Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity

Chloé Schmidt, Stéphane Dray, Colin J Garroway

Title: Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity

Running title: Geography of nuclear genetic diversity

Authors: Chloé Schmidt[1*], Stéphane Dray[2], Colin J. Garroway[1*]

Affiliations:

[1]Department of Biological Sciences, 50 Sifton Rd, University of Manitoba, Winnipeg, Manitoba R3T 2N2 Canada

[2]Univ Lyon, Université Claude Bernard Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive, F-69100, Villeurbanne, France

*Correspondence to:
Chloé Schmidt
Department of Biological Sciences
50 Sifton Rd
University of Manitoba
Winnipeg, MB R3T 2N2
email: schmid46@myumanitoba.ca

Colin J Garroway
Department of Biological Sciences
50 Sifton Rd
University of Manitoba
Winnipeg, MB R3T 2N2
email: colin.garroway@umanitoba.ca

Abstract: The processes that give rise to species richness gradients are not well understood, but may be linked to resource-based limits on the number of species a region can support. Ecological limits placed on regional species richness would also limit population sizes, suggesting that these processes could also generate genetic diversity gradients. If true, we might better understand how broad-scale biodiversity patterns are formed by identifying the common causes of genetic diversity and species richness. We develop a hypothetical framework based on the consequences of regional variation in ecological limits to simultaneously explain spatial patterns of species richness and neutral genetic diversity. Repurposing raw genotypic data spanning 38 mammal species sampled across 801 sites in North America, we show that estimates of genome-wide genetic diversity and species richness share spatial structure. Notably, species richness hotspots tend to harbor lower levels of within-species genetic variation. A structural equation model encompassing eco-evolutionary processes related to resource availability, habitat heterogeneity, and human disturbance explained 78% of variation in genetic diversity and 74% of the variation in species richness. These results suggest we can infer broad-scale patterns of species and genetic diversity using two simple environmental measures of resource availability and ecological opportunity.

Keywords: more individuals hypothesis, heterogeneity, Anthropocene, latitudinal diversity gradient, carrying capacity, macroecology, macrogenetics

Introduction

Genetic diversity and species richness are the most fundamental levels of biodiversity because they reflect within- and across-species contributions to ecosystem functioning (Oliver et al. 2015; Des Roches et al. 2021b). Genetic diversity underlies a population's capacity to adapt in response to environmental change, and species richness enhances ecosystem resiliency to perturbation. If we are to manage the current high rates of biodiversity loss, we need to better understand how patterns of biodiversity are produced and how they interact across levels of biological organization. Patterns of species richness are well-described, but because several independent processes are capable of generating these patterns, their origins remain puzzling. We know less about multi-species patterns of genetic diversity. However, there is good reason to think that the processes forming patterns of species richness could simultaneously produce spatial patterns in neutral genetic diversity (Vellend 2005; Evanno et al. 2009). This is because spatial variation in neutral genetic diversity should reflect how local population-level demographic and evolutionary processes interact with environments to produce species richness gradients. If true, we may be able to infer processes underlying biodiversity patterns at both genetic and species levels by attempting to understand their common causes. The accumulation of open data now allows us to tackle these types of questions by repurposing and synthesizing publicly archived raw data (e.g., Leigh et al. in press; Miraldo et al. 2016; Manel et al. 2020; Schmidt et al. 2020a; Theodoridis et al. 2020; Schmidt and Garroway 2021). Here we produce a continental map of spatial variation in neutral nuclear genetic diversity for North American mammals, show that genetic diversity and species richness covary spatially and are negatively correlated, and find empirical support suggesting that measures of resource availability and

heterogeneity predict both genetic diversity and species richness patterns through their effects on demography.

We developed a conceptual framework to explain how genetic diversity and species richness patterns could emerge from common causes. This framework extends predictions from well-supported hypotheses for species richness patterns to the population genetic level. Hypotheses for species richness gradients fall into three general categories related to evolutionary time, evolutionary rates, and ecological limits (Mittelbach et al. 2007; Worm and Tittensor 2018; Pontarp et al. 2019). We focus on ecological limits hypotheses—these posit that variation in resource availability limits the number of species able to coexist in a particular area (Rabosky and Hurlbert 2015). Here the speciation, extinction, and colonization dynamics of species are analogous to the birth, death, and immigration dynamics that set carrying capacities at the population level. Simulations suggest multiple hypotheses can produce species richness gradients (Etienne et al. 2019), but the preponderance of theory suggests that ecological limits produce the strongest and most stable gradients (Vellend 2005; Worm and Tittensor 2018; Etienne et al. 2019). There is also good empirical support for the likely importance of ecological limits in the formation of species richness patterns (reviewed in Rabosky and Hurlbert 2015; Brodie 2019). We thus considered ecological limits hypotheses as parsimonious starting expectations when exploring the causes of biodiversity patterns (Etienne et al. 2019).

It is relatively straightforward to extend the consequences of ecological limits on community size to the population genetic level. If environments limit the number of supportable species, they must also limit the population sizes of species, and therefore affect the strength of genetic drift.

The first ecological limits hypothesis we consider is the more individuals hypothesis (Wright 1983). In terms of community composition, the more individuals hypothesis suggests that resource availability imposes an upper limit on the number of individuals, and as a consequence, the number of species an area can support (Currie 1991; Rabosky and Hurlbert 2015; Storch et al. 2018). Diversity tends to increase with the number of individuals in an assemblage both in terms of genetic diversity within populations and the number of species in a community (Kimura 1983; Hubbell 2001). Thus, the more individuals hypothesis predicts neutral genetic diversity and species richness will be positively correlated and increase with resource availability (Fig. 1).

The second ecological limits hypothesis we consider pertains to environmental heterogeneity, which includes variation in resources, habitat types, and habitat complexity (Stein et al. 2014). Here we assume heterogeneity equates to niche availability. The idea of area-heterogeneity trade-offs suggests that heterogeneous areas can support richer communities of more specialized species, but these species should tend to have smaller population sizes because resources and species are divided among niches (Kadmon and Allouche 2007; Allouche et al. 2012). Local adaptation, and subsequently specialization, can also occur within and across species distributed across heterogeneous environments. As increasingly specialized populations diverge, genetic variation would be partitioned among locally adapted populations that may eventually no longer interbreed. Compared to larger populations, these smaller populations would also more rapidly lose genetic diversity due to genetic drift. If this were the case, we expect environmental heterogeneity would be positively associated with species richness and negatively associated with neutral genetic diversity (Fig. 1).

Contemporary rapid environmental change also affects biodiversity patterns, yet it is not typically modelled in a way that makes it comparable to historical processes acting over long

periods. A major contemporary ecological limit on diversity is human land transformation. Human activities such as urbanization reduce the amount of habitat available to wild populations (McKinney 2006; Grimm et al. 2008) with consequences at genetic and species levels (Ceballos et al. 2015; WWF 2018; Leigh et al. 2019; Schmidt et al. 2020a). Habitat loss, fragmentation, and homogenization resulting from human land use alters resource and niche availability, thus processes associated with ecological limits should play out in populations and communities of urban wildlife. By reducing habitable area and resource heterogeneity, we predicted that the effects of urbanization for mammals should also cause species richness and genetic diversity to decrease in more heavily disturbed areas (Fig. 1).

The effects of resource availability and heterogeneity are not mutually exclusive, and in our framework they can act in concert to produce biodiversity patterns. The links among our hypotheses and their predictions are diagrammed in full in Figure 1. We jointly model both hypotheses with a method that allows us to assess their relative importance for shaping genetic diversity and species richness. Our predictions for the ways resource availability and heterogeneity interact are consistent with previous work on species richness in North American mammals (Kerr and Packer 1997), where heterogeneity becomes a more important determinant of species richness as resource availability increases. If our model successfully captures known relationships between species richness and environments, and genetic diversity behaves in the ways we predict, we will have strong empirical evidence supporting the contention that continental patterns of neutral genetic diversity and species richness are both are in part governed by ecological limits.

Our specific objectives were threefold. Because biogeographic patterns of neutral nuclear genetic diversity have not yet been mapped, we first produced a continental map of spatial patterns of genetic diversity in North American mammals. To do this we repurposed publicly archived, raw, neutral nuclear genetic data spanning 38 species and >34,000 individuals at 801 sample sites in the United States and Canada. We then tested the degree to which patterns of genetic diversity matched those of species richness. Having established shared patterns of spatial variation, we then tested our proposed conceptual model based on ecological limits hypotheses where genetic diversity and species richness are caused by common environmental factors (Fig. 1). We tested our hypothetical model using structural equation modelling (SEM), a modelling framework that fits hypothesis networks by accommodating multiple predictor and response variables. Our approach (Fig. 2) allowed us to assess the relative importance of both hypotheses and the effects of contemporary environmental change while accounting for species-level variation using hierarchical models.

Methods

Biodiversity data

Genetic diversity. We used raw genotypic data compiled by Schmidt et al. (2020a,b). This data set is comprised of repurposed raw microsatellite data from 34,841 individuals from 38 mammalian species sampled at 801 sites in the United States and Canada. With it we could consistently calculate measures of gene diversity (Nei 1973) and population-specific F_{ST} across sites (Weir and Goudet 2017). See Table 1 for a summary of the dataset. Microsatellite markers estimate genome-wide diversity well (e.g., microsatellite and genome-wide diversity are correlated at $R^2 \sim 0.83$; Mittell et al. 2015). They are commonly used in wildlife population

genetic studies because they are cost-effective and do not require a reference genome, which allowed us to maximize sample size. Detailed methods for assembling this dataset can be found in (Schmidt et al. 2020a). Briefly, we performed a systematic search for species names of native North American mammals with keywords "microsat*", "single tandem*", "short tandem*", and "str" using the dataone R package, which interfaces with the DataONE platform to search online open data repositories (Jones et al. 2017). We discarded search results that did not meet our criteria for inclusion and removed results where study design may have influenced genetic diversity. For example, we excluded non-neutral data and samples taken after a recent bottleneck, translocations, managed or captive populations, or island populations. We additionally removed populations with fewer than 5 individuals sampled. Gene diversity estimates the richness and evenness of alleles in a population, and we used it here as our metric for genetic diversity because it is minimally affected by sample size (Charlesworth and Charlesworth 2010) (Fig. S1). Sample sites are treated as point locations.

Species richness. We downloaded range maps for terrestrial mammals native to North America from the IUCN Red List database (IUCN 2019). We filtered these maps to retain ranges for extant, native, resident, mainland species in R version 4.0.1 (R Core Team 2020). To generate a map of species richness coincident with genetic sample sites, we estimated species richness by summing the number of ranges overlapping each site.

