Accepted Article

Accounting for the nested nature of genetic variation across levels of organization improves our understanding of biodiversity and community ecology

Quentin D. Read[1], Sean M. Hoban[2], Maarten B. Eppinga[3], Jennifer A. Schweitzer[1] and Joseph K. Bailey[1]

[1]*Dept of Ecology and Evolutionary Biology, Univ. of Tennessee, Knoxville, TN 37996, USA. QDR also at: Rocky Mountain Biological Laboratory, PO Box 519, Crested Butte, CO 81224, USA.*

[2]*National Inst. of Mathematical and Biological Synthesis, Knoxville, TN 37996, USA.*

[3]*Dept of Environmental Sciences, Copernicus Inst. of Sustainable Development, Utrecht University, PO Box 80115, NL-3508 TC Utrecht, the Netherlands.*

Corresponding author: Quentin D. Read, Dept of Ecology and Evolutionary Biology, Univ. of Tennessee, Knoxville, TN 37996, Tennessee, United States. E-mail: qread@utk.edu

Decision date: 24-Nov-2015

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1111/oik.02760].

'This article is protected by copyright. All rights reserved.'

Accepted Article

Abstract

Recent work has demonstrated that the presence or abundance of specific genotypes, populations, species and phylogenetic clades may influence community and ecosystem properties such as resilience or productivity. Many ecological studies, however, use simple linear models to test for such relationships, including species identity as the predictor variable and some measured trait or function as the response variable without accounting for the nestedness of genetic variation across levels of organization. This omission may lead to incorrect inference about which source of variation influences community and ecosystem properties. Here, we explicitly compare this common approach to alternative ways of modeling variation in trait data, using simulated trait data and empirical results of common-garden trials using multiple levels of genetic variation within *Eucalyptus*, *Populus*, and *Picea*. We show that: 1) when nested variation is ignored, an incorrect conclusion of species effect is drawn in up to 20% of cases; 2) overestimation of the species effect increases – up to 60% in some scenarios – as the nested term explains more of the variation; and 3) the sample sizes needed to overcome these potential problems associated with aggregating nested hierarchical variation may be impractically large. In common-garden trials, incorporating nested models increased explanatory power twofold for mammal browsing rate in *Eucalyptus*, threefold for leaf area in *Populus*, and tenfold for branch number in *Picea*. Thoroughly measuring intraspecific variation and characterizing hierarchical genetic variation beyond the species level has implications for developing more robust theory in community ecology, managing invaded natural systems, and improving inference in biodiversity–ecosystem functioning research.

'This article is protected by copyright. All rights reserved.'

Accepted Article

(Introduction)

Biodiversity is inherently nested. Genetic variation is differentially expressed within sub-populations and populations, within and among species, and, most broadly, across phylogenetic clades. At any level of biological organization, this variation may be expressed in traits that may in turn influence community structure or ecosystem properties. For example, two recent experimental studies showed that population-level genetic differentiation drives community structure in natural arthropod communities (Barbour et al. 2009), and that trait variation among species and phylogenetic clades influence a diversity-stability relationship in the overall accumulation of plant biomass in artificial communities (Cadotte et al. 2012). Understanding the relationship between genetic diversity, species diversity, and ecosystem properties has been fundamental to basic research linking ecological and evolutionary disciplines for decades (Tilman et al. 1997, Hooper et al. 2005, Fargione et al. 2007), including the growing field of community and ecosystem genetics (Whitham et al. 2006, Hughes et al. 2008, Genung et al. 2012, Bailey et al. 2014). Recently, studies in eco-evolutionary dynamics have shown experimentally that genotype changes in response to selective environmental pressures often have ecosystem consequences (Turcotte et al. 2013, Hiltunen et al. 2014). This knowledge is critical for predicting the consequences of climate change (Singer and Thomas 1996, Savolainen et al. 2007, Visser 2008), assessing the effects of exotic species (Hobbs and Huenneke 1992, Levine and D'Antonio 1999, Levine et al. 2003, Fargione and Tilman 2005), and boosting food production, among other ecosystem services (Hoehn et al. 2008; Feld et al. 2009; de Bello et al. 2010; Cardinale et al. 2012). Correct inference regarding the source of genetic and trait variation, and its consequences, is critical for making predictions grounded in theory about how biodiversity changes will affect ecosystem functioning.

