recently used to describe fossil taxa more closely related to tetrapods than other sarcopterygians. A recent cladistic redefinition of Stegocephalia includes all vertebrates more closely related to temnospondyls than *Panderichthys* (Laurin, 1998). Here, we use the definitions of Laurin (1998) for a monophyletic Stegocephalia

and of Gauthier et al. (1989) for Tetrapoda, which refers specifically to the crown group. We use Tetrapodomorpha to refer to all taxa closer to the tetrapod crown-group than the lungfish crown-group (Ahlberg, 1998). We additionally use Elpistostegalia (= Panderichthyida) to refer to the common ancestor of all stegocephalians and *Panderichthys* as well as Eotetrapodiformes to refer to the common ancestor of all tristichopterids, elpistostegalians, and tetrapods (Coates and Friedman, 2010).

2.2. Supertree

We inferred a supertree of 69 early tetrapodomorph taxa from five edited, published morphological data matrices, focusing on tetrapodomorphs whose previously inferred phylogenetic position bracket the water-land transition (Clack et al., 2017; Friedman et al., 2007; Pardo et al., 2017; Swartz, 2012; Zhu et al., 2017). Since downstream analyses might be sensitive to unequal sample sizes between taxa pre- and post-water-land transition, we did not include several crownward stem-tetrapodomorphs from the original matrices (see Supplementary Material). For each matrix, we generated a posterior distribution of phylogenetic trees using MrBayes 3.2.6 (Ronquist et al., 2012b). In each case, we ran two Markov chain Monte Carlo (MCMC) replicates for 20,000,000 generations with 25% burn-in, each with four chains and a sampling frequency of 1,000. We used one partition, except for Clack et al.'s (2017) matrix, which was explicitly divided into cranial and postcranial characters. To time-calibrate the trees, we constrained the root ages and employed a tip-dating approach (Ronquist et al., 2012a). Tip dates (last occurrence) were acquired from the Paleobiology Database (PBDB; <https://paleobiodb.org/>) and the literature (see Supplementary Table 2). Root calibrations (minimum and soft maximum age estimates) were collected from the PBDB and Benton et al. (2015). We also used the fossilized birth-death model as the branch length prior (Didier et al., 2017, 2012; Didier and Laurin, 2018;

Gavryushkina et al., 2014; Heath et al., 2014; Stadler, 2010; Zhang et al., 2016). All pairs of MCMC replicates converged as demonstrated by low average standard deviation of split frequencies (<0.005 ; Lakner et al., 2008; see Supplementary Table 3).

Next, we used the five maximum clade credibility trees (source trees; Supplementary Fig. 1-10) to compute a distance supermatrix using SDM 2.1 (Criscuolo et al., 2006). We then inferred an unweighted neighbor-joining tree (UNJ by Gascuel, 1997) from the distance supermatrix using PhyD* 1.1 (Criscuolo and Gascuel, 2008). The UNJ* algorithm is preferable for matrices based on morphological characters. Unlike most supertree methods, the SDM-PhyD* combination produces a supertree with branch lengths. We rooted the supertree using phytools 0.6.60 (Revell, 2012) by adding an arbitrary branch length of 0.00001 to break the trichotomy at the basal-most node in R 3.5.2 (R Core Team, 2018), designating the dipnomorph *Glyptolepis* as the outgroup.

We qualitatively compared the supertree topology with the published source trees and Marjanović and Laurin's (2019) Paleozoic limbed vertebrate topologies. We also calculated normalized Robinson-Foulds (nRF) distances (Robinson and Foulds, 1981) using phangorn 2.4.0 (Schliep, 2011) in R to assess the congruency of topologies. In each comparison, polytomies in the supertree or the source tree were resolved in all possible ways using phytools. We then calculated all nRF distances and took an average (see Supplementary Table 4). The supplementary materials include a more detailed description of this approach.

2.3. Phylogeography

We obtained paleocoordinate data (paleolatitude and paleolongitude) for 63 early tetrapodomorphs from the PBDB using the GPlates software setting (<https://gws.gplates.org/>). By default, GPlates estimates paleocoordinates from the midpoint of each taxon's age range. For 16 taxa that did not have direct paleocoordinate data in the PBDB, we searched for the geological

formations and geographic regions within the time range from which they are known and averaged the paleolocations across each valid taxonomic occurrence in the PBDB. If the paleolocation of the formation was not listed in the PBDB, we used published geographic locations of the formations. This level of precision is adequate for world-wide phylogeographic analyses, such as conducted here. Present-day coordinates for these geographic locations were obtained from Google Earth and matched with PBDB entries that date within each taxon's age range (see Supplementary Table 5). Four additional taxa, *Kenichthys*, *Koilops*, *Ossirarus*, and *Tungsenia*, had occurrences in the PBDB but the GPlates software could not estimate their paleocoordinates. For *Koilops* and *Ossirarus*, we used all tetrapodomorph occurrences from the Ballagan Formation of Scotland, UK—a formation in which these two taxa are found (Clack et al., 2017). For *Kenichthys* and *Tungsenia*, we calculated paleocoordinate data from the GPlates website directly using the present-day coordinates from the PBDB (<https://gws.gplates.org/#recon-p>). This approach did not work for the 16 previously mentioned taxa (see Supplementary Table 5). We therefore obtained paleocoordinate data from nearby entries in the PBDB. We excluded the following taxa from our analyses due to the lack of data and comparable entries in the PBDB: *Jarvikina*, *Koharalepis*, *Spodichthys*, and *Tinirau*. We excluded the outgroup taxon, *Glyptolepis*, in our analysis to focus on the dispersal trends within early Tetrapodomorpha. We also excluded *Eusthenodon* and *Strepsodus* because their high estimated dispersal rates—being reported from multiple continents—masked other rate variation throughout the phylogeny and inhibited our downstream analyses from converging on a stable likelihood. We do, however, discuss their geographic implications in Section 4.