Maps and spatial variation partitioning

Genetic diversity and species richness maps. Our first step was to map spatial patterns in genetic diversity. We accomplished this using distance-based Moran's eigenvector maps (MEMs) in the package adespatial (Dray et al. 2017). MEMs detect spatial patterns in data using a matrix of

distances between sites—a neighbor matrix—whose eigenvalues are proportional to Moran's I index of spatial autocorrelation (Borcard and Legendre 2002; Borcard et al. 2004; Dray et al. 2006). MEMs are spatial eigenvectors that represent relationships between sites at all spatial scales detectable by the sampling scheme. Multiple MEMs can be included in linear models to identify spatial patterns in data because they are orthogonal. They are appropriate for use in genetics because Moran's I is a direct analog of Malécot's estimator of spatial autocorrelation of allele frequencies (Malécot 1955; Epperson 2005) which accurately summarizes neutral variation in gene flow and allele frequencies (e.g., Sokal and Oden 1978; Epperson 2005). Distance-based MEM analysis produces $n - 1$ MEMs (n being the number of sample sites), but only eigenvectors corresponding to positive spatial autocorrelation are used. MEMs are ordered according to spatial scale explained, with the first eigenvector explaining the broadest autocorrelation pattern. We used linear regressions and the forward selection procedure described in (Blanchet et al. 2008) to select two sets of MEMs: one describing spatial patterns in genetic diversity and the other describing species richness. Thirteen MEMs, ranging from broad to fine scales, explained important spatial variation in gene diversity. Forty-three MEMs were important predictors of species richness, and 8 of these patterns were shared by genetic diversity (significant MEMs are listed in Fig. S3).

To restrict ourselves to broad spatial patterns, we focused on genetic and species MEMs with Moran's I values >0.25. We fit individual linear regression models for species richness and genetic diversity with corresponding broad-scale MEMs as covariates and plotted model predicted values representing spatial patterns on maps of North America (Fig. 3). These MEMs describe the broadest-scale spatial patterns at both levels of diversity. By using values of genetic diversity and species richness described by these MEMs, we can visualize pure spatial variation

at the continental scale without local spatial patterns that may be due to environmental idiosyncrasies, and without considering non-spatial variation in genetic diversity. We also provided maps of raw genetic diversity and species richness values in Figure S2.

Variation partitioning. We next quantified the extent to which genetic diversity and species richness covary spatially. Because MEMs for species richness and genetic diversity were computed from the same set of coordinates, they were directly comparable. This allowed us to identify shared spatial MEMs. We used linear regressions and variance partitioning to determine what fraction of the total variation in species richness and genetic diversity could be attributed to: (1) non-spatial variation, (2) non-shared spatial variation, and (3) shared spatial variation. We partitioned variation as follows:

$$y_{SR} \sim \alpha + \beta_{1S}(\mathrm{MEM}_{1S}) + \beta_{2S}(\mathrm{MEM}_{2S}) + \cdots + \beta_{43S}(\mathrm{MEM}_{43S}) + \epsilon$$

$$y_{GD} \sim \alpha + \beta_{1G}(\mathrm{MEM}_{1G}) + \beta_{2G}(\mathrm{MEM}_{2G}) + \cdots + \beta_{13G}(\mathrm{MEM}_{13G}) + \epsilon$$

where α is the grand mean, and y_{SR} and y_{GD} are species richness and genetic diversity at sites. MEM_S and MEM_G refer to the set of MEMs explaining spatial variation in species richness (43 MEMs) and genetic diversity (13 MEMs), respectively, and βs are their slopes. The coefficients of variation (R^2) for these models give us the proportion of variation in each response variable attributable to spatial variation. Subtracting these values from 1 gives the amount of non-spatial variation.

To determine the amount of shared variation, we used the set of 8 MEMs shared between species richness and genetic diversity (MEM_{SG}) as predictors in the regressions below:

$$y_{SR} \sim \alpha + \beta_{1SG}(\mathrm{MEM}_{1SG}) + \beta_{2SG}(\mathrm{MEM}_{2SG}) + \cdots + \beta_{8SG}(\mathrm{MEM}_{8SG}) + \epsilon$$

$$y_{GD} \sim \alpha + \beta_{1SG}(\mathrm{MEM}_{1SG}) + \beta_{2SG}(\mathrm{MEM}_{2SG}) + \cdots + \beta_{8SG}(\mathrm{MEM}_{8SG}) + \epsilon$$

R^2 values from these models estimate the proportion of variation in genetic diversity and species richness explained by shared spatial variation. Subtracting these values from the total spatial variation in species richness and genetic diversity gives the proportion of non-shared spatial variation.

Structural equation modeling

Population size data. The more individuals hypothesis is most applicable at broad spatial scales, and when considering the total number of individuals that comprise a species (Storch et al. 2018). In place of census sizes for the species in our dataset, which are not consistently available, we craft our hypothesis around species' long-term effective population sizes. The effective population size is a concept defined in population genetics as the number of individuals in an idealized population that experiences the same rate of genetic drift as the real population (Charlesworth and Charlesworth 2010). Populations lose genetic diversity to drift at a rate inversely proportional to the effective population size. Body size is routinely used as a proxy for long-term effective population size at the species level (Frankham 1996; Corbett-Detig et al. 2015; Mackintosh et al. 2019; Buffalo 2021). Large bodied species that tend to have long lifespans and produce few offspring generally have smaller effective population sizes than small, fecund, short-lived species (Romiguier et al. 2014; Mackintosh et al. 2019). Thus body size measured at the species level is an imperfect, but nevertheless useful substitution for census size. We recorded adult body mass (g) for each species in our genetic dataset from the PanTHERIA database (Jones et al. 2009) and log-transformed values before analysis. There were no obvious outliers in these data.

255 *Environmental variables.* We used potential evapotranspiration as a surrogate for total ecosystem
256 resource availability (Currie 1991; Rabosky and Hurlbert 2015). Potential evapotranspiration is
257 an indicator of atmospheric energy availability and is one of the strongest environmental
258 correlates of species richness in North American mammals (Currie 1991; Kreft and Jetz 2007;
259 Fisher et al. 2011; Jiménez-Alfaro et al. 2016). As we predict, at the species level, that resource
260 availability across a range sets the long-term effective population size, we estimated mean range-
261 wide potential evapotranspiration (mm/yr) using annual data from 1970-2000 available via the
262 CGIAR Consortium for Spatial Information (Trabucco and Zomer 2019). For comparison, we
263 also measured mean range-wide actual evapotranspiration, an alternative measure of resource
264 availability, and present those results in the Supplementary Information.

265 We quantified heterogeneity and niche availability using a 250 m resolution map of land cover
266 types in North America (CEC et al. 2010). This map includes 19 land cover categories based on
267 satellite imagery collected in 2010 with multiple categories of forest, shrubland, grassland, polar
268 habitat types, wetland, cropland, barren land, built up land, and open water. Because the
269 heterogeneity hypothesis suggests species specialize on different resources within their range, we
270 quantified heterogeneity at sites rather than at the species level. We measured heterogeneity
271 using Simpson's diversity index. To assess scale dependence, we calculated Simpson's index
272 within four buffer zones around each site: 5000, 20000, 50000, and 100000 km^2. Lastly, we
273 quantified human disturbance at each site using human population density (CIESIN 2016)
274 measured within a 10 km buffer following Schmidt et al. (2020a).

275 *Analysis.* Structural equation modeling accommodates multiple dependent and independent
276 variables in a model network, and directional paths connecting variables represent causal

277 relationships. The strengths of direct paths are regression coefficients (Shipley 2016), and
278 indirect effects can be quantified by multiplying coefficients along paths of direct effects.
279 We constructed a graph of our conceptual model laid out in the introduction (Fig. 1), which we
280 then translated into a network of three linear models for body size, species richness, and genetic
281 diversity. In it, body size is predicted by resource availability, and species richness and genetic
282 diversity are each predicted by body size, heterogeneity, and human disturbance (Fig. 4a). We fit
283 structural equation models using piecewiseSEM in R (Lefcheck 2016; Lefcheck et al. 2019)
284 because this package accommodates complex model structures. We used a linear mixed-effects
285 model with a random intercept for species to account for species-level variation in genetic
286 diversity. PiecewiseSEM fits hierarchical models using the lme4 package (Bates et al. 2015).
287 Body size and species richness models were initially fit as linear regressions, but residuals from
288 both models were spatially autocorrelated at broad scales. We refit these regressions using
289 simultaneous autoregressive models in spatialreg (Bivand et al. 2013) and this successfully
290 removed spatial autocorrelation from the residuals. All variables were scaled and centered before
291 analysis to obtain standardized regression coefficients, allowing us to compare the strength of
292 relationships and the relative support for hypotheses across genetic and species levels.

293 Our model includes variables measured at the site level (genetic diversity, species richness,
294 heterogeneity, and human disturbance) and species level (body size, resource availability; Fig.
295 2). This hierarchical data structure introduces spurious correlations between variables sampled at
296 different levels that we know are not causal. For example, regressing human disturbance at sites
297 on species body size would estimate an artefactual experiment (the size of species researchers
298 choose to sample near cities)—not the effects of disturbance on body size. We can account for
299 these known non-causal relationships by allowing variables to have correlated errors (Lefcheck

2016). Correlated errors indicate that a relationship exists between variables, but allow the direction of causality to be ambiguous: both could be caused by another factor not included in the model (e.g., researcher species choice). We specified correlated errors between body size and human population density, and body size and heterogeneity.

The conceptual model is evaluated by testing whether additional links are needed between variables to make the proposed causal structure more consistent with the data. In piecewiseSEM, missing links are tested using tests of directed separation (Shipley 2016), where the null hypothesis is that two variables are independent conditional on other predictors in the model. A *p* value for the model network is obtained by testing Fisher's C calculated from the *p* values summed across directed separation tests (Lefcheck 2016; Shipley 2016). A model-wide $p < 0.05$ means the causal structure is not a good fit to the data and additional links are needed to resolve dependencies. If $p > 0.05$, the model is considered acceptable because we fail to reject our causal structure. This means that although we start with a focus on our conceptual model, the data can suggest the addition or removal of links and our hypotheses can be updated for future testing with new data.

We assessed model fit using R^2 values for each response variable in the model network. For genetic diversity, we used marginal R^2 (R^2_m) which measures the total variation explained by fixed effects, and conditional R^2 (R^2_c) which is the variation explained by both fixed and random effects. For spatial body size and species richness regressions, we report Nagelkerke pseudo-R^2.

Effect of heterogeneity on population divergence

After detecting a negative effect of heterogeneity on intraspecific genetic diversity in our structural equation model, we performed a post hoc analysis to test whether environmental

heterogeneity also caused greater population differentiation within species. To test for differentiation we calculated population-specific F_{ST} (Weir and Goudet 2017) as a measure of genetic divergence using the hierfstat package (Goudet and Jombart 2015). Population-specific F_{ST} can be interpreted as a relative estimate of the time since a population has diverged from a common ancestor. This metric required at least 2 sampled populations in the original studies to estimate, and due to this constraint 16 sites were excluded from this analysis. We controlled for isolation-by-distance by including MEMs significantly related to F_{ST} to account for spatial structure. We scaled and centered all variables, then used a linear mixed model controlling for species differences by including it as a random effect.