Accounting for the nested nature of biodiversity (e.g., that genotypes are nested within populations, populations exist within species, and species are nested within phylogenies) is rarely undertaken in studies measuring the functional consequences of biodiversity (but see Cook-Patton et al. 2011; Schöb et al. 2015). Without considering variation in a trait at a lower level (e.g., genotype), estimates of differences at a higher level (e.g., species) may potentially be inflated. Similarly, examining trait differences among phylogenetic clades without accounting for variation among species would yield similarly increased estimates of effect size. In both cases, this is simply because only considering higher-order variation (i.e., species or phylogenetic clade, respectively) means that much of the variation among individuals is assigned incorrectly, either to higher-level differences or to residual variation (i.e., variation among species may simply be due to variation among genotypes or populations and not reflect true

'This article is protected by copyright. All rights reserved.'

Accepted Article

differences among species). Despite the potential for inflated effect sizes and incorrect inference, estimates of differences in traits or functions among populations, or species, in a range of contexts routinely do not account for the nested nature of biodiversity. For example, many landmark studies in biodiversity-ecosystem functioning relationships have treated species richness as synonymous with diversity, neglecting any genotypically driven or within-species variation (e.g., Tilman *et al.* 1997; Caliman *et al.* 2013; Isbell *et al.* 2013). One alternative in biodiversity-ecosystem functioning experiments is to manipulate genotype identity or diversity within species, as several recent experimental studies have done (Hiltunen *et al.* 2014; Rudman *et al.* 2015). However, this labor-intensive approach is admittedly prohibitive in some systems. In such cases, we recommend either indirectly addressing intraspecific variation with a functional trait approach (Díaz *et al.* 2007) or, at the minimum, explicitly stating that taking into account variation at lower levels could change the magnitude or significance of the findings.

Trait distributions of different species within an assemblage can overlap greatly; therefore, genotypic variation within species may represent the bulk of variation in traits that drives variation in ecosystem function (Crutsinger *et al.* 2006, 2014). Unfortunately, variation between genotypes is not captured in an aggregated estimate of species-level variation. Appropriate aggregation requires much larger sample sizes than would be expected from sampling a single distribution. Hence, stratifying trait samples only at the species level may lead to an increased chance of the investigator concluding that there is an effect of species identity on a response trait or ecosystem process. In many cases, the result is more properly attributed to variation at the within-species genetic level. If sampled individuals are only binned by species, the uneven sampling of variation at levels lower than species can lead to Type I error where trait variation is attributed solely to species identity. This phenomenon is not merely statistical nuance – this misleading and incorrect inference may have profound effects on how we interpret ecological patterns and how theory is developed. For example, Schöb *et al.* (2015) found that for barley plants and associated weed species that were planted in monoculture and mixture, the effect on overall productivity of increasing genotypic variation in barley was qualitatively different than the effect of increasing species richness of associated weeds. This result is different from what one would predict from a non-nested model. It suggests, as do other similar studies (e.g., Pante *et al.* 2015), that aggregating or averaging out intraspecific variation can lead investigators to overlook critical processes that structure natural systems.

An approach that goes beyond a narrow focus on species identity has important implications for extending the scope and power of numerous fields of ecological research. For example, in invasion biology, genetically based trait variation can explain variation in the success of invasive species

'This article is protected by copyright. All rights reserved.'

Accepted Article

(Lindholm *et al.* 2005; Roman & Darling 2007), as well as variation in the ability of resident communities to resist invaders (Crutsinger *et al.* 2008). More generally for community ecology, the inclusion of genetically based intraspecific variation is vital for increasing our understanding of the forces driving species coexistence. For example, individual-level trait diversity is necessary for species coexistence in temperate forests (Clark 2010), and underlies the relationship between environmental stress and the net balance between competition and facilitation among tree species (Coyle *et al.* 2014). Predictions of species responses to ongoing global change may also be made more accurate by incorporating genotypic and phenotypic trait variation within species, which has been shown in the context of trait variation within a single species (Anderson & Gezon 2015) as well as across species at the landscape level (Fitzpatrick & Keller 2015). Finally, meta-analyses (Cardinale *et al.* 2012) and reviews (Hooper *et al.* 2005) in the biodiversity-ecosystem functioning field have synthesized studies assessing the consequences of changes to biodiversity. However, diversity metrics that are partitioned more finely than at the species level have been largely neglected in the studies that are synthesized in these meta-analyses. If the published studies from which estimates of effect size in these meta-analyses are derived fail to account for the nested nature of biodiversity, the grand-mean effect size estimates may be inflated or incorrect (Borenstein *et al.* 2009).