A model that incorporates phylogeny is crucial for paleobiogeographic reconstruction because it accounts for both species relationships and the amount of evolutionary divergence

(branch lengths). Using continuous paleocoordinate data, rather than discretely-coded regions, allows dispersal trends to be estimated at finer resolutions. Discretely-coded geographic regions also limit ancestral states to the same regions inhabited by descendant species. However, standard phylogenetic comparative methods for continuous data assume a flat Earth because they do not account for spherically structured coordinates (i.e., the proximity of -179° and 179° longitudes). Recently-developed phylogenetic comparative methods for modeling continuous paleocoordinate data, implemented as the ‘geo’ model in the program BayesTraits V3, overcome this hurdle by “evolving” continuous coordinate data on the surface of a globe (O’Donovan et al., 2018). The model is implemented with a Bayesian reversible jump MCMC algorithm to estimate rates of geographic dispersal and ancestral paleolocations simultaneously. To account for the spheroid shape of the globe, the ‘geo’ model converts latitude and longitude data into three-dimensional coordinates while prohibiting moves that penetrate the inside of the globe. Ancestral states, which are converted back to standard latitude and longitude, are estimated for each node of the phylogeny. The method includes a variable rates model to estimate variation in dispersal rate (Venditti et al., 2011). The ‘geo’ model makes no assumptions about the location of geographic barriers or coastlines, but a study on dinosaur biogeography found 99.2% of mean ancestral state reconstructions to be located within the bounds of landmasses specific to the time at which they occurred (O’Donovan et al., 2018). We ran three replicate independent analyses using the Bayesian phylogenetic ‘geo’ model for 100 million iterations each with a 25% burn-in and sampling every 1,000 iterations. We estimated log marginal likelihoods using the Stepping Stone algorithm with 250 stones sampling every 1,000 iterations (Xie et al., 2011). We used Bayes factors (BF) to test whether a variable rates model explained the data better than a uniform rates model. Bayes factors greater than two are considered good evidence in support of the model with the greater log

marginal likelihood. We compared estimated rate scalars and ancestral states among the three independent variable rates analyses to check for consistency in our results. Rates of dispersal were estimated for each branch by dividing the average rate scalars by the original branch lengths (scaled by time). We assessed the MCMC convergence of all analyses using Tracer 1.7 (Rambaut et al., 2018).

To test for the effect of sampling bias on dispersal rates, we developed a sampling bias proxy that incorporates geographic context: regional-level formation count. Formation counts are meant to capture multiple biases: uneven global rock exposure, uneven fossil collection and database efforts, and global variation in sediment deposition in environments conducive to preservation. Stage-level (stage-specific) formation count represents the mean number of formations, or distinct rock units, globally known to produce relevant fossils along each terminal branch of a phylogeny. Following the protocol of Sakamoto et al. (2016) and O'Donovan et al. (2018), stage-level formation counts are calculated by taking the average number of formations known from each geological age across the globe that encompass the time period between the taxon's tip date and its preceding node. These average stage-level formation counts are weighted by the proportion that each terminal branch length covers each geological age. For example, if a terminal branch covers two geological ages (e.g., Frasnian and Famennian) at 30% and 70%, respectively, then the stage-level formation counts from each geological age are weighted by those proportions and then divided by the number of geological ages covered:

$$\text{Stage – Level FormationCount} = \frac{\text{FrasnianCount} \times 0.3 + \text{FamennianCount} \times 0.7}{2}$$

Stage-level formation count is not informed by geography; it is a global metric. It is

therefore an inadequate proxy if bias has a strong geographic component (e.g., if the majority of formations recorded are from a specific region or if few formations are exposed within a region). The number of fossil-bearing geological formations, accounting for geographic distribution, is expected to be an important confounding bias in the fossil record. We developed a proxy that includes geographic sampling bias. Our approach breaks down stage-level formation count by geographic region. To account for the arrangement of the continents during the Devonian, Carboniferous, and Permian, we recognized five major regions: Northern Euramerica (including Northeastern Eurasia and Central Asia), Southern Euramerica (North America, Greenland, and Western Europe), Western Gondwana (South America and Africa), Eastern Gondwana (Antarctica, Australia, and Southern Asia), and East Asia (e.g., China). For each branch in the phylogeny, we used the average ancestral state and taxon paleolocation estimates to determine if the branch crossed multiple geographic regions. The number of formations within this time window are totaled for every region covered by the branch and then divided by the number of regions covered. For example, if ancestral state estimates at node 1 and 2 are located in Eastern Gondwana and Southern Euramerica, respectively, then the number of formations recorded in Eastern Gondwana, Southern Euramerica, and the regions in between (i.e., Western Gondwana or Northern Euramerica + East Asia) are counted for that geological age; this total is then divided by the number of geographic regions covered by the entire branch (three for the Western Gondwana route and four for the Northern Euramerica + East Asia route). If the dispersal path between two consecutive ancestral states does not cross any of the five regions, then the number of formations in the inhabited region is counted alone. Figure 1 illustrates an example of how this proxy is measured. This results in the average number of formations present along the dispersal path (at geographic region scale) for each branch in the phylogeny. As with stage-level formation counts, the regional-

level formation counts are weighted by the proportion that the branch length covers each geological age. We hypothesize that dispersal rate will inversely correlate with regional-level formation count because we expect that the lack of formations in intermediate regions will lead to inflated dispersal rates. The ‘geo’ model will increase the dispersal rate along a branch to account for the geographic variation observed when there is a lack of intermediate geographic fossil occurrences. This hypothesis can be falsified if high dispersal rates are associated with larger average numbers of formations along dispersal paths. Benton et al. (2013) provide a global sample of tetrapod-bearing rock formations known for each geological age from the Middle Devonian through the Triassic. We supplemented these lists with stratigraphic units known to produce sarcopterygian fossils entered in the PBDB (collected on December 10th, 2018).