Results

Spatial patterns in genetic diversity and species richness

There was no obvious relationship between latitude and nuclear genetic diversity (Fig. 3). Similar to patterns of species richness, a longitudinal gradient in genetic diversity was the dominant pattern for North American mammals—however, it appears regions with high species richness have lower genetic diversity. We detected spatial patterns at genetic and species levels of diversity. Sixty-five percent of the total variation in species richness and 24% of variation in genetic diversity was spatially structured (Fig. 3). Variance partitioning suggested that 85% of the total spatial variation in genetic diversity, and 32% of spatial variation in species richness was accounted for by spatial patterns shared at both levels of diversity. This shared variation implies that, to an extent, neutral genetic diversity and species richness are simultaneously shaped by spatially structured environmental factors, and these shared processes account for most of the spatial variation in genetic diversity.

346

347 *Joint environmental causes of genetic diversity and species richness*

348 We present results from the model using a 5000 km^2 heterogeneity buffer in the main text.
349 Results from wider heterogeneity buffers can be found in SI Tables S1-S4. Our conceptual
350 model, updated according to tests of directed separation, fit the data well (SEM p= 0.33, Fisher's
351 C= 2.245; Fig. 4, Table S1). Note that for structural equation models, $p > 0.05$ indicates that we
352 fail to reject our model. There was no residual spatial autocorrelation in body size and species
353 richness model residuals, and genetic diversity model residuals were spatially autocorrelated at
354 local scales (genetic diversity Moran's I = 0.01). These Moran's I values indicate very weak
355 spatial structure in the data, and so we decided not to integrate it into our model. Positive spatial
356 autocorrelation at such short distances is likely an artifact of irregular site locations and the
357 hierarchical nature of the data. A lack of strong spatial autocorrelation in the model residuals
358 suggests that the spatial structure of the diversity data was well captured by our model's
359 environmental covariates (Fig. S3). All predicted links in our conceptual model were supported
360 except that between body size and species richness (Fig. 4). Tests of directed separation
361 suggested additional direct links from resource availability to species richness, and genetic
362 diversity to species richness (Fig. 4). Effects of heterogeneity on genetic diversity were not
363 detectable at broader scales (Tables S2-S4). These relationships were consistent using actual
364 evapotranspiration as an alternative measure of resource availability (Table S5).

365 Resource availability, heterogeneity, and human disturbance, acting both directly and indirectly
366 through species population size, explained 20% of the variation in genetic diversity. The species-
367 level variation explained by the random effect brought the total variation in genetic diversity
368 explained by our model to 78%. The same model explained 74% of the variation in species

richness. Genetic diversity was strongly negatively related to body size. The direct effects of resource heterogeneity on species richness and genetic diversity were of similar magnitude, and directions of effects were as expected if processes associated with greater resource heterogeneity reduce population sizes, lead to increased genetic drift, and facilitate local adaptation and coexistence (Fig. 4, Table S1). However, because gene diversity is not a measure of divergence, we additionally tested whether environmental heterogeneity predicted evolutionary divergence at the population level. Heterogeneity was positively related to genetic divergence, but the effect was not significant ($\beta = 0.06 \pm 0.04$ SE). Finally, human disturbance negatively affected both species richness and genetic diversity, but its effects were stronger for genetic diversity (Fig. 4, Table S1).

Discussion

We found striking continental spatial gradients in nuclear genetic diversity, and show that these patterns are negatively correlated with patterns of species richness in North America (Simpson 1964) (Fig. 3). A considerable portion of the variation in genetic diversity and species richness patterns could be explained by just three environmental factors: resource availability, resource heterogeneity, and human disturbance. This is strong empirical evidence suggesting that genetic diversity and species richness patterns emerge, in part, from the same environmental processes.

Both our maps and our structural equation model suggest that resource availability and heterogeneity interact to produce biodiversity patterns at genetic and species levels. In North America, the threshold where environmental heterogeneity presumably becomes a more important determinant of species richness than resource availability lies roughly along the US-

392 Canada border where potential evapotranspiration reaches ~1000 mm/yr (Kerr and Packer 1997).
393 Near this threshold is also where we see longitudinal patterns of genetic diversity emerge.
394 Although the negative correlation between spatial patterns of genetic diversity and species
395 richness is most apparent in species richness hotspots (particularly in the southwest), structural
396 equation modeling incorporating both hypotheses gives us a more nuanced view of the
397 connections between these patterns. Indeed, effects related to resource availability and
398 heterogeneity were of similar magnitude (Fig. 4). Population size and genetic diversity increased
399 with resource availability, and though the link between species' long-term effective population
400 sizes and species richness was unsupported, species richness also increased with resource
401 availability. Moreover, we detected a positive relationship between genetic diversity and species
402 richness as predicted if population and community size increase with resource availability. It
403 may be that effective population size, measured using species body size, is too coarse an
404 indicator of census population size to detect an effect on species richness at sites—if so, site-
405 level measures of genetic diversity could be a better indicator of local population sizes.
406
407 Our results suggest that once a minimum energy threshold is reached, populations can afford to
408 specialize in heterogeneous environments while maintaining viable population sizes. In this way,
409 the interplay between ecological limits and ecological opportunity simultaneously produces
410 biogeographic patterns in genetic diversity and species richness. This interpretation of our results
411 assumes that an environmentally set equilibrium between speciation, immigration and extinction
412 has been reached. There is good evidence for this in North American mammals, where
413 diversification rates have slowed as diversity increased (Alroy 2009; Brodie 2019). However, the
414 specific ways environments shape nuclear genetic- and species-level diversity will likely differ

415 across taxa and regions depending on whether or not they have reached equilibrium (e.g.,
416 Schmidt et al. 2021). Though we measure contemporary genetic diversity, historical variation in
417 resource availability and heterogeneity likely contribute to the patterns we detect because they
418 reflect whether populations have experienced large contractions in the recent past (Hewitt 2000).
419 However, in the past when communities may not have been at equilibrium, it seems likely that
420 other processes could have been the predominant drivers of biodiversity patterns. Indeed,
421 hypotheses about species richness patterns have likely been a topic of debate for so long because
422 several processes operating with different importance across the timeline of diversification are
423 capable of producing gradients (Etienne et al. 2019). It has been suggested that time for
424 speciation should be most detectable more immediately following broad-scale environmental
425 change, and when all regions are colonized, habitats that provide more opportunities for
426 speciation should over time become the most diverse (Pontarp and Wiens 2017). As diversity
427 increases, diversification rates slow as regions approach equilibrium (Brodie 2019). It follows
428 that the relative importance of evolutionary time and diversification rates as contributors to
429 biodiversity patterns varies with time with patterns ultimately affected by variation in ecological
430 limits (Rabosky and Hurlbert 2015; Pontarp and Wiens 2017; Storch et al. 2018).
431
432 Contemporary environmental change is our chance to explore pre-equilibrium processes. Cities
433 are the newest and most rapidly expanding biome, and it is clear that they have already
434 profoundly affected biodiversity patterns (Palumbi 2001; WWF 2018; Schmidt et al. 2020a). At
435 this early stage of colonization it is unlikely that urban communities have reached equilibrium,
436 suggesting processes related to evolutionary time and diversification will predominate until more
437 niches are occupied. Indeed, there is some evidence that following an initial extinction debt after

rapid urbanization, older cities support higher species richness (Aronson et al. 2014). Human disturbance had a negative effect on genetic diversity in our model, and also reduces gene flow in mammals (Schmidt et al. 2020a). This suggests that there is potential for population divergence and local adaptation if new urban niches are exploited and spatially varying selection is sufficiently strong in cities. The extent to which urban populations adapt to local environmental conditions is an ongoing and active field of study, and no consensus has been reached (Lambert et al. 2021). Equilibrium levels of genetic diversity and species richness in urban communities thus seem likely to strongly depend on resource availability and heterogeneity both within and across cities, but these aspects of urban environments are not yet well defined or understood (Norton et al. 2016; Des Roches et al. 2021a).

Notably, the negative correlation we find between spatial patterns of species richness and nuclear genetic diversity runs opposite the relatively consistent positive correlations found between species richness and mitochondrial genetic diversity gradients (Martin and McKay 2004; Adams and Hadly 2012; Miraldo et al. 2016; Manel et al. 2020; Theodoridis et al. 2020). Mitochondrial DNA has several idiosyncrasies associated with the specific biology of mitochondria that distinguish it from genetic diversity measured with neutral nuclear DNA (Schmidt and Garroway 2021). The most commonly used mitochondrial markers are the protein-coding genes *cytochrome oxidase I* and *cytochrome b*, which very likely do not evolve under neutrality (Galtier et al. 2009). Unlike neutral nuclear DNA, mitochondrial genetic diversity is not consistently related to life history, ecological traits, or census and effective population sizes (Bazin et al. 2006; Nabholz et al. 2008; James and Eyre-Walker 2020). Mitochondrial genetic diversity is thus a very different quantity than the neutral nuclear diversity estimates we use here,

and its lack of relationship with population size makes it unsuited for testing hypotheses based on ecological limits. Using genetic diversity metrics estimated from neutral nuclear DNA allows us to more directly link environments to species richness through demography, population size, and by extension, species life history traits which partly set the effective population size.

Ecosystem sustainability, given environmental perturbations occurring more frequently due to human causes, depends on the resiliency of landscapes, communities, and populations (Oliver et al. 2015). Our framework and the results presented herein suggest that we can understand continental patterns of species richness and genetic diversity using two simple measures of resource availability and heterogeneity. This is potentially informative for conservation practices aiming to conserve both of these levels of biodiversity at once. Maps of neutral nuclear genetic diversity can identify regions where long-term effective population sizes may have been historically small, indicating areas where low levels of neutral genetic diversity are not necessarily of immediate conservation concern (e.g., Yates et al. 2019). However, population declines due to recent human disturbance in areas with historically low genetic diversity may warrant specific attention. Furthermore, designing protected area networks based on species richness to maintain beta diversity, or variation between sites (Bush et al. 2016; Socolar et al. 2016), will likely also capture differentiated populations with complementary genetic compositions. The connections between environments, species richness, and genetic diversity we find here suggest we should be able to make informed decisions for the joint conservation of species and genetic diversity with knowledge of few environmental parameters.

Author contributions: C.J.G. and C.S. conceptualized the study. C.S., S.D. and C.J.G. designed the study and C.S. conducted the statistical analysis with input from S.D. and C.J.G. All authors contributed to data interpretation. C.S. wrote the first draft of the manuscript and all authors participated in editing subsequent manuscript drafts.

Acknowledgements: We would like to thank the Population Ecology and Evolutionary Genetics group for their feedback on this manuscript. We are also grateful to the authors whose work provided the raw data for this synthesis. C.S. and C.J.G. were supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant to C.J.G. C.S. was also supported by a U. Manitoba Graduate Fellowship, and a U. Manitoba Graduate Enhancement of Tri-council funding grant to C.J.G.

Data availability: Synthesized genetic data is available from the Dryad Data Repository (DOI: 10.5061/dryad.cz8w9gj0c). Species range boundary files and environmental data are available from open online sources (see Methods).

References

Adams, R. I., and E. A. Hadly. 2012. Genetic diversity within vertebrate species is greater at lower latitudes. Evol. Ecol. 27:133–143.