Using simulated trait data and three empirical examples as case studies, we examine how inferences drawn from models including nested hierarchical variation differ from those drawn using models that do not incorporate this variation. We show that: (1) when hierarchically nested variation is ignored and only one level of variation, such as the species level, is considered, an incorrect conclusion of a species identity effect where one does not exist is drawn in up to 20% of cases (i.e., Type I error); (2) overestimation of the proportion of variation explained by species identity increases up to 60% under some scenarios; and (3) the sample sizes needed to overcome these potential problems associated with aggregating nested hierarchical variation may be too large to feasibly achieve, demonstrating the importance of partitioning variation among individuals at multiple levels of diversity. Such potentially high rates of type I error and inflated effect size suggest that taking the hierarchically nested nature of biodiversity into account is an important step in biodiversity research and should be considered in future theoretical and empirical studies.

Simulated Data Sets

Generation of simulated data. Our goal was to determine the statistical effect of nesting compared to an approach without nesting when a genetic hierarchy (i.e., genotypes nested within species) exists and is available to incorporate into ecological field or laboratory studies. Specifically, we examined variation in genotypes when comparing trait means in a system in which some variation in a continuous

'This article is protected by copyright. All rights reserved.'

Accepted Article

quantitative trait is due to among-genotype differences within a species and some is due to among-species differences. The hierarchical levels simulated need not necessarily be limited to genotype nested within species; our results could equally apply to any other level of genetic hierarchy, such as species nested within phylogenetic clades.

We simulated trait data as follows. First, we generated an allele matrix *M* with variable numbers of binary loci with a value of either 1 or -1 that together represent a simulated genotype for each individual within a species. In each replicate simulation, *M* is an $m \times n$ matrix, where *m* is the number of individuals in the simulated population and *n* is the number of genetic loci. The total additive genetic variation *G* for each individual was calculated by taking the cross product of the transposed matrix containing allele values and a vector with length *m* of normally distributed random variables *U* representing variation across genotypes: $G = M^T \cdot U;\ U \sim N(0, \sigma_g)$. This results in a vector of length *m*. Each individual was assigned a phenotype or trait value *Y* by adding a normally distributed environmental noise term *V* (also a vector of length *m*) to its genotype value: $Y = G + V;\ V \sim N(\mu_s,\ (1 - \sqrt{H^2}) / s_g)$; in this way, individuals of the same genotype may have differing phenotypes. For each replicate simulation run, we generated a new allele matrix *M* and vectors *U* and *V*. The heritability value H^2 represents the proportion of variation in the trait value due to variation in the genotype (i.e., broad-sense heritability). The simulation procedure is outlined in Figure 1.

All simulations consisted of two species. The parameters that were varied across simulations included the standard deviation σ_g of the genotype value, the mean trait value μ_s for each species, the heritability H^2, the number of individuals, and the number of binary loci. We simulated 100 datasets for each combination of parameters, including 10, 50, and 100 individuals per species; 1, 2, and 3 loci (yielding 2, 4, and 8 possible genotypes per species); heritability values $H^2 = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$; species means $\mu_1 = 0$ and $\mu_2 = \{0, 0.1, 0.25, 0.5, 1, 2, 5, 10\}$; and genotype standard deviation $\sigma_g = \{0.1, 0.25, 0.5, 1, 2, 5, 10\}$. In addition, to assess the influence of increasing the number of loci per individual while heritability is held constant, we conducted an additional simulation with 10 datasets for each of the following combinations of parameters: 100 individuals per species; each possible number of loci from 1 to 10; $H^2 = 0.5$; species means $\mu_1 = 0$ and $\mu_2 = \{0, 1, 5, 10\}$; and genotype standard deviation $\sigma_g = \{0.1, 1, 5, 10\}$. Data were simulated in the R 3.0.1 statistical programming language (R Core Team 2013).

We were additionally interested in assessing the sample size needed to appropriately aggregate trait differences at the species level while assuming different magnitudes of genotype-level variation. We considered the specific case in which two species had identical mean trait values, but each species

'This article is protected by copyright. All rights reserved.'