Period	Epoch	Age	End Time (Ma)	Northern Euramerica	Southern Euramerica	Western Gondwana	Eastern Gondwana	East Asia	Total
Permian	Cisuralian	Kungurian	272.95	10	44	11	0	0	65
Permian	Cisuralian	Artinskian	283.5	2	39	9	0	0	50
Permian	Cisuralian	Sakmarian	290.1	2	44	4	0	0	50
Permian	Cisuralian	Asselian	295	2	47	3	0	0	52
Pennsylvanian	Late	Gzhelian	298.9	1	42	0	0	0	43
Pennsylvanian	Late	Kasimovian	303.7	0	33	0	0	0	33
Pennsylvanian	Middle	Moscovian	307	0	16	0	0	0	16
Pennsylvanian	Early	Bashkirian	315.2	0	28	0	0	0	28
Mississippian	Late	Serpukhovian	323.2	0	16	0	0	0	16
Mississippian	Middle	Viséan	330.9	0	14	0	1	0	15
Mississippian	Early	Tournaisian	346.7	0	7	0	1	0	8
Devonian	Late	Famennian	358.9	1	9	1	5	1	17
Devonian	Late	Frasnian	372.2	1	11	0	3	2	17
Devonian	Middle	Givetian	382.7	1	8	1	4	1	15
Devonian	Middle	Eifelian	387.7	1	8	0	5	2	16
Devonian	Early	Emsian	393.3	2	5	0	7	3	17
Devonian	Early	Pragian	407.6	1	6	0	4	2	13
Devonian	Early	Lochkovian	410.8	1	3	0	3	1	8
Silurian	Prídolí	Prídolí	419.2	0	0	0	0	1	1
Silurian	Ludlow	Ludfordian	423	0	0	0	0	2	2
Silurian	Ludlow	Gorstian	425.6	0	0	0	0	2	2

Table 1: Regional- and stage-level (total) formation counts through time.

To test for the effect of regional-level formation count bias on dispersal rate, we conducted a non-parametric two-sample, upper-tailed Mann-Whitney *U*-test using the base package ‘stats’ in R (R Core Team, 2018). This approach ranks all branches of the phylogeny by their regional-level formation count and tests if the branches with lower dispersal rates rank higher on average than branches with higher rates. We define “high” vs “low” dispersal rates based on whether or not they are two standard deviations greater than the average rate across the tree. Due to the vast difference in sample size between the two groups (“high rates”: $n = 9$, “low rates”: $n = 111$), we bootstrapped the regional-level formation counts from each group with 100,000 replicates. From this bootstrap analysis, we obtained a 95% confidence interval for the summed ranks of the branches with low dispersal rates ($n = 100,000$ *U*-statistic values). The expected *U*-statistic is 499.5 given the null hypothesis that only 50% of the regional-level formation counts along branches with low rates rank higher than the formation counts with high rates (half of all possible combinations = $\frac{9 \times 111}{2}$). A 95% confidence interval of bootstrapped *U*-statistics that does not include the null expected *U*-statistic is considered good evidence for higher mean dispersal rates along branches with lower regional-level formation counts. The full dataset and code for the phylogeographic analyses can be requested by email to the corresponding author.

Estimated ancestral states do not identify specific dispersal routes, so we conducted sensitivity analyses to test if the dispersal route chosen for counting formations influenced our results. We conceived of three scenarios for dispersal routes between Eastern Gondwana and Southern Euramerica or vice versa: 1) a dispersal route through Western Gondwana; 2) a route through Northern Euramerica and East Asia; and 3) a direct route between Eastern Gondwana and Southern Euramerica. For the first scenario, we averaged the number of formations found in

Eastern and Western Gondwana and Southern Euramerica for a given time period. The second scenario is similar to the first but included formation counts from Northern Euramerica and East Asia in place of Western Gondwana. The third scenario only averaged formation counts from Eastern Gondwana and Southern Euramerica.

3. Results

3.1. Supertree

Topological differences resulted among our supertree, the published source trees, and Marjanović and Laurin's (2019) tree (Figure 2). In our tree, a polyphyletic “Megalichthyiformes” is the basal-most tetrapodomorph group instead of Rhizodontida (Swartz, 2012; Zhu et al., 2017). Canowindrids and rhizodontids formed an unexpected sister clade to Eotetrapodiformes. Clack et al.'s (2017) five Tournaisian tetrapod taxa cluster together. Colosteidae is rootward of *Crassigyrinus*. *Caerorhachis* is next to Baphetidae. Baphetidae moved crownward compared to previous topologies (likely because of a small character sample size [Marjanović and Laurin, 2019]). Two crownward nodes are unresolved (polytomous). We retained *Tungsenia* and *Kenichthys* as the oldest and second oldest tetrapodomorphs. Tristichopteridae, Elpistostegalia, Stegocephalia, Aïstopoda, Whatcheeriidae, Colosteidae, Anthracosauria, Dendrerpetidae, and Baphetidae remain monophyletic. Aïstopoda (*Lethiscus* and *Coloraderpeton*) fell rootward to Tetrapoda as reported in Pardo et al. (2017; 2018). The average nRF distances quantify differences in topology (see Supplementary Table 4). On average, there are 39.7% different or missing bipartitions in the source trees compared to the supertree.