Allouche, O., M. Kalyuzhny, G. Moreno-Rueda, M. Pizarro, and R. Kadmon. 2012. Area-heterogeneity tradeoff and the diversity of ecological communities. Proc. Natl. Acad. Sci. U. S. A. 109:17495–17500.

Alroy, J. 2009. Speciation and extinction in the fossil record of North American mammals. Pp. 301–323 *in* R. K. Butlin, J. R. Bridle, and D. Schluter, eds. Speciation and Patterns of Diversity. Cambridge University Press, Cambridge.

Aronson, M. F. J., F. A. La Sorte, C. H. Nilon, M. Katti, M. A. Goddard, C. A. Lepczyk, P. S. Warren, N. S. G. Williams, S. Cilliers, B. Clarkson, C. Dobbs, R. Dolan, M. Hedblom, S.

510 Klotz, J. L. Kooijmans, I. Kühn, I. MacGregor-Fors, M. McDonnell, U. Mörtberg, P. Pyšek,
511 S. Siebert, J. Sushinsky, P. Werner, and M. Winter. 2014. A global analysis of the impacts
512 of urbanization on bird and plant diversity reveals key anthropogenic drivers. Proc. R. Soc.
513 B Biol. Sci. 281:20133330.

514 Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using
515 lme4. J. Stat. Softw. 67.

516 Bazin, E., S. Glémin, and N. Galtier. 2006. Population size does not influence mitochondrial
517 genetic diversity in animals. Science (80-.). 312:570–572.

518 Bivand, R. S., E. Pebesma, and V. Gomez-Rubio. 2013. Applied spatial data analysis with R,
519 Second edition. Springer, NY.

520 Blanchet, G. F., P. Legendre, and Borcard. 2008. Forward selection of explanatory variables.
521 Ecology 89:2623–2632.

522 Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of
523 principal coordinates of neighbour matrices. Ecol. Modell. 153:51–68.

524 Borcard, D., P. Legendre, C. Avois-Jacquet, and H. Tuomisto. 2004. Dissecting the spatial
525 structure of ecological data at multiple scales. Ecology 85:1826–1832.

526 Brodie, J. F. 2019. Environmental limits to mammal diversity vary with latitude and global
527 temperature. Ecol. Lett. 22:480–485.

528 Buffalo, V. 2021. Why do species get a thin slice of π? Revisiting Lewontin's Paradox of
529 Variation. bioRxiv 2021.02.03.429633.

530 Bush, A., T. Harwood, A. J. Hoskins, K. Mokany, and S. Ferrier. 2016. Current Uses of Beta-
531 Diversity in Biodiversity Conservation: A response to Socolar et al. Trends Ecol. Evol.
532 31:337–338.

533 Ceballos, G., P. R. Ehrlich, A. D. Barnosky, A. García, R. M. Pringle, and T. M. Palmer. 2015.
534 Accelerated modern human–induced species losses: Entering the sixth mass extinction. Sci.
535 Adv. 1:e1400253.

536 CEC, NRCan/CCMEO, USGS, INEGI, CONABIO, and CONAFOR. 2010. 2010 North
537 American Land Cover at 250 m spatial resolution.

538 Charlesworth, B., and D. Charlesworth. 2010. Elements of evolutionary genetics. Roberts &
539 Company Publishers, Greenwood Village, Colorado, USA.

540 CIESIN. 2016. Gridded Population of the World, Version 4 (GPWv4): Population Density.
541 NASA Socioeconomic Data and Applications Center (SEDAC).

542 Corbett-Detig, R. B., D. L. Hartl, and T. B. Sackton. 2015. Natural Selection Constrains Neutral
543 Diversity across A Wide Range of Species. PLoS Biol. 13:1–25.

544 Currie, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. Am.
545 Nat. 137:27–49.

546 Des Roches, S., K. I. Brans, M. R. Lambert, L. R. Rivkin, A. M. Savage, C. J. Schell, C. Correa,

L. De Meester, S. E. Diamond, N. B. Grimm, N. C. Harris, L. Govaert, A. P. Hendry, M. T. J. Johnson, J. Munshi-South, E. P. Palkovacs, M. Szulkin, M. C. Urban, B. C. Verrelli, and M. Alberti. 2021a. Socio-eco-evolutionary dynamics in cities. Evol. Appl. 14:248–267.

Des Roches, S., L. H. Pendleton, B. Shapiro, and E. P. Palkovacs. 2021b. Conserving intraspecific variation for nature's contributions to people. Nat. Ecol. Evol., doi: 10.1038/s41559-021-01403-5.

Dray, S., G. Blanchet, D. Borcard, S. Clappe, G. Guenard, T. Jombart, G. Larocque, P. Legendre, N. Madi, and H. H. Wagner. 2017. adespatial: Multivariate Multiscale Spatial Analysis.

Dray, S., P. Legendre, and P. R. Peres-Neto. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol. Modell. 196:483–493.

Epperson, B. K. 2005. Estimating dispersal from short distance spatial autocorrelation. Heredity (Edinb). 95:7–15.

Etienne, R. S., J. S. Cabral, O. Hagen, F. Hartig, A. H. Hurlbert, L. Pellissier, M. Pontarp, and D. Storch. 2019. A minimal model for the latitudinal diversity gradient suggests a dominant role for ecological limits. Am. Nat. 194:E122–E133.

Evanno, G., E. Castella, C. Antoine, G. Paillat, and J. Goudet. 2009. Parallel changes in genetic diversity and species diversity following a natural disturbance. Mol. Ecol. 18:1137–1144.

Fisher, J. B., R. J. Whittaker, and Y. Malhi. 2011. ET come home: Potential evapotranspiration in geographical ecology. Glob. Ecol. Biogeogr. 20:1–18.

Frankham, R. 1996. Relationship of Genetic Variation to Population Size in Wildlife. Conserv. Biol. 10:1500–1508.

Galtier, N., B. Nabholz, S. Glémin, and G. D. D. Hurst. 2009. Mitochondrial DNA as a marker of molecular diversity: A reappraisal. Mol. Ecol. 18:4541–4550.

Goudet, J., and T. Jombart. 2015. hierfstat: Estimation and Tests of Hierarchical F-Statistics.

Grimm, N. B., S. H. Faeth, N. E. Golubiewski, C. L. Redman, J. Wu, X. Bai, and J. M. Briggs. 2008. Global Change and the Ecology of Cities. Science (80-.). 319:756–760.

Hewitt, G. M. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

Hubbell, S. P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton NJ.

IUCN. 2019. The IUCN Red List of Threatened Species. Version 2019-1.

James, J., and A. Eyre-Walker. 2020. Mitochondrial DNA sequence diversity in mammals: a correlation between the effective and census population sizes. Genome Biol. Evol., doi: 10.1093/gbe/evaa222.

Jiménez-Alfaro, B., M. Chytrý, L. Mucina, J. B. Grace, and M. Rejmánek. 2016. Disentangling vegetation diversity from climate-energy and habitat heterogeneity for explaining animal geographic patterns. Ecol. Evol. 6:1515–1526.

584 Jones, K. E., J. Bielby, M. Cardillo, S. A. Fritz, J. O'Dell, C. D. L. Orme, K. Safi, W. Sechrest,
585 E. H. Boakes, C. Carbone, C. Connolly, M. J. Cutts, J. K. Foster, R. Grenyer, M. Habib, C.
586 A. Plaster, S. A. Price, E. A. Rigby, J. Rist, A. Teacher, O. R. P. Bininda-Emonds, J. L.
587 Gittleman, G. M. Mace, and A. Purvis. 2009. PanTHERIA: a species-level database of life
588 history, ecology, and geography of extant and recently extinct mammals. Ecology 90:2648–
589 2648.

590 Jones, M. B., P. Slaughter, R. Nahf, C. Boettiger, C. Jones, J. Read, L. Walker, E. Hart, and S.
591 Chamberlain. 2017. dataone: R Interface to the DataONE REST API.

592 Kadmon, R., and O. Allouche. 2007. Integrating the effects of area, isolation, and habitat
593 heterogeneity on species diversity: A unification of island biogeography and niche theory.
594 Am. Nat. 170:443–454.

595 Kerr, J. T., and L. Packer. 1997. Habitat heterogeneity as a determinant of mammal species
596 richness. Nature 385:253–254.

597 Kimura, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press,
598 Cambridge.

599 Kreft, H., and W. Jetz. 2007. Global patterns and determinants of vascular plant diversity. Proc.
600 Natl. Acad. Sci. 104:5925–5930.

601 Lambert, M. R., K. I. Brans, S. Des Roches, C. M. Donihue, and S. E. Diamond. 2021. Adaptive
602 Evolution in Cities: Progress and Misconceptions. Trends Ecol. Evol. 36:239–257. Elsevier
603 Ltd.

604 Lefcheck, J., J. Byrnes, and J. Grace. 2019. piecewiseSEM: Piecewise Structural Equation
605 Modeling.

606 Lefcheck, J. S. 2016. piecewiseSEM: Piecewise structural equation modelling in r for ecology,
607 evolution, and systematics. Methods Ecol. Evol. 7:573–579.

608 Leigh, D. M., A. P. Hendry, E. Vázquez-Domínguez, and V. L. Friesen. 2019. Estimated six per
609 cent loss of genetic variation in wild populations since the industrial revolution. Evol. Appl.
610 12:1505–1512.

611 Mackintosh, A., D. R. Laetsch, A. Hayward, B. Charlesworth, M. Waterfall, R. Vila, and K.
612 Lohse. 2019. The determinants of genetic diversity in butterflies. Nat. Commun. 10:1–9.

613 Malécot, G. 1955. Decrease of relationship with distance. Cold Spring Harb. Symp. Quant. Biol.
614 20:52–53.

615 Manel, S., P. E. Guerin, D. Mouillot, S. Blanchet, L. Velez, C. Albouy, and L. Pellissier. 2020.
616 Global determinants of freshwater and marine fish genetic diversity. Nat. Commun. 11:1–9.
617 Springer US.

618 Martin, P. R., and J. K. McKay. 2004. Latitudinal variation in genetic divergence of populations
619 and the potential for future speciation. Evolution (N. Y). 58:938–945.

620 McKinney, M. L. 2006. Urbanization as a major cause of biotic homogenization. Biol. Conserv.
621 127:247–260.

Miraldo, A., S. Li, M. K. Borregaard, A. Florez-Rodriguez, S. Gopalakrishnan, M. Rizvanovic, Z. Wang, C. Rahbek, K. A. Marske, and D. Nogues-Bravo. 2016. An Anthropocene map of genetic diversity. Science (80-.). 353:1532–1535.

Mittelbach, G. G., D. W. Schemske, H. V. Cornell, A. P. Allen, J. M. Brown, M. B. Bush, S. P. Harrison, A. H. Hurlbert, N. Knowlton, H. A. Lessios, C. M. McCain, A. R. McCune, L. A. McDade, M. A. McPeek, T. J. Near, T. D. Price, R. E. Ricklefs, K. Roy, D. F. Sax, D. Schluter, J. M. Sobel, and M. Turelli. 2007. Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. Ecol. Lett. 10:315–331.