Accepted Article

consisted of eight genotypes that differed in trait means,, with heritability $H^2 = 0.5$, and genotype standard deviation $\sigma_g = \{0.1, 0.25, 0.5, 1, 2, 5, 10\}$. The more appropriate (i.e., larger) the sample size, the closer to zero the difference in estimated species means, $|\mu_1 - \mu_2|$, should be. We conducted this separate simulation with larger variation of sample sizes (10^1, 10^2, 10^3, 10^4, and 10^5 individuals per species). We generated 100 replicate datasets for each of the above combinations of sample size and genotype standard deviation. For each replicated dataset, we calculated the observed difference between the species-level sample means and compared them to the expected difference of zero, assuming a normal distribution, $N(0, \sigma_g)$, to estimate the standard error of the difference between means across datasets. The sample size simulation was carried out in the MATLAB v8.4 programming language (MathWorks, 2014).

Statistical analysis of simulated data. We used two different approaches to statistically model the variation in trait means within each data set: (1) non-nested linear models with the trait as response variable and species as a categorical fixed effect, a common method of analyzing trait data across a broad range of studies and sub-disciplines, and (2) nested mixed models with the trait as response variable, species as a categorical fixed effect, and genotype as a categorical random effect nested within species. It is important to note that we did not explicitly use species or genotype diversity as an effect in either model; we concerned ourselves only with the way individuals were binned when statistically modelling variation in trait values. The models were fit using the R language (R Core Team 2013); in particular, mixed models were fit with the *lme4* function in the *lmer* package (D. Bates *et al.* 2015) using the restricted maximum likelihood method.

The coefficient of determination (R^2) and Akaike's Information Criterion (AIC) were calculated for both types of model. We calculated the R^2 for nested mixed models using a variance decomposition technique (Nakagawa & Schielzeth 2013); we partitioned the total variance estimated by the model into the variance of the random effects and of the fixed effects, and the residual variance. The R^2 value is decomposed into two components: the marginal R^2, or the percent variation explained by fixed effects, is the variance of the fixed effects divided by the total variance (fixed + random + residual variance). The conditional R^2, accounting for both fixed and random effects, is the sum of the fixed and random effect variance divided by the total variance term. We compared the estimates of the proportion of variation explained by species in non-nested models and nested mixed models by subtracting the marginal R^2 of the mixed model from the R^2 value of the non-nested model. We calculated the variance components with the *rsquared.GLMM* function in the *MuMIn* package (Kamil Bartoń 2015). For both the non-nested and the mixed nested models, we calculated the *p*-value using a likelihood ratio test, comparing the full fitted

'This article is protected by copyright. All rights reserved.'

Accepted Article

model with a model only fitting an intercept for the non-nested models, and a model fitting the intercept and the random effect for the nested models (*lmtest* package; Achim Zeileis & Torsten Hothorn 2002).

We quantified how each parameter of the simulation influenced (1) the R^2 value associated with the species term in each model type, (2) the difference in AIC scores (hereafter ΔAIC) between the nested and non-nested models, and (3) the proportion of times that each model included a species coefficient significantly different from zero at $\alpha = 0.05$. Scripts from all of our analyses are included as supplementary information.

Influence of variance in parameters on inference drawn from simulated datasets. Fitting both non-nested and nested mixed models to our simulated datasets revealed that non-nested models uniformly overestimate the proportion of variation explained by species (ΔR^2, the R^2 of the non-nested model subtracted from the marginal R^2 of the nested model) if genotype-level variation is ignored, with overestimates of marginal R^2 ranging from 0 to 0.62. Overestimates of species-level variation decreased with increasing sample size and increased with heritability (Figure 2a), although ΔR^2 values varied most at the lowest sample size of 10 individuals, occasionally exceeding 0.6 in some simulation runs (Supplemental Figure S1). In addition, overestimates were highest when the variance in genotype means was similar to the difference between species means (i.e., both very small or both very large; Figure 2b). When either variance parameter was much larger than the other, the non-nested model estimated the proportion of variation due to species accurately, even at high heritability values (Table 1; Supplemental Figure S1).

In addition, despite the increased number of parameters required to model variation by genotype and the penalty assessed for model complexity by Akaike's Information Criterion, model selection most often chose the more complex model incorporating genotype; ΔAIC between non-nested and nested models was positive 83% of the time and increased with heritability (Figure 3). The ΔAIC increased with increasing sample sizes from 10 to 100 individuals, but did not systematically change with the absolute difference in species means or the among-genotype standard deviation (Table 1; Supplemental Figure S2). This indicates that a nested model is a more parsimonious fit to the data regardless of the differences in trait values.