3.2. Phylogeography

We found overwhelming support for a variable rates model of geographic dispersal in early tetrapodomorphs (BF = 632.3; Figure 3). The estimated rates across the three replicate runs are

consistent (out of 122 branches, only three had a median rate scalar with an absolute value difference among the three runs greater than 3). All rate shifts that were two standard deviations greater than the average dispersal rate were reconstructed dispersal events moving from East Asia to Southern Euramerica, from Eastern Gondwana to Southern Euramerica, or Southern Euramerica to Eastern Gondwana. The fastest estimated dispersal rate occurs along the branch leading to Eotetrapodiformes, moving from Eastern Gondwana to Southern Euramerica (14.34x the average rate). As Long et al. (2018) suggest, we find evidence for an East Asian origin for Tetrapodomorpha but with moderate uncertainty (average estimate \pm standard deviation of posterior distribution; $\text{longitude}_{\text{avg}} = 81.5^\circ \pm 10.1^\circ$, $\text{latitude}_{\text{avg}} = -6.4^\circ \pm 8.5^\circ$). We also reconstruct an origin for “Megalichthyiformes” that borderlines East Asia and Eastern Gondwana ($\text{longitude}_{\text{avg}} = 107.2^\circ \pm 14.1^\circ$, $\text{latitude}_{\text{avg}} = -22.6^\circ \pm 8.7^\circ$), along with an Eastern Gondwana origin for the clade uniting “Canowindridae” and Rhizodontida ($\text{longitude}_{\text{avg}} = 137.1^\circ \pm 8.2^\circ$, $\text{latitude}_{\text{avg}} = -32.0^\circ \pm 4.7^\circ$). We recover a Southern Euramerican origin for Eotetrapodiformes, consistent with previous studies ($\text{longitude}_{\text{avg}} = -12.5^\circ \pm 7.0^\circ$, $\text{latitude}_{\text{avg}} = -19.4^\circ \pm 6.4^\circ$). A Southern Euramerican origin was also found for Tristichopteridae ($\text{longitude}_{\text{avg}} = -12.7^\circ \pm 6.9^\circ$, $\text{latitude}_{\text{avg}} = -19.7^\circ \pm 6.3^\circ$) and Elpistostegalia ($\text{longitude}_{\text{avg}} = -12.3^\circ \pm 5.5^\circ$, $\text{latitude}_{\text{avg}} = -13.5^\circ \pm 5.3^\circ$). As expected in a phylogenetic comparative analysis, uncertainty in estimated node states increases toward the root. However, despite the level of uncertainty within a single run, only three nodes have mean ancestral state values that are greater than an absolute value of 5° among the replicate three runs.

We find good evidence that geographic sampling bias influences dispersal rate estimates, regardless of the route used (95% CI: Western Gondwana route $U = [800, 928]$; Northern Euramerica + East Asia route $U = [832, 946]$; direct route $U = [729, 889]$; no scenario includes the null $U = 499.5$; Figure 4 and Supplementary Figures 12-13). A U -statistic considerably higher than

499.5 suggests that branches with high dispersal rates have lower regional-level formation counts, on average, than branches with low rates. One can also interpret the null U -statistic of 499.5 as a 50% probability that a random branch with a low dispersal rate will rank higher in its regional-level formation count than a random branch with a high dispersal rate. With bootstrapping, we are 95% confident that the probability of a random branch with a low dispersal rate having a higher regional-level formation count than a random branch with a high rate is 72.97–88.99% for the more conservative ‘direct route’ scenario. Under the more liberal ‘Northern Euramerica + East Asia route’ scenario, the probabilities are 83.28–94.69%. In sum, branches with high dispersal rates (two standard deviations greater than average) have a smaller number of recorded formations, on average, along their reconstructed dispersal path.

Our results cannot be explained by a fossil record that is more complete through time (Pull of the Recent). A regression model relating regional-level formation count to the minimum age of each branch shows only a weak relationship (slope = -0.044, $r^2 = 0.1$, $P < 0.001$). However, total global (stage-level) formation count (which does not account for geographic variation) does show potential bias from Pull of the Recent (slope = -0.3, $r^2 = 0.71$, $P < 0.0001$). If dispersal rates are biased by the increase in number of formations globally, we would also expect to see elevated dispersal rates decrease toward the tips, but a regression model relating stretched branch lengths with time is not supported (slope = -0.025, $r^2 = 0.006$, $P = 0.41$).

4. Discussion

We expected to infer high dispersal rates for closely related taxa that are distributed across the globe. Our results, unadjusted for geographic bias in the fossil record, confirm this notion. However, we also find a compelling statistical association between high dispersal rates and a low number of formations along dispersal paths—a patchy fossil record is driving inferences of high

dispersal rates. Although we did not test for a correlation between dispersal rate and previously used proxies, such as valid taxon count and stage-level formation count, these proxies do not offer clear predictions for explaining dispersal rate variation. High dispersal rate variation is inferred when closely related taxa are geographically separate. For example, valid taxon count cannot explain geographic rate variation because spatial information is lacking in this bias proxy and because sister taxa are likely to have similar counts (these data are phylogenetically structured). Stage-level formation counts will also not explain dispersal rate variation, particularly if high rate variation exists within the same geological age. Assuming geological formations are evenly exposed and sampled worldwide, low stage-level formation counts should yield geographically variable fossil species and, therefore, drive high dispersal rate variation. However, formations are not evenly exposed or recorded in geological/paleontological databases, including the PBDB. Our formation count table demonstrates this bias (Table 1). Without geographic context, stage-level formation count cannot distinguish between global and local regions. For example, the geological ages that have the highest recorded number of formations are restricted to Southern Euramerica where the majority of eotetrapodiform taxa have been discovered. The association between high formation counts in specific regions and high paleobiodiversity in those regions is likely not a coincidence and has a clear impact on how we interpret dispersal history. The earliest tetrapodomorphs are known from China and Australia at geological ages where relatively few formations are recorded outside of East Asia and Eastern Gondwana. The basal-most ancestral state estimates reconstruct paleolocations in East Asia (not surprisingly). This inference (hypothesis) is predicated on the lack of geological formations recorded outside of East Asia during this time period. In addition, the majority of more crownward taxa and their reconstructed ancestral states are located in North America and Europe at geological ages in which relatively fewer