Mittell, E. A., S. Nakagawa, and J. D. Hadfield. 2015. Are molecular markers useful predictors of adaptive potential? Ecol. Lett. 18:772–778.

Nabholz, B., J. F. Mauffrey, E. Bazin, N. Galtier, and S. Glemin. 2008. Determination of mitochondrial genetic diversity in mammals. Genetics 178:351–361.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U. S. A. 70:3321–3323.

Norton, B. A., K. L. Evans, and P. H. Warren. 2016. Urban biodiversity and landscape ecology: patterns, processes and planning. Curr. Landsc. Ecol. Reports 1:178–192. Current Landscape Ecology Reports.

Oliver, T. H., M. S. Heard, N. J. B. Isaac, D. B. Roy, D. Procter, F. Eigenbrod, R. Freckleton, A. Hector, C. D. L. Orme, O. L. Petchey, V. Proença, D. Raffaelli, K. B. Suttle, G. M. Mace, B. Martín-López, B. A. Woodcock, and J. M. Bullock. 2015. Biodiversity and resilience of ecosystem functions. Trends Ecol. Evol. 30:673–684.

Palumbi, S. R. 2001. Humans as the world's greatest evolutionary force. Science (80-.). 293:1786–1790.

Pontarp, M., L. Bunnefeld, J. S. Cabral, R. S. Etienne, S. A. Fritz, R. Gillespie, C. H. Graham, O. Hagen, F. Hartig, S. Huang, R. Jansson, O. Maliet, T. Münkemüller, L. Pellissier, T. F. Rangel, D. Storch, T. Wiegand, and A. H. Hurlbert. 2019. The latitudinal diversity gradient: novel understanding through mechanistic eco-evolutionary models. Trends Ecol. Evol. 34:211–223.

Pontarp, M., and J. J. Wiens. 2017. The origin of species richness patterns along environmental gradients: uniting explanations based on time, diversification rate and carrying capacity. J. Biogeogr. 44:722–735.

R Core Team. 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Rabosky, D. L., and A. H. Hurlbert. 2015. Species richness at continental scales is dominated by ecological limits. Am. Nat. 185:572–583.

Romiguier, J., P. Gayral, M. Ballenghien, A. Bernard, V. Cahais, A. Chenuil, Y. Chiari, R. Dernat, L. Duret, N. Faivre, E. Loire, J. M. Lourenco, B. Nabholz, C. Roux, G. Tsagkogeorga, A. A. T. Weber, L. A. Weinert, K. Belkhir, N. Bierne, S. Glémin, and N. Galtier. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature 515:261–263.

662 Schmidt, C., M. Domaratzki, R. P. Kinnunen, J. Bowman, and C. J. Garroway. 2020a. Continent-
663 wide effects of urbanization on bird and mammal genetic diversity. Proc. R. Soc. B Biol.
664 Sci. 287:20192497.

665 Schmidt, C., M. Domaratzki, R. P. Kinnunen, J. Bowman, and C. J. Garroway. 2020b. Data
666 from: Continent-wide effects of urbanization on bird and mammal genetic diversity. Dryad
667 Data Repository.

668 Schmidt, C., and C. J. Garroway. 2021. The conservation utility of mitochondrial genetic
669 diversity in macrogenetic research. Conserv. Genet. 22:323–327.

670 Schmidt, C., J. Munshi-South, and C. J. Garroway. 2021. Determinants of genetic diversity and
671 species richness of North American amphibians. bioRxiv 1–26.

672 Shipley, B. 2016. Cause and correlation in biology. 2nd ed. Cambridge University Press,
673 Cambridge.

674 Simpson, G. G. 1964. Species density of North American recent mammals. Syst. Zool. 13:57–73.

675 Socolar, J. B., J. J. Gilroy, W. E. Kunin, and D. P. Edwards. 2016. How Should Beta-Diversity
676 Inform Biodiversity Conservation? Trends Ecol. Evol. 31:67–80. Elsevier Ltd.

677 Sokal, R. R., and N. L. Oden. 1978. Spatial autocorrelation in biology 1. Methodology. Biol. J.
678 Linn. Soc. 10:199–228.

679 Stein, A., K. Gerstner, and H. Kreft. 2014. Environmental heterogeneity as a universal driver of
680 species richness across taxa, biomes and spatial scales. Ecol. Lett. 17:866–880.

681 Storch, D., E. Bohdalková, and J. Okie. 2018. The more-individuals hypothesis revisited: the role
682 of community abundance in species richness regulation and the productivity–diversity
683 relationship. Ecol. Lett. 21:920–937.

684 Theodoridis, S., D. A. Fordham, S. C. Brown, S. Li, C. Rahbek, and D. Nogues-Bravo. 2020.
685 Evolutionary history and past climate change shape the distribution of genetic diversity in
686 terrestrial mammals. Nat. Commun. 11:2557.

687 Trabucco, A., and R. Zomer. 2019. Global Aridity Index and Potential Evapotranspiration (ET0)
688 Climate Database v2. , doi: 10.6084/m9.figshare.7504448.v3.

689 Vellend, M. 2005. Species diversity and genetic diversity: parallel processes and correlated
690 patterns. Am. Nat. 166:199–215.

691 Weir, B. S., and J. Goudet. 2017. A unified characterization of population structure. Genetics
692 206:2085–2103.

693 Worm, B., and D. P. Tittensor. 2018. A theory of global biodiversity. Princeton University Press,
694 Princeton, New Jersey.

695 Wright, D. H. 1983. Species-Energy Theory: An Extension of Species-Area Theory. Oikos
696 41:496–506.

697 WWF. 2018. Living Planet Report - 2018: Aiming higher. WWF, Gland, Switzerland.

698 Yates, M. C., E. Bowles, and D. J. Fraser. 2019. Small population size and low genomic

699 diversity have no effect on fitness in experimental translocations of a wild fish. Proc. R.
700 Soc. B Biol. Sci. 286.

701

Table 1. Data summary. Summary of aggregated raw genetic data: mean gene diversity, mean number of loci, median number of individuals at sites per species. Species mass (kg); species richness = mean species richness at sites; energy = mean potential evapotranspiration across species' ranges (mm/yr), heterogeneity = mean land cover diversity (Simpson's Index) within a 5,000 km^2 zone around a site; human population = mean human population density across sites. Ranges of values are given in parentheses for species with multiple sample sites.

Species (# sites)	Gene diversity	# loci	# Individuals	Mass	Species richness	Energy	Heterogeneity	Human population
Alces alces (2)	0.47 (0.43–0.51)	10	44.5 (40–49)	481	34.5 (34–35)	694.5	0.81 (0.78–0.84)	134.79 (1.45–268.14)
Antilocapra americana (1)	0.67	19	175	46.9	63	1615.39	0.73	3.42
Bison bison (8)	0.47 (0.43–0.51)	29	26.5 (7–31)	620	47.62 (32–69)	745.94	0.68 (0.4–0.82)	1.3 (0.52–3.42)
Canis latrans (41)	0.77 (0.69–0.82)	10	7 (5–10)	12	48.39 (40–55)	1039.97	0.7 (0.27–0.85)	134.92 (1.54–1463.37)
Canis lupus (1)	0.66	12	62	35	44	642.5	0.78	5.85
Glaucomys volans (8)	0.75 (0.65–0.81)	7	18.5 (6–120)	0.07	49.12 (47–51)	1287.09	0.68 (0.58–0.8)	7.31 (1.04–33.59)
Lasionycteris noctivagans (1)	0.83	18	87	0.01	53	1252.72	0.71	2.4
Lasiurus cinereus (1)	0.88	19	132	0.03	53	1126.83	0.71	2.4
Leopardus pardalis (2)	0.47 (0.36–0.58)	10	35 (28–42)	11.9	51 (49–53)	2287.6	0.77 (0.76–0.77)	27.12 (18.73–35.51)
Lepus americanus (39)	0.66 (0.48–0.76)	8	15 (7–100)	1.57	50.44 (2–73)	750.2	0.65 (0.27–0.81)	79.95 (0.69–2711.29)
Lynx canadensis (33)	0.72 (0.46–0.75)	14.15 (14–15)	26 (13–328)	9.77	43.82 (6–50)	655.47	0.73 (0.61–0.81)	4.59 (1.04–56.48)
Lynx rufus (65)	0.73 (0.56–0.79)	14.37 (9–17)	27 (7–141)	6.39	55.6 (33–81)	1459.73	0.62 (0.19–0.83)	189.69 (1.04–3540.4)
Martes americana (29)	0.63 (0.54–0.67)	12	22 (11–47)	0.88	45.86 (42–50)	709.47	0.73 (0.65–0.79)	1.21 (0.69–2.27)
Mephitis mephitis (1)	0.81	9	345	2.4	53	1230.39	0.73	34.19
Microdipodops megacephalus (3)	0.78 (0.73–0.82)	11	62 (49–69)	0.01	64.33 (59–70)	1780.52	0.45 (0.37–0.56)	1.04 (1.04–1.04)
Microdipodops pallidus (2)	0.73 (0.73–0.73)	10	52.5 (42–63)	0.01	59 (58–60)	2045.67	0.25 (0.24–0.27)	1.04 (1.04–1.04)
Myotis lucifugus (65)	0.83 (0.72–0.9)	9.2 (8–11)	33 (11–167)	0.01	44.45 (32–66)	971.21	0.75 (0.21–0.86)	25.96 (0–493.88)
Myotis septentrionalis (15)	0.87 (0.85–0.88)	5	54 (34–110)	0.01	42.13 (41–43)	952.73	0.82 (0.74–0.85)	16.05 (1.04–123.03)
Odocoileus hemionus (67)	0.62 (0.2–0.72)	10.55 (10–18)	29 (7–262)	83.8	57.72 (2–87)	1356.11	0.61 (0.08–0.84)	36.97 (0.83–1213.54)
Odocoileus virginianus (64)	0.81 (0.76–0.84)	14	32.5 (10–79)	75	48.95 (46–52)	1243.86	0.55 (0.32–0.74)	57.42 (6.13–351.21)
Oreamnos americanus (1)	0.52	22	102	71.3	37	743.39	0.79	2.06
Otospermophilus beecheyi (3)	0.75 (0.72–0.78)	11	61 (40–104)	0.6	64.33 (57–69)	1634.31	0.6 (0.54–0.71)	6.18 (3.08–9.03)
Ovis canadensis (16)	0.61 (0.48–0.67)	40.12 (16–210)	42.5 (10–276)	74.6	49.5 (44–63)	1553.5	0.38 (0.15–0.73)	1.27 (1.04–3.42)
Pekania pennanti (34)	0.62 (0.52–0.66)	16	20.5 (7–48)	3.75	45.56 (3–51)	808.85	0.65 (0.25–0.8)	130.82 (0–2620)