Regardless of the actual proportion of variation explained, many studies draw biological inferences solely from the p-value of a null-hypothesis significance test. When the trait variation is modeled using only the fixed factor of species, an investigator would draw a different inference compared to the nested mixed model. When we evaluated the statistical significance of the likelihood ratio of each model fit at $\alpha = 0.05$, a significant effect of species on the trait mean was found much more often in the

'This article is protected by copyright. All rights reserved.'

Accepted Article

non-nested model than the nested model (44% of cases versus 27%, respectively, for all parameter combinations). As heritability increased, the proportion of discrepancies in null hypothesis rejection between the non-nested and nested mixed model increased, ranging from 14% at $H^2 = 0$ to 20% at $H^2 = 0.9$ (Figure 4a). As before, the discrepancies among models were most frequent when the species and genotype-level variance terms were either both small or both large; both statistical techniques generally agreed when one variance term was much larger than the other (Figures 4b and 4c). Finally, increasing sample size tended to cause models to support opposing inference more often, as increased statistical power meant that p-values for the non-nested model were more likely to drop below the significance threshold, all else being held constant (Table 1; Supplemental Figure S3).

The previous datasets were generated from simulated genotypes with few loci (1, 2, or 3); the additional simulations we ran with number of loci ranging from 1 to 10 showed that increasing the number of loci decreased the effect of accounting for nested variation on ecological inference (Table 1; Figure 5). Increasing the number of loci had little effect on the model selection criteria at low numbers of loci: across all combinations of parameter values, median ΔAIC was highest (1153.9) with 3 loci. At higher numbers of loci, median ΔAIC steadily decreased: the median value with 10 loci was 75.1 (Figure 5a). While this difference is still large enough to signify that a nested mixed model is more parsimonious than a non-nested linear model, the increase in explanatory power may not be great enough to justify the additional effort in some cases, as both the ΔR^2 value and proportion of false positive results both approached zero as number of loci increased (Figure 5b and 5c).

Finally, our simulation of observed difference in sample means with increasing sample size showed the pattern that as sample size increased, the effect size associated with species decreased (Figure 6). The absolute value of the mean effect size decreased despite no change in true difference in sample means as sample size increased. In our particular example, the spuriously high observed difference in species-level means was present even at relatively high levels of within-genotype variance, and did not disappear until impractically large sample sizes were assumed (Figure 6).

Empirical Data Sets

Common-garden studies. We used data from three previously conducted common-garden studies to illustrate similar patterns to the ones we observed in simulated datasets. Each data set represents different levels of hierarchically nested biological variation (for the first, family nested within species, and for the second two, genotype nested within locality), illustrating that the nestedness of variation—and potential pitfalls for inference—are independent of the particular level examined. The first, established in Tasmania, was a controlled cross trial of multiple *Eucalyptus* species in three clades within

'This article is protected by copyright. All rights reserved.'

Accepted Article

the subgenus *Symphyomyrtus*, with a total of 332 individuals measured across 85 families nested within 16 cross types (Gorman *et al.* 2014). The second study consisted of genotypes of cottonwood (*Populus angustifolia*) distributed in the arid Southwestern United States, with a total of 74 individuals measured across 37 genotypes nested within six localities along an elevation gradient (Whitham *et al.* 2006 and references therein). See the above two references for specific methods. Finally, in the third study, which consists of previously unpublished data, seeds from 10 red spruce (*Picea rubens*) genotypes from 6 different mountaintop populations were collected in Great Smoky Mountains National Park (TN/NC, USA). The seeds were cold-stratified, germinated, and potted in a randomized single-block design in a greenhouse. After one year, the seedlings were measured for a number of properties relating to plant architecture. Here, we present one response variable from each set of trials: level of mammal browsing from the *Eucalyptus* trials, leaf area measurements from the *Populus* trials, and number of branches from the *Picea* trials. Each of the traits we selected represents an easily measured property of a plant; the traits were chosen to demonstrate patterns of variation similar to those observed in the simulated data.

We analyzed each of the empirical data sets in the same way as the simulated data sets, modeling each one using a non-nested model with species or locality as a categorical fixed effect and a nested mixed model with family or genotype nested within species or locality (see above). We generated bootstrapped standard errors for each estimate of R^2 by drawing with replacement from the trait values until reaching the original sample size, calculating the marginal and conditional R^2 values, and repeating the procedure 9999 times (*boot* package; Angelo Canty & Brian Ripley 2015).