formations are known elsewhere. This bias may heavily influence any conclusions made on the location and habitat of the tetrapod water-land transition. Recently discovered taxa could help mitigate this problem by increasing the power of taxon sampling (Heath et al., 2014), such as *Tutusius* and *Umzantsia* from South Africa (Gess and Ahlberg, 2018). However, the current lack of cladistic coding for these taxa excludes them from phylogeny-based analyses. The taxonomic resolution of globally-occurring species, like *Eusthenodon* and *Spodichthys*, also impacts current models of species dispersal history because of their relatively uniform distribution (Long et al., 2018). *Eusthenodon* and *Spodichthys* represent possible cases where taxonomic resolution is too coarse for phylogeographic analyses. Including these species inhibited our MCMC algorithms from reaching convergence. Widely distributed cosmopolitan species that lack intermediate geographic occurrences increase the uncertainty of parameter estimates within phylogeographic models, as is the case here for these two species.

Phylogenetic studies on macroevolution also often fail to incorporate data from the fossil record itself, such as trace fossil occurrences. Non-anatomical data often contribute to our understanding of taxonomic originations, including chirodian (or digit-possessing) tetrapodomorphs for which trace fossil evidence exists about 10 million years before the first elpistostegalian body fossils (Niedźwiedzki et al., 2010). The inclusion of additional data from trace fossils could radically alter our current models of species dispersal history. Finally, it is important to note that the sampling bias proxies are also constrained by database curation biases. Phylogenetic studies on macroevolutionary trends now regularly leverage public databases, such as the PBDB, which allows larger and broader studies. It is unclear how patchy entries, on taxonomic occurrences and geological formations, for example, interact with other biases inherent in the fossil record. Caution is therefore warranted when these databases are mined, as is the case

here.

5. Conclusions

Phylogenetic studies on macroevolution have not previously incorporated geographic context, which could influence a wide variety of analyses. We demonstrate here that phylogeographic methods are influenced by geographic sampling variability. We develop a simple sampling bias proxy that incorporates geographic information and show that it explains variation in estimated dispersal rates. The majority of elevated dispersal rates are associated with large-scale movements between major landmasses that have very few, if any, relevant geological formations in between. Our analysis is also unlikely to be influenced by “Pull of the Recent”-like effects. Although not the first supertree for early tetrapodomorphs (Ruta et al., 2003), this study presents the first (to our knowledge) with branch lengths, making it useable for phylogenetic comparative analyses. The new supertree comprises many of the major clades previously inferred, but also recovers new ones that will be subject to scrutiny in future studies (discussed further in the Supplementary Material). This supertree should be useful to researchers who aim to use phylogenetic comparative methods to test hypotheses on the evolution of early tetrapodomorphs. In sum, our study estimates ancestral geographical reconstructions consistent with previously hypothesized dispersal patterns in early tetrapodomorphs. We also find that rates of dispersal are strongly influenced by geographic sampling bias. We suggest that researchers incorporate this proxy in phylogeny-based macroevolutionary studies that could be influenced by spatial distribution of the fossil record.

Acknowledgements

We thank the MSU Macroevolution Lab, Jack Wilson, and Matt Lavin for helpful discussions, as

well as David Marjanović and John Long for their helpful reviews. We also thank Nathalie Bardet, Eric Buffetaut, Annelise Folie, Emmanuel Gheerbrant, Alexandra Houssaye, and Michel Laurin for the invitation to contribute to this special issue celebrating the life and accomplishments of Jean-Claude Rage.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Ahlberg, P.E., 1998. Postcranial stem tetrapod remains from the Devonian of Scat Craig, Morayshire, Scotland. *Zool. J. Linn. Soc.* 122, 99–141. <https://doi.org/10.1006/zjls.1997.0115>
- Alroy, J., Marshall, C.R., Bambach, R.K., Bezusko, K., Foote, M., Fürsich, F.T., Hansen, T.A., Holland, S.M., Ivany, L.C., Jablonski, D., Jacobs, D.K., Jones, D.C., Kosnik, M.A., Lidgard, S., Low, S., Miller, A.I., Novack-Gottshall, P.M., Olszewski, T.D., Patzkowsky, M.E., Raup, D.M., Roy, K., Sepkoski, J.J., Sommers, M.G., Wagner, P.J., Webber, A., 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *PNAS* 98, 6261–6266. <https://doi.org/10.1073/pnas.111144698>
- Benson, R.B.J., Butler, R.J., 2011. Uncovering the diversification history of marine tetrapods: ecology influences the effect of geological sampling biases. Geological Society, London, Special Publications 358, 191–208. <https://doi.org/10.1144/SP358.13>
- Benson, R.B.J., Butler, R.J., Lindgren, J., Smith, A.S., 2010. Mesozoic marine tetrapod diversity:

mass extinctions and temporal heterogeneity in geological megabiases affecting vertebrates.

Proceedings of the Royal Society B: Biological Sciences 277, 829–834.