Peromyscus leucopus (36)	0.82 (0.75–0.87)	13.19 (10–18)	20 (5–38)	0.02	44.83 (12–53)	1475.67	0.63 (0.3–0.83)	2412.04 (4.86–10000)
Peromyscus maniculatus (10)	0.77 (0.75–0.8)	10.9 (10–11)	12.5 (6–31)	0.02	48.3 (47–51)	1139.04	0.75 (0.62–0.83)	9.44 (1.04–62.41)
Procyon lotor (1)	0.84	10	330	6.37	51	1330.93	0.69	68.91
Puma concolor (13)	0.48 (0.33–0.58)	33.62 (10–46)	51 (21–739)	53.9	61.77 (50–82)	1561.58	0.64 (0.43–0.85)	104.17 (1.04–691.74)
Rangifer tarandus (82)	0.77 (0.45–0.87)	16.96 (14–21)	24 (5–283)	108	40.26 (5–63)	546.54	0.65 (0.28–0.83)	1.11 (0.69–5.11)
Sylvilagus transitionalis (3)	0.42 (0.32–0.47)	10	48 (6–103)	0.81	50.33 (50–51)	1070.02	0.75 (0.7–0.8)	161.12 (43.25–244.49)
Tamiasciurus douglasii (14)	0.65 (0.58–0.72)	9	10.5 (7–24)	0.22	65 (54–70)	1237.95	0.6 (0.53–0.82)	7.76 (1.04–93)
Tamiasciurus hudsonicus (12)	0.66 (0.48–0.77)	9	11 (5–48)	0.2	62.58 (54–68)	804.72	0.61 (0.49–0.73)	1.71 (1.04–5.6)
Taxidea taxus (12)	0.73 (0.43–0.82)	17.33 (12–20)	34.5 (19–649)	7.84	49.25 (39–56)	1403.56	0.59 (0.25–0.81)	44.68 (1.04–297.98)
Ursus americanus (43)	0.72 (0.32–0.82)	15 (8–20)	18 (5–2444)	111	51.35 (0–86)	791.39	0.7 (0.23–0.84)	50.31 (0.52–1065.89)
Ursus arctos (19)	0.67 (0.51–0.77)	9.89 (8–20)	48 (14–729)	196	36.68 (1–68)	568.29	0.6 (0.22–0.75)	1.25 (1.04–5.11)
Ursus maritimus (35)	0.69 (0.56–0.8)	15.09 (8–24)	31 (6–1050)	375	9.17 (2–33)	369.94	0.55 (0.27–0.76)	0.82 (0–1.04)
Vulpes lagopus (3)	0.72 (0.68–0.78)	9	28 (8–42)	3.6	25 (22–27)	455.51	0.73 (0.72–0.75)	1.04 (1.04–1.04)
Vulpes vulpes (16)	0.65 (0.53–0.74)	11.44 (8–13)	30 (9–116)	4.84	44.81 (24–54)	836.89	0.71 (0.39–0.83)	882.65 (1.04–3667.27)

707

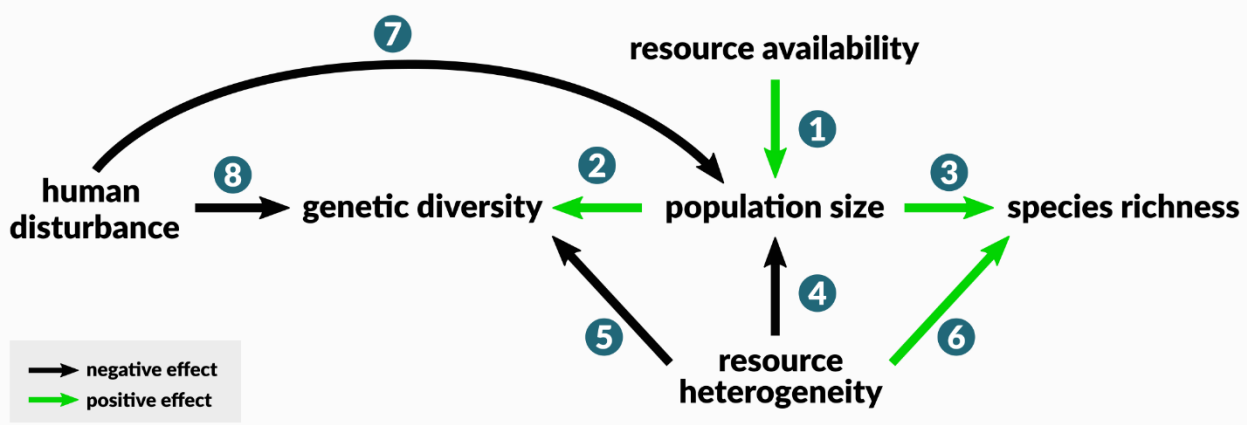


Fig. 1. Framework integrating genetic diversity into ecological limits hypotheses. We focus on two major ecological limits pathways which stem from resource availability (the more individuals hypothesis) and resource heterogeneity. Under the more individuals hypothesis, resource availability across a species range positively affects species' long-term effective population sizes (1). Nuclear genetic diversity increases with the effective population size (2). If population size is regulated by resource availability, it will be positively associated with community size and thus species richness (3). In heterogeneous environments, populations and species specialize to different niches but have smaller population sizes (4). Specialization reduces gene flow (5) and enhances species' ability to coexist (6). Human land transformation reduces wildlife habitat (7) and gene flow among populations (8).

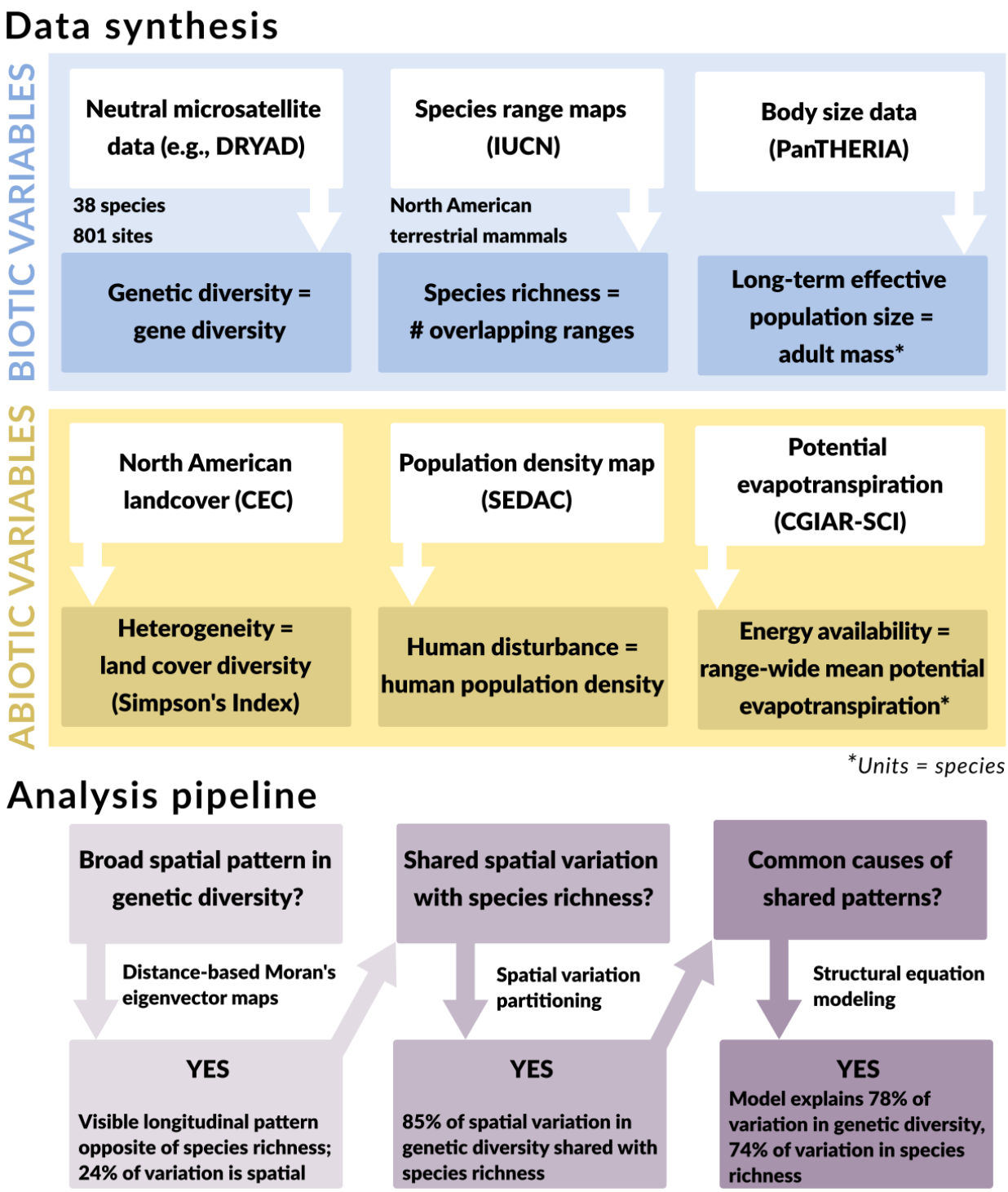
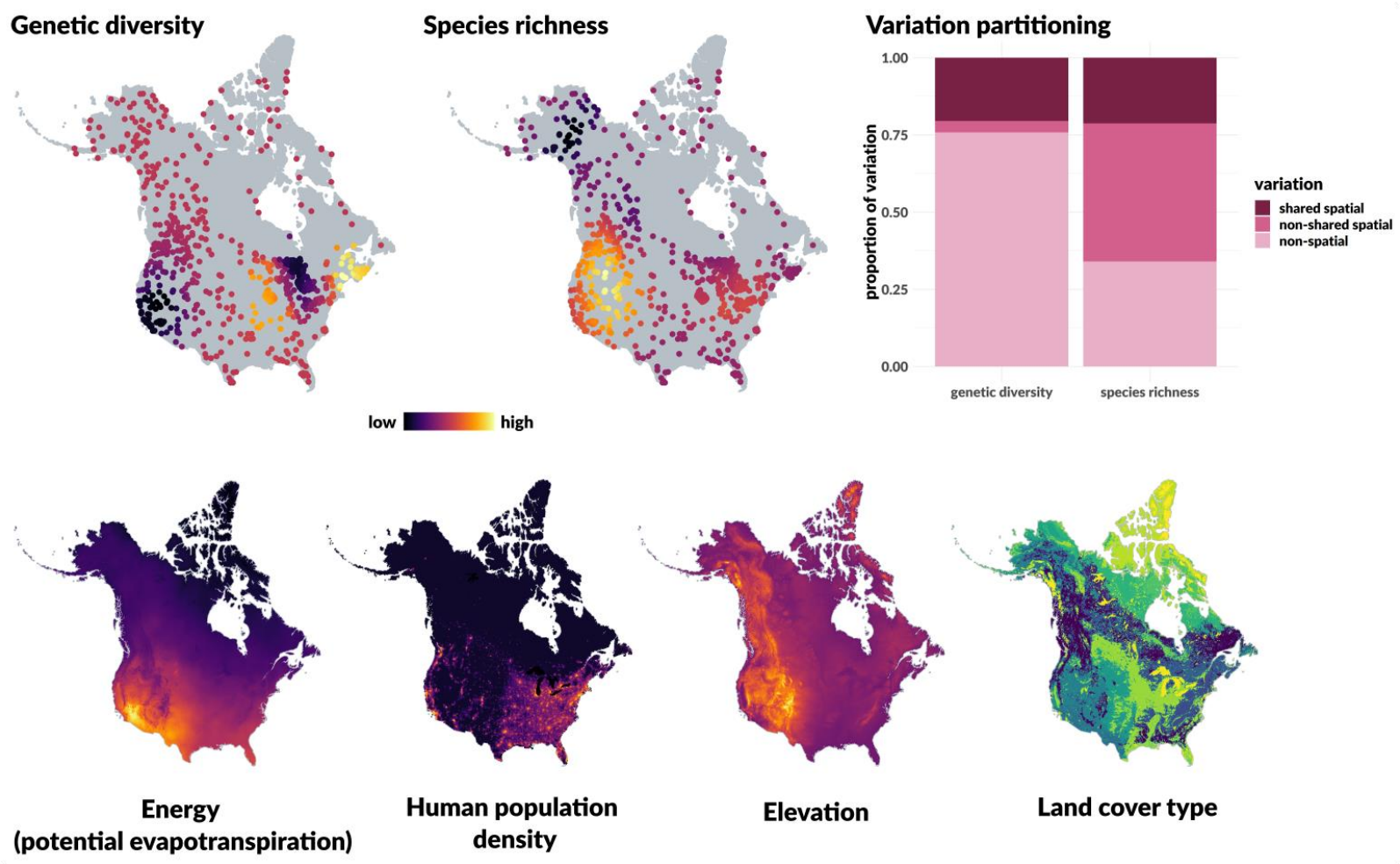