Application of different statistical models: **Eucalyptus.** A significant amount of variation in mammal browsing percentage in *Eucalyptus spp.* was explained by species in the non-nested ANOVA model. However, accounting for family in a mixed model reveals that the species term accounts for a relatively small proportion of the total variation explained by the nested mixed model (non-nested model $R^2 = 0.13$, mixed-model conditional $R^2 = 0.28$; Figure 7). This indicates an approximate twofold increase in explanatory power for the mammal browsing trait in *Eucalyptus*. Additionally, both the non-nested and the nested mixed model indicate that roughly the same proportion of variation is explained by the non-nested factor, species, across both models. Despite this, each model would lead to a different inference from a frequentist standpoint because the species effect was significant at $\alpha = 0.05$ in the non-nested model but not in the nested mixed model (non-nested model, $p = 0.00010$; nested mixed model, $p = 0.06$).

Application of different statistical models: **Populus.** A significant proportion of variation in leaf area was explained by locality of origin in the *Populus* common-garden trial. When genotype was accounted for in a nested mixed model, the explanatory power increased roughly threefold relative to the

'This article is protected by copyright. All rights reserved.'

Accepted Article

non-nested model (non-nested model, $R^2 = 0.25$; nested mixed model, conditional $R^2 = 0.70$; Figure 7). As in the previous case, although the non-nested and nested mixed models both apportion a similar amount of variation to the outermost level – in this case locality – the locality effect was not significant at $\alpha = 0.05$ in the nested mixed model (non-nested model, $p = 0.0086$; nested mixed model, $p = 0.22$).

Application of different statistical models: **Picea.** As in the previous examples, a significant proportion of variation was explained by locality of origin, but when genotype was also accounted for in a nested mixed model, explanatory power increased by an order of magnitude (non-nested model, $R^2 = 0.03$; nested mixed model, $R^2 = 0.24$; Figure 7). However, both would lead to a similar qualitative inference, albeit stronger for the non-nested model (non-nested model, $p = 2 \times 10^{-8}$; nested mixed model, $p = 0.02$). These three empirical results collectively suggest that the non-nested model leads to an incorrect inference in a traditional null hypothesis significance testing framework, as is commonly employed.

Discussion and Implications

Here we show that in a simulated system, a nested mixed model incorporating genotypic variation to partition variation in trait means is more appropriate than a non-nested model that does not nest other levels of genetic variation. The simulated system was constructed with no interactions among species or genotype and no environmental effects, which could either mask or amplify the biodiversity signal in natural systems. Specifically, the proportion of detected significant species effects was up to 20% greater in non-nested versus nested models. In other words, one would conclude that there are species effects when in fact those effects were related to lower-order levels of variation. Non-nested linear models resulted in effect sizes associated with the species term from 20% to 50% greater than the nested models; additionally, the model selection criterion showed that nested models were superior regardless of parameter values.

The results of our simulations demonstrate the need to employ a nested approach. Even if data are not available to do so explicitly, the binning of individuals into species should be acknowledged as one choice among multiple ways of partitioning variation among individuals, instead of being regarded as synonymous with diversity itself. By the same reasoning, genotype identity should be acknowledged as an important basis for inference in its own right. We also show the utility of the nested approach for analyzing data in field experiments. When fitting nested and non-nested models to empirical data taken from *Eucalyptus*, *Populus*, and *Picea* common-garden trials, the nested and non-nested models performed similarly in terms of estimating the proportion of total variation due to species, but the likelihood ratio test showed that the non-nested models would erroneously conclude the species effect to be significant in

'This article is protected by copyright. All rights reserved.'

Accepted Article

two of the three cases. Whether the system is experimental, natural, or simulated, a hierarchical nesting approach should help reduce the false attribution of variation and improve our understanding of the sources of variation among species or among genotypes. One caveat to our findings is that adding additional parameters to a model – in this case by including an additional nested level of genetic variation – will raise the proportion of variation explained, potentially leading to overfitting (Ginzburg & Jensen 2004). However, in our study, analysis of AIC values showed that nested models are much more parsimonious despite the additional parameters, relaxing the above caveat. The conclusions we draw here on the importance of nesting apply not only to comparisons of intraspecific and interspecific trait variation, but to any system with hierarchically nested variation, including at the genotype, population, species, or phylogenetic level.