<https://doi.org/10.1098/rspb.2009.1845>

Benson, R.B.J., Upchurch, P., 2013. Diversity trends in the establishment of terrestrial vertebrate

ecosystems: Interactions between spatial and temporal sampling biases. *Geology* 41, 43–

46. <https://doi.org/10.1130/G33543.1>

Benton, M.J., Donoghue, P.C.J., Asher, R.J., Friedman, M., Near, T.J., Vinther, J., 2015.

Constraints on the timescale of animal evolutionary history. *Palaeontol. Electron.* 18, 1–

106. <https://doi.org/10.26879/424>

Benton, M.J., Ruta, M., Dunhill, A.M., Sakamoto, M., 2013. The first half of tetrapod evolution,

sampling proxies, and fossil record quality. *Palaeogeography, Palaeoclimatology,*

Palaeoecology 372, 18–41. <https://doi.org/10.1016/j.palaeo.2012.09.005>

Budd, G.E., Mann, R.P., 2018. History is written by the victors: The effect of the push of the past

on the fossil record. *Evolution* 72, 2276–2291. <https://doi.org/10.1111/evo.13593>

Clack, J.A., Bennett, C.E., Carpenter, D.K., Davies, S.J., Fraser, N.C., Kearsey, T.I., Marshall,

J.E.A., Millward, D., Otoo, B.K.A., Reeves, E.J., Ross, A.J., Ruta, M., Smithson, K.Z.,

Smithson, T.R., Walsh, S.A., 2017. Phylogenetic and environmental context of a

Tournaisian tetrapod fauna. *Nat. Ecol. Evol.* 1, 0002. <https://doi.org/10.1038/s41559-016->

0002

Coates, M.I., Friedman, M., 2010. *Litoptychus bryanti* and characteristics of stem tetrapod

neurocrania, in: Elliot, D.K., Maisey, J.G., Yu, X., Miao, D. (Eds.), *Morphology, Phylogeny*

and Paleobiogeography of Fossil Fishes. Verlag Dr. Friedrich Pfeil, München, Germany,

pp. 389–416.

- 495 Cope, E.D., 1868. Synopsis of the Extinct Batrachia of North America. Proceedings of the
496 Academy of Natural Sciences of Philadelphia. 20, 208–221.
- 497 Criscuolo, A., Berry, V., Douzery, E.J.P., Gascuel, O., 2006. SDM: A fast distance-based approach
498 for (super)tree building in phylogenomics. Syst. Biol. 55, 740–755.
499 <https://doi.org/10.1080/10635150600969872>
- 500 Criscuolo, A., Gascuel, O., 2008. Fast NJ-like algorithms to deal with incomplete distance matrices.
501 BMC Bioinformatics 9, 166. <https://doi.org/10.1186/1471-2105-9-166>
- 502 Didier, G., Fau, M., Laurin, M., 2017. Likelihood of tree topologies with fossils and diversification
503 rate estimation. Syst. Biol. 66, 964–987. <https://doi.org/10.1093/sysbio/syx045>
- 504 Didier, G., Laurin, M., 2018. Exact distribution of divergence times from fossil ages and topologies.
505 bioRxiv 490003. <https://doi.org/10.1101/490003>
- 506 Didier, G., Royer-Carenzi, M., Laurin, M., 2012. The reconstructed evolutionary process with the
507 fossil record. J. Theor. Biol. 315, 26–37. <https://doi.org/10.1016/j.jtbi.2012.08.046>
- 508 Dunhill, A.M., Benton, M.J., Newell, A.J., Twitchett, R.J., 2013. Completeness of the fossil record
509 and the validity of sampling proxies: a case study from the Triassic of England and Wales.
510 Journal of the Geological Society 170, 291–300. <https://doi.org/10.1144/jgs2012-025>
- 511 Dunhill, A.M., Benton, M.J., Twitchett, R.J., Newell, A.J., 2014a. Testing the fossil record:
512 Sampling proxies and scaling in the British Triassic–Jurassic. Palaeogeography,
513 Palaeoclimatology, Palaeoecology 404, 1–11. <https://doi.org/10.1016/j.palaeo.2014.03.026>
- 514 Dunhill, A.M., Hannisdal, B., Benton, M.J., 2014b. Disentangling rock record bias and common-
515 cause from redundancy in the British fossil record. Nature Communications 5, 4818.
516 <https://doi.org/10.1038/ncomms5818>
- 517 Foote, M., 2003. Origination and Extinction through the Phanerozoic: A New Approach. The

518 Journal of Geology 111, 125–148. <https://doi.org/10.1086/345841>

519 Friedman, M., Coates, M.I., Anderson, P., 2007. First discovery of a primitive coelacanth fin fills
520 a major gap in the evolution of lobed fins and limbs. *Evol. Dev.* 9, 329–337.
521 <https://doi.org/10.1111/j.1525-142X.2007.00169.x>

522 Gascuel, O., 1997. Concerning the NJ algorithm and its unweighted version, UNJ, in: Roberts, F.,
523 Rzhetsky, A. (Eds.), *Mathematical Hierarchies and Biology*. American Mathematical Soc.,
524 Providence, RI, pp. 149–170.

525 Gauthier, J., Cannatella, D., de Queiroz, K., Kluge, A.G., Rowe, T., 1989. Tetrapod phylogeny, in:
526 Fernholm, B., Bremer, K., Jörnvall, H. (Eds.), *The Hierarchy of Life*. Elsevier Science
527 Publishers B. V. (Biomedical Division), Amsterdam, Netherlands.