Fig. 2. Methodological workflow detailing data sources and our series of analyses. For structural equation models, variables were either measured at each of 801 sample sites for which genetic diversity data was available, or at the species level (n = 38 species).



724

725 **Fig. 3.** Spatial patterns of biodiversity and environmental factors. *(Top row)* Locations of 801 North American mammal populations
726 for which raw microsatellite data was available in public repositories. Point color indicates predicted values of genetic diversity and
727 species richness based on spatial patterns detected in the data. The variation partitioning plot shows the proportion of variation in
728 genetic diversity and species richness which can be explained by spatial factors. Spatial variation is further broken down into shared

and non-shared spatial variation. (*Bottom row*) Major environmental features of North America. Note land cover is categorical and colors represent different types. Elevation is shown for reference, but was not included in our models.

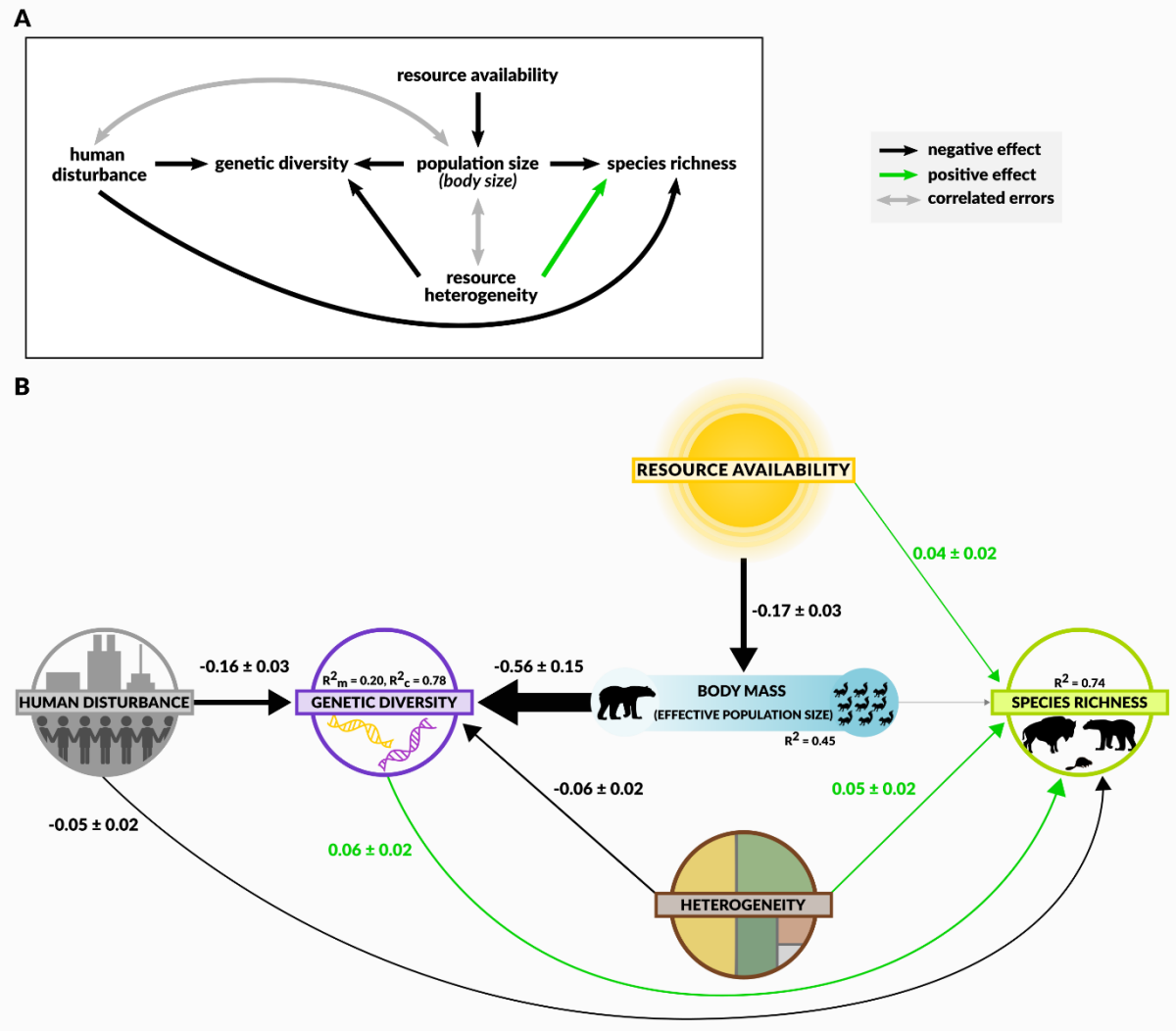


Fig. 4. Structural equation models. (a) Our conceptual hypothesis network, modified from Figure 1 to accommodate variables measured at species and site levels. Single-headed arrows represent unidirectional relationships between variables. Grey double-headed arrows indicate variables with correlated errors that were excluded from model evaluation (see Methods). (b) Structural equation model results. Green and black lines are positive and negative relationships, respectively, and the grey line is an unsupported link. Line widths reflect partial regression coefficients, which are listed for each path with standard errors. R^2 values are the amount of variation explained for each response variable. Genetic diversity was fit with a random effect for species: R^2_m is the variation explained by fixed effects only, and R^2_c is the variation explained by fixed and random effects.

Supplementary Information for: Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity

Includes:
Figures S1-S3
Tables S1-S5

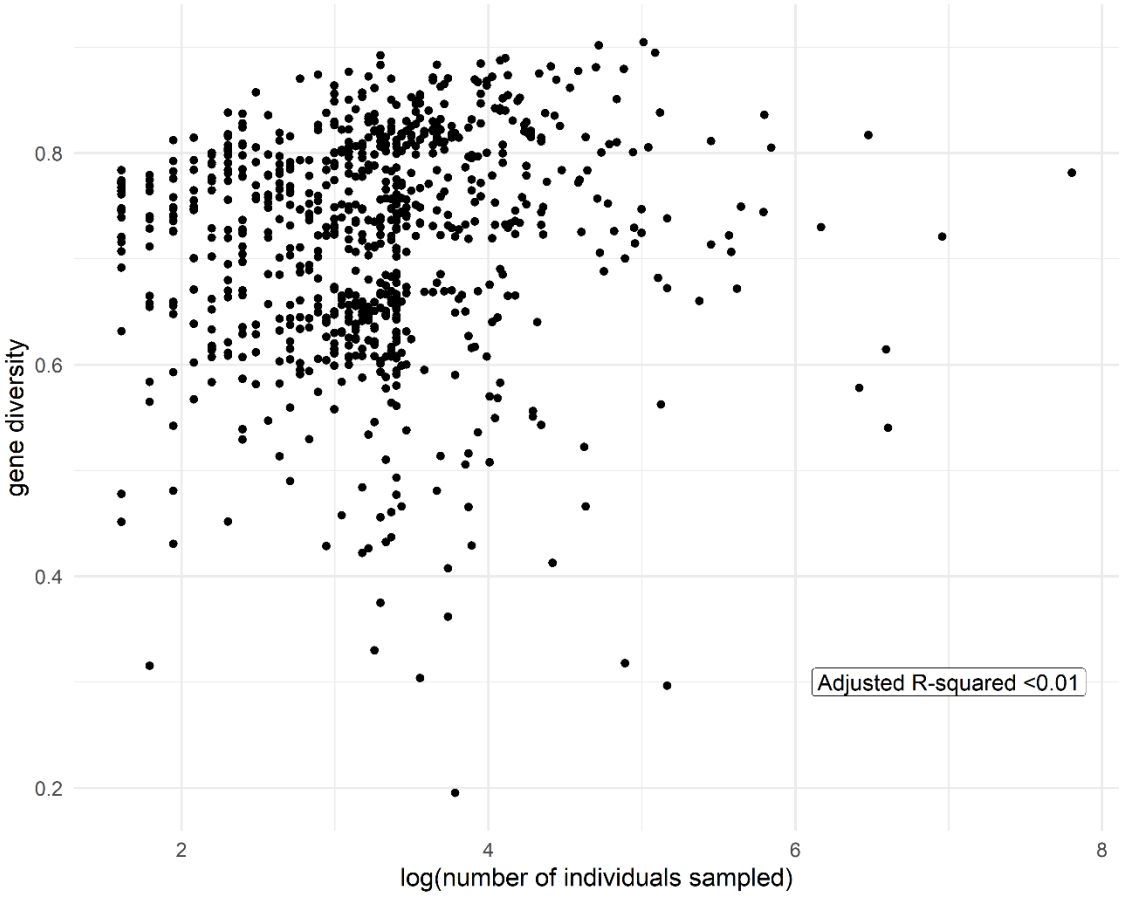


Figure S1. Plot of gene diversity vs. sample size. Gene diversity as a metric of genetic diversity depends on allele frequencies and is minimally affected by sample size. Larger populations have more rare alleles, which contribute little to gene diversity.