Importantly, our conclusions on the importance of placing individuals in a properly nested hierarchy of bins are independent of the statistical approach chosen by the investigator (such as frequentist, Bayesian, or maximum-likelihood methods). Although we employ frequentist methods to support our point, neglecting to incorporate a particular level of variation into any type of statistical model will result in impoverished inference, regardless of the method. Along these lines, we found that a species effect may be incorrectly detected at relatively large sample sizes (Figure 6). As increasing sample size does not tend to decrease the incidence of Type I errors, this indicates that the most effective way to avoid such errors is to distinguish species-level from genotype-level variation. The sample sizes required to alleviate the problem of falsely attributing a pattern to the higher nested level may be several orders of magnitude higher than what is feasible to collect in natural systems. Explicitly accounting for variation at different nested levels of genetic variation would help solve this problem even at the relatively low sample sizes characteristic of most ecological studies.

Our simulated datasets also showed cases in which incorporating nested genetic variation may not be as important to consider from a functional perspective. In the algorithm we used to generate trait values from genetic loci, increasing the number of loci means that more extreme trait values are possible within a species. This effectively yields a wider distribution of trait values around the mean, with a larger proportion of values concentrated in the tails of the distribution. However, if more extreme trait values, both positive and negative, are possible in both species, both the nested and non-nested models are less likely to detect a significant effect of intraspecific variation. As the effect of intraspecific genetic variation on phenotype decreases, accounting for nested genetic variation at levels below species may be less useful for characterizing the ecological function of a population as measured by its mean trait value (Figure 5). However, in many cases the mean trait value may not be what drives community assembly or

'This article is protected by copyright. All rights reserved.'

Accepted Article

ecosystem function; instead, the variance in trait values or the extremes of the trait distribution may be more important (Violle *et al.* 2012). In those cases, other statistical methods than the ones we consider here may be more appropriate (e.g. Violle *et al.* 2012; Siefert *et al.* 2015).

It is becoming increasingly clear that genotype identity can be as important as species identity for determining the outcome of ecological processes. In the field of biodiversity-ecosystem function research, experimental and observational tests (Bailey *et al.* 2009; Latta *et al.* 2011; Avolio & Smith 2013; Crawford & Rudgers 2013; Genung *et al.* 2013) have shown that genotypic diversity within foundational species may explain as much variation in important ecosystem functions as variation in species diversity, further emphasizing the importance of taking a nested approach. Similar techniques that account for identity and diversity of both species and genotypes have been recently applied in fields including invasion biology (Bossdorf *et al.* 2005; Dlugosch & Parker 2008; Kimbro *et al.* 2013) and sustainable agriculture (Newton *et al.* 2010; Tester & Langridge 2010). Analogous to the acknowledgement that sampling effects and genotypic effects can mechanistically explain many ecological patterns and therefore should be accounted for, hierarchically partitioning variation also significantly changes how we interpret data when nested genetic variation is taken into account.

Accounting for nested levels of variation in functional traits and for the effects of different levels of diversity is important for theoretical and applied work. Historically, the approach has been more common in applied contexts such as production forestry (Jordan *et al.* 1993; Hartley 2002). However, because nested genetic effects have often been neglected in basic biodiversity research, we should acknowledge that existing syntheses and estimates of biodiversity effect sizes in meta-analyses may not tell the whole story. If the full range of variation at all levels were incorporated and nested properly, we might reach different conclusions about the mechanisms by which biodiversity influences ecosystem functioning and the strength of those effects, as our simple model suggests that effect sizes associated with non-nested models may be inflated by 20% or more.

Our results clearly demonstrate the changes in interpretation of biological relationships that can result from incorporating nested, genetically based variation into studies of natural systems. However, we admit that some biological systems are more amenable to incorporating genetic variation into experimental or observational studies than others. We therefore strongly recommend that researchers take one of the following three approaches. First, if possible, genetic variation should be quantified at the within- and among-population level and either experimentally manipulated or accounted for as a random effect in observational studies. If this is not possible, however, a promising alternative is to address intraspecific variation with a functional trait-based approach -- recent work has shown that intraspecific

'This article is protected by copyright. All rights reserved.'

Accepted Article

trait variability accounts for a large proportion of the variation in functional traits both within and among communities (Siefert *et al.* 2015), and that this variability has important consequences for both species abundance and species performance (Umaña *et al.* 2015). Finally, if neither of these approaches is tractable and intraspecific variability is neither quantified directly at the genetic level nor indirectly via functional traits, we feel that researchers should qualify their results with the explicit statement that variation at levels below species was assumed *a priori* not to affect the outcome of their study.