528 Gavryushkina, A., Welch, D., Stadler, T., Drummond, A.J., 2014. Bayesian Inference of Sampled
529 Ancestor Trees for Epidemiology and Fossil Calibration. *PLOS Comput Biol* 10, e1003919.
530 <https://doi.org/10.1371/journal.pcbi.1003919>

531 Gess, R., Ahlberg, P.E., 2018. A tetrapod fauna from within the Devonian Antarctic Circle. *Science*
532 360, 1120–1124. <https://doi.org/10.1126/science.aag1645>

533 Heath, T.A., Huelsenbeck, J.P., Stadler, T., 2014. The fossilized birth–death process for coherent
534 calibration of divergence-time estimates. *PNAS* 111, E2957–E2966.
535 <https://doi.org/10.1073/pnas.1319091111>

536 Jablonski, D., Roy, K., Valentine, J.W., Price, R.M., Anderson, P.S., 2003. The Impact of the Pull
537 of the Recent on the History of Marine Diversity. *Science* 300, 1133–1135.
538 <https://doi.org/10.1126/science.1083246>

539 Koch, C.F., 1978. Bias in the Published Fossil Record. *Paleobiology* 4, 367–372.

540 Lakner, C., van der Mark, P., Huelsenbeck, J.P., Larget, B., Ronquist, F., 2008. Efficiency of

- 541 Markov chain Monte Carlo tree proposals in Bayesian phylogenetics. *Syst. Biol.* 57, 86–
542 103. <https://doi.org/10.1080/10635150801886156>
- 543 Laurin, M., 1998. The importance of global parsimony and historical bias in understanding
544 tetrapod evolution. Part I. Systematics, middle ear evolution and jaw suspension. *Ann. Sci.*
545 *Nat. Zoo.* 19, 1–42. [https://doi.org/10.1016/S0003-4339\(98\)80132-9](https://doi.org/10.1016/S0003-4339(98)80132-9)
- 546 Lloyd, G.T., 2012. A refined modelling approach to assess the influence of sampling on
547 palaeobiodiversity curves: new support for declining Cretaceous dinosaur richness. *Biol.*
548 *Lett.* 8, 123–126. <https://doi.org/10.1098/rsbl.2011.0210>
- 549 Long, J.A., Clement, A.M., Choo, B., 2018. New insights into the origins and radiation of the mid-
550 Palaeozoic Gondwanan stem tetrapods. *Earth and Environmental Science Transactions of*
551 *The Royal Society of Edinburgh* 1–17. <https://doi.org/10.1017/S1755691018000750>
- 552 Marjanović, D., Laurin, M., 2019. Phylogeny of Paleozoic limbed vertebrates reassessed through
553 revision and expansion of the largest published relevant data matrix. *PeerJ* 6, e5565.
554 <https://doi.org/10.7717/peerj.5565>
- 555 Marshall, J.E.A., Reeves, E.J., Bennett, C.E., Davies, S.J., Kearsey, T.I., Millward, D., Smithson,
556 T.R., Browne, M.A.E., 2019. Reinterpreting the age of the uppermost “Old Red Sandstone”
557 and Early Carboniferous in Scotland. *Earth Env. Sci. T. R. So.* 109, 265–278.
558 <https://doi.org/10.1017/S1755691018000968>
- 559 Niedźwiedzki, G., Szrek, P., Narkiewicz, K., Narkiewicz, M., Ahlberg, P.E., 2010. Tetrapod
560 trackways from the early Middle Devonian period of Poland. *Nature* 463, 43–48.
561 <https://doi.org/10.1038/nature08623>
- 562 O’Donovan, C., Meade, A., Venditti, C., 2018. Dinosaurs reveal the geographical signature of an
563 evolutionary radiation. *Nature Ecology & Evolution* 2, 452.

- <https://doi.org/10.1038/s41559-017-0454-6>
- Pardo, J.D., Szostakiwskyj, M., Ahlberg, P.E., Anderson, J.S., 2017. Hidden morphological diversity among early tetrapods. *Nature* 546, 642–645. <https://doi.org/10.1038/nature22966>
- R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst Biol* 67, 901–904. <https://doi.org/10.1093/sysbio/syy032>
- Raup, D.M., Boyajian, G.E., 1988. Patterns of Generic Extinction in the Fossil Record. *Paleobiology* 14, 109–125.
- Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3, 217–223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>
- Robinson, D.F., Foulds, L.R., 1981. Comparison of phylogenetic trees. *Math. Biosci.* 53, 131–147. [https://doi.org/10.1016/0025-5564\(81\)90043-2](https://doi.org/10.1016/0025-5564(81)90043-2)
- Ronquist, Fredrik, Klopstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L., Rasnitsyn, A.P., 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Syst. Biol.* 61, 973–999. <https://doi.org/10.1093/sysbio/sys058>
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*.
- Rothkugel, S., Varela, S., 2015. paleoMap: An R-package for getting and using paleontological

maps.

Sakamoto, M., Benton, M.J., Venditti, C., 2016a. Dinosaurs in decline tens of millions of years before their final extinction. *PNAS* 113, 5036–5040. <https://doi.org/10.1073/pnas.1521478113>

Sakamoto, M., Venditti, C., Benton, M.J., 2016b. ‘Residual diversity estimates’ do not correct for sampling bias in palaeodiversity data. *Methods Ecol Evol* n/a-n/a. <https://doi.org/10.1111/2041-210X.12666>

Schliep, K.P., 2011. phangorn: Phylogenetic analysis in R. *Bioinformatics* 27, 592–593. <https://doi.org/10.1093/bioinformatics/btq706>

Signor, P.W., Lipps, J.H., 1982. Sampling bias, gradual extinction patterns, and catastrophes in the fossil record. *Geological Society of America Special Publication* 190, 291–296.