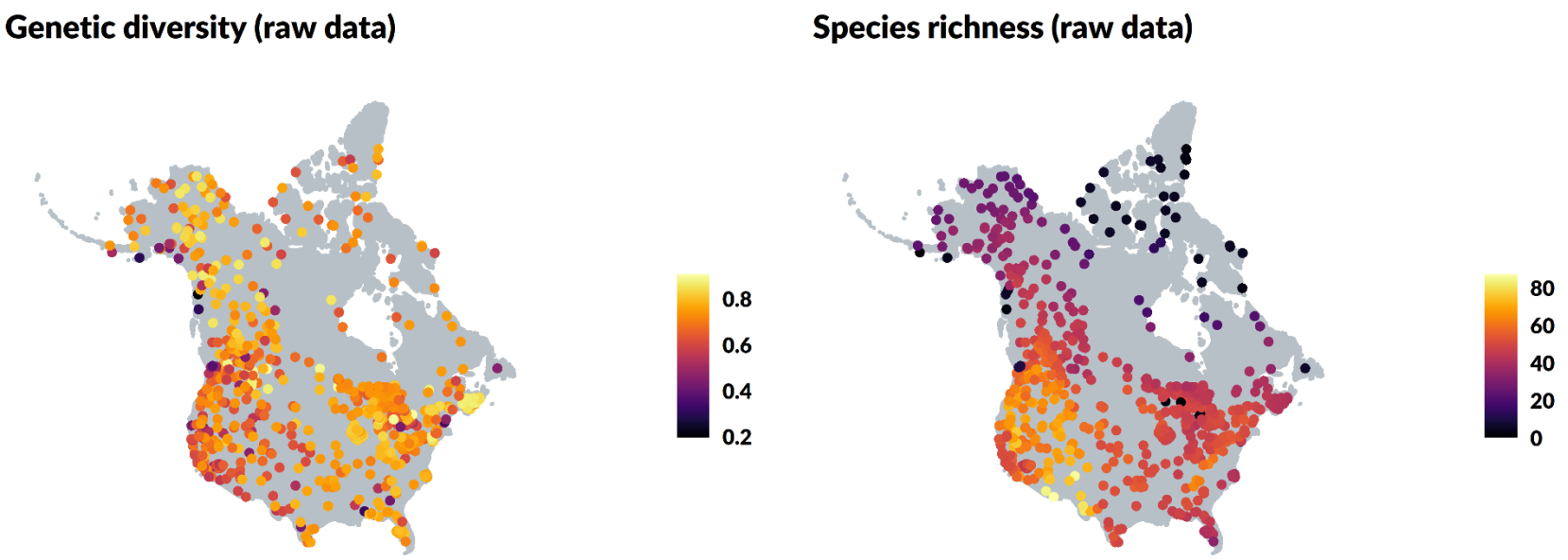


Figure S2. Maps of raw genetic diversity (gene diversity) and species richness data for each site. Clear gradients are visible in species richness but not genetic diversity. This is because most (65%) variation in species richness was spatial, in contrast to genetic diversity where comparatively less (24%) variation was spatially structured (Fig. 3).

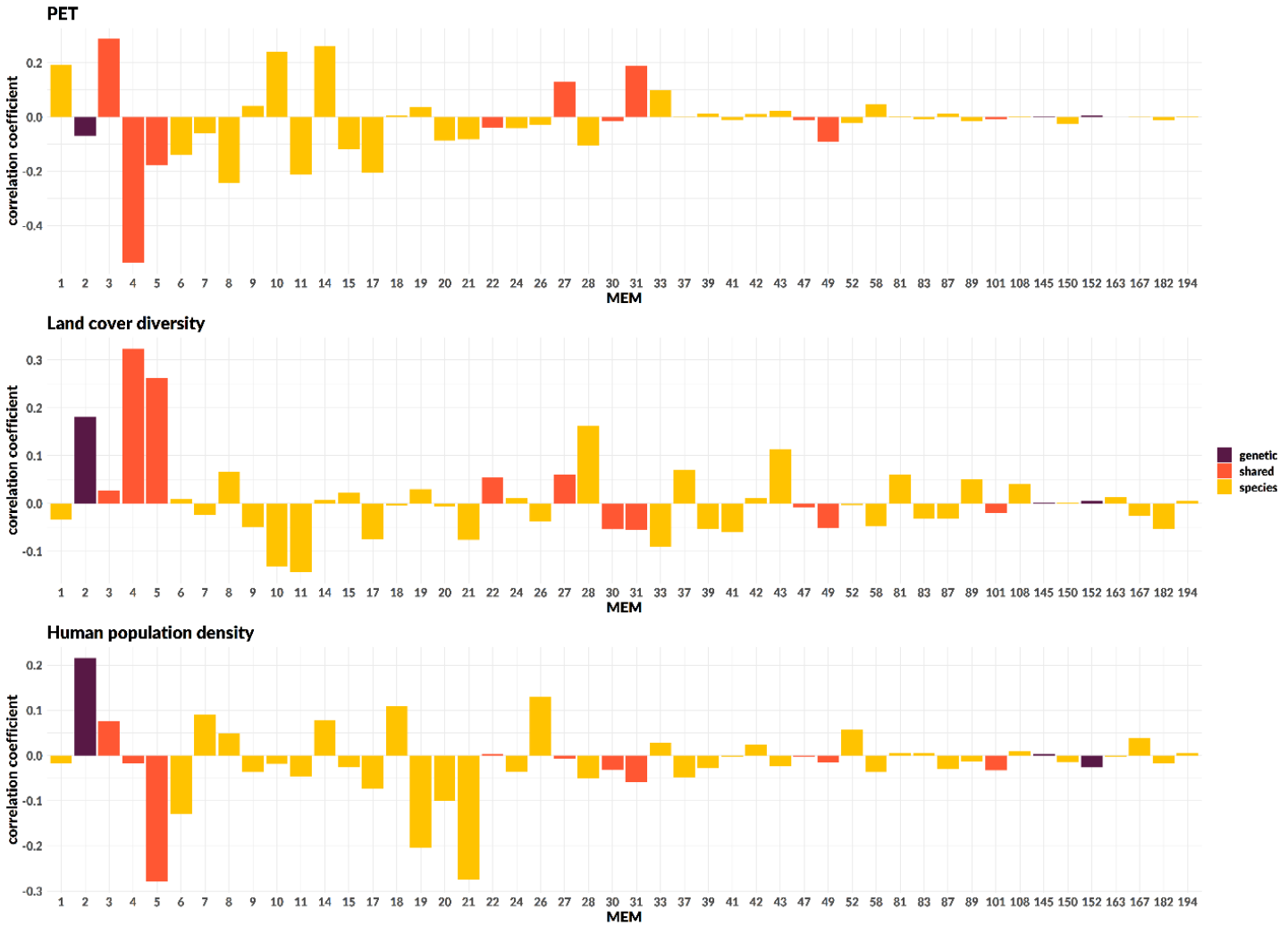


Figure S3. Correlation coefficients for spatial patterns (MEMs) and environmental variables measured at the site level: potential evapotranspiration (PET), land cover diversity, and human population density. MEMs describe spatial patterns in genetic diversity, species richness, or both (shared spatial patterns). MEMs are ordered from broad (MEM1) to fine scale (MEM194) patterns. Strong correlations indicate that environmental variables included in structural equation models account for broad scale spatial patterns present in genetic diversity and species richness.

Table S1. Path coefficients and standard errors for structural equation model presented in main text (Fisher's C = 2.25, p = 0.33, 2 degrees of freedom). Heterogeneity is measured within 5000 km^2 of a site.

Response	Predictor	Estimate ± SE
genetic diversity	human population density	-0.16 ± 0.03
genetic diversity	body size	-0.56 ± 0.15
genetic diversity	heterogeneity	-0.06 ± 0.02
body size	PET	-0.17 ± 0.03
species richness	body size	0.02 ± 0.02
species richness	heterogeneity	0.05 ± 0.02
species richness	human population density	-0.05 ± 0.02
species richness	PET	0.04 ± 0.02
species richness	genetic diversity	0.06 ± 0.02
Correlated errors		**Partial correlation coefficient**
body size	human population density	0.10*
body size	heterogeneity	-0.02

Table S2. Path coefficients and standard errors for structural equation model; heterogeneity is measured within 20000 km^2 of a site (Fisher's C = 2.09, p = 0.35, 2 degrees of freedom).

Response	Predictor	Estimate ± SE
genetic diversity	human population density	-0.16 ± 0.03
genetic diversity	body size	-0.57 ± 0.15
genetic diversity	heterogeneity	-0.04 ± 0.03
body size	PET	-0.17 ± 0.03
species richness	body size	0.02 ± 0.02
species richness	heterogeneity	0.06 ± 0.02
species richness	human population density	-0.05 ± 0.02
species richness	PET	0.05 ± 0.02
species richness	genetic diversity	0.06 ± 0.02
Correlated errors		**Partial correlation coefficient**
body size	human population density	0.10*
body size	heterogeneity	-0.03

Table S3. Path coefficients and standard errors for structural equation model; heterogeneity is measured within 50000 km^2 of a site (Fisher's C = 1.78, p = 0.41, 2 degrees of freedom).

Response	Predictor	Estimate ± SE
genetic diversity	human population density	-0.17 ± 0.03
genetic diversity	body size	-0.57 ± 0.15
genetic diversity	heterogeneity	0.00 ± 0.03
body size	PET	-0.17 ± 0.03
species richness	body size	0.02 ± 0.02
species richness	heterogeneity	0.06 ± 0.02
species richness	human population density	-0.06 ± 0.02
species richness	PET	0.05 ± 0.02
species richness	genetic diversity	0.06 ± 0.02
Correlated errors		**Partial correlation coefficient**
body size	human population density	0.10*
body size	heterogeneity	-0.03

12 **Table S4.** Path coefficients and standard errors for structural equation model; heterogeneity is
13 measured within 100000 km^2 of a site (Fisher's C = 1.54, p = 0.46, 2 degrees of freedom).

Response	Predictor	Estimate ± SE
genetic diversity	human population density	-0.18 ± 0.03
genetic diversity	body size	-0.56 ± 0.15
genetic diversity	heterogeneity	0.04 ± 0.03
body size	PET	-0.17 ± 0.03
species richness	body size	0.02 ± 0.02
species richness	heterogeneity	0.05 ± 0.02
species richness	human population density	-0.05 ± 0.02
species richness	PET	0.04 ± 0.02
species richness	genetic diversity	0.06 ± 0.02

Correlated errors		Partial correlation coefficient
body size	human population density	0.10*
body size	heterogeneity	-0.03

14

15

Table S5. Path coefficients and standard errors for structural equation model using actual evapotranspiration (AET) as a measure of resource availability. Heterogeneity is measured within 5000 km^2 of a site (Fisher's C = 0.53, p = 0.77, 2 degrees of freedom). Genetic diversity R^2_m = 0.20; R^2_c = 0.78; body size R^2 = 0.48; species richness R^2 = 0.74.

Response	Predictor	Estimate ± SE
genetic diversity	human population density	-0.16 ± 0.03
genetic diversity	body size	-0.56 ± 0.15
genetic diversity	heterogeneity	-0.06 ± 0.02
body size	AET	-0.33 ± 0.03
species richness	body size	0.03 ± 0.02
species richness	heterogeneity	0.04 ± 0.02
species richness	human population density	-0.05 ± 0.02
species richness	AET	0.05 ± 0.02
species richness	genetic diversity	0.05 ± 0.02
Correlated errors		**Partial correlation coefficient**
body size	human population density	0.09*
body size	heterogeneity	-0.01