As a discipline, it is important to recognize the importance of hierarchically nested genetic variation and incorporate it into analyses to advance ecological research, if we are to fully understand which of the many facets of diversity have important consequences for ecosystem function. Given the growing research emphasis linking ecology and evolutionary biology (Schoener 2011), it is crucial to ensure that quantitative or molecular genetic variation is accounted for. The approach demonstrated here is particularly important for studies that examine quantitative or molecular genetic divergence, as well as species mixture effects, as we demonstrate that incorporating a properly nested variance structure may dramatically affect the way the results are interpreted. From an applied perspective, it is possible that if managers take local adaptation and intraspecific trait variation into account when managing or restoring communities affected by anthropogenically driven species loss or gain (Zenni *et al.* 2014), the functional importance of variation within species may be as important as the presence or absence of particular species for the success of management and restoration efforts. Finally, our results suggest that the response of ecological communities and the ecosystem services they provide to species gains and losses are likely more subtle than we would expect by simply observing which species invade or are extirpated. Correctly determining the nature of the relationship between biodiversity and ecosystem functioning is critical for understanding the direct and indirect effects of losses in species diversity and genetic diversity in natural systems.

Acknowledgments

The authors would like to thank J. Jiang, J. Moore, L. Moorhead, J. Henning, and anonymous reviewers for their helpful advice and comments during the simulation and writing process. The Department of Ecology and Evolutionary Biology at the University of Tennessee (Q.D.R.; J.K.B.; J.A.S.), the Rocky Mountain Biological Laboratory Jean Langenheim Fellowship (Q.D.R.), and the National Institute of Mathematical and Biological Synthesis (S.M.H.) funded this work.

Authorship Statement

'This article is protected by copyright. All rights reserved.'

Accepted Article

Q.D.R. and J.K.B. developed the idea, Q.D.R. and M.B.E. carried out the statistical analysis, Q.D.R. was the primary writer of the manuscript, and S.M.H., M.B.E, J.A.S, and J.K.B. all worked on writing and revising the manuscript.

'This article is protected by copyright. All rights reserved.'

Accepted Article

Works Cited

Anderson, J. T. and Gezon, Z. J. 2015. Plasticity in functional traits in the context of climate change: a case study of the subalpine forb Boechera stricta (Brassicaceae). - Global Change Biology 21:1689-703.

Avolio, M. L. and Smith, M. D. 2013. Correlations between genetic and species diversity: effects of resource quantity and heterogeneity. - Journal of Vegetation Science 24:1185-1194.

Bailey, J. K. et al. 2014. Indirect genetic effects: an evolutionary mechanism linking feedbacks, genotypic diversity and coadaptation in a climate change context. - Functional Ecology 28:87-95.

Bailey, J. K. et al. 2009. From genes to ecosystems: a synthesis of the effects of plant genetic factors across levels of organization. - Philosophical Transactions of the Royal Society B-Biological Sciences 364:1607-1616.

Barbour, R. C. et al. 2009. A geographic mosaic of genetic variation within a foundation tree species and its community-level consequences. - Ecology 90:1762-1772.

Borenstein, M. et al. 2009. Introduction to meta-analysis. - John Wiley & Sons, Ltd.

Bossdorf, O. et al. 2005. Phenotypic and genetic differentiation between native and introduced plant populations. - Oecologia 144:1-11.

Cadotte, M. W. et al. 2012. Phylogenetic diversity promotes ecosystem stability. - Ecology 93:S223-S233.

Caliman, A. et al. 2013. Biodiversity effects of ecosystem engineers are stronger on more complex ecosystem processes. - Ecology 94:1977-1985.

Cardinale, B. J. et al. 2012. Biodiversity loss and its impact on humanity. - Nature 486:59-67.

Clark, J. S. 2010. Individuals and the variation needed for high species diversity in forest trees. - Science 327:1129-32.

Cook-Patton, S. C. et al. 2011. A direct comparison of the consequences of plant genotypic and species diversity on communities and ecosystem function. - Ecology 92:915-923.

Coyle, J. R. et al. 2014. Using trait and phylogenetic diversity to evaluate the generality of the stress-dominance hypothesis in eastern North American tree communities. - Ecography 37:814-826.

Crawford, K. M. and Rudgers, J. A. 2013. Genetic diversity within a dominant plant outweighs plant species diversity in structuring an arthropod community. - Ecology 94:1025-1035.

Crutsinger, G. M. et al. 2006. Plant genotypic diversity predicts community structure and governs an ecosystem process. - Science 313:966-8.

Crutsinger, G. M. et al. 2008. Intraspecific diversity and dominant genotypes resist plant invasions. - Ecology Letters 11:16-23.

'This article is protected