Stadler, T., 2010. Sampling-through-time in birth-death trees. *J. Theor. Biol.* 267, 396–404. <https://doi.org/10.1016/j.jtbi.2010.09.010>

Swartz, B., 2012. A marine stem-tetrapod from the Devonian of Western North America. *PLoS ONE* 7, e33683. <https://doi.org/10.1371/journal.pone.0033683>

Tennant, J.P., Mannion, P.D., Upchurch, P., 2016a. Environmental drivers of crocodyliform extinction across the Jurassic/Cretaceous transition. *Proc. R. Soc. B* 283, 20152840. <https://doi.org/10.1098/rspb.2015.2840>

Tennant, J.P., Mannion, P.D., Upchurch, P., 2016b. Sea level regulated tetrapod diversity dynamics through the Jurassic/Cretaceous interval. *Nature Communications* 7, 12737. <https://doi.org/10.1038/ncomms12737>

Venditti, C., Meade, A., Pagel, M., 2011. Multiple routes to mammalian diversity. *Nature* 479, 393–396. <https://doi.org/10.1038/nature10516>

- Xie, W., Lewis, P.O., Fan, Y., Kuo, L., Chen, M.-H., 2011. Improving Marginal Likelihood Estimation for Bayesian Phylogenetic Model Selection. *Syst Biol* 60, 150–160. <https://doi.org/10.1093/sysbio/syq085>
- Zhang, C., Stadler, T., Klopstein, S., Heath, T.A., Ronquist, F., 2016. Total-Evidence Dating under the Fossilized Birth–Death Process. *Syst Biol* 65, 228–249. <https://doi.org/10.1093/sysbio/syv080>
- Zhu, M., Ahlberg, P.E., Zhao, W.-J., Jia, L.-T., 2017. A Devonian tetrapod-like fish reveals substantial parallelism in stem tetrapod evolution. *Nat. Ecol. Evol.* 1, 1470–1476. <https://doi.org/10.1038/s41559-017-0293-5>

Figure Captions

1. Example of how the regional-level formation count proxy is calculated. A) Five major geographic regions are highlighted by color in the Devonian map. Red arrows represent a branch-specific dispersal path to species A, beginning in Southern Euramerica and ending in Eastern Gondwana. The blue arrow represents the dispersal path to species B. B) The phylogeny of species A and B scaled by time, with equal branch lengths to both species, and colored to represent the rate of dispersal (red is fast, blue is slow). For every branch of the tree, the number of formations is counted for every region and for each geological age covered by the dispersal pathway. It is then weighted by the number of geological ages and geographic regions covered. Under the Western Gondwana route scenario, the branch to species A covers three geographic regions, while the branch to species B only covers one. Assuming both branches cover only one geological age, the high dispersal rate for species A can be explained by the lack of recorded geological formations in Western Gondwana. C) A line plot of the formation counts through time, colored by geographic

region according to the Devonian map above, shows temporal and geographic variability.

2. The time-scaled tetrapodomorph supertree. Taxonomic groups in quotes are not monophyletic. Here, *Glyptolepis*, a dipnomorph, is the outgroup. We downloaded the silhouettes from phylopic.org: *Eucritta* and *Greererpeton* by Dmitry Bogdanov (vectorized by Michael Keesey), *Eusthenopteron* by Steve Coombs (vectorized by Michael Keesey), and *Gogonasus* and *Tiktaalik* by Nobu Tamura (CC BY-SA 3.0).

3. A) Trimmed tetrapodomorph phylogeny with mapped rates of dispersal. Cooler (bluish) colors represent slower rates and warmer (reddish) colors represent faster rates. B) Non-eotetrapodiform (left in blue) and eotetrapodiform (right in green) trees and taxon paleolocations plotted on a map of the Middle Devonian. Transparent polygons illustrate broad geographic regions of sampled taxa in Southern Euramerica, Eastern Gondwana, and East Asia. Numbers show the total number of geological formations recorded from each major geographic region (Eastern Gondwana and East Asia combined). Colored circles show average paleolocations of major clades estimated by the ‘geo’ model and indicated in the tree above. Red circle: Tetrapodomorpha, orange: “Megalichthyiformes”, yellow: “Canowindridae” + Rhizodontidae, green: Tristichopteridae, and blue: Elpistostegalia. Phylogeny with mapped dispersal rates was produced in BayesTrees (<http://www.evolution.rdg.ac.uk/BayesTrees.html>). Middle Devonian tree and paleolocation plots were made using the ‘phylo-to-map’ function in the R package, phytools (Revell, 2012). Middle Devonian map was sourced from the R package, paleoMap (Rothkugel and Varela, 2015). Tetrapodomorph silhouettes were sourced from phylopic.org: *Eucritta* by Dmitry Bogdanov (vectorized by T. Michael Keesey), *Osteolepis* by Nobu Tamura, and *Acanthostega* by Mateus Zica.

656

657 4. A) Scatter-plot of the average dispersal rates over the regional-level formation counts for each
 658 branch of the phylogeny, using the Northern Euramerica + East Asia route scenario. Points colored
 659 by the dispersal rate being above (red) or below (blue) two standard deviations greater than the
 660 average rate across the tree. B) Histogram of the bootstrapped U -statistics. Values outside of the
 661 95% confidence interval are grayed out. The median and null expected U -statistics are indicated
 662 by the red and blue dotted lines, respectively. The null expected U -statistic is based on the null
 663 hypothesis that 50% of the branches with low dispersal rates will have a greater regional-level
 664 formation count than branches with higher rates. Rejecting the null hypothesis suggests that
 665 estimated dispersal rates are biased and correlate with regional-level formation count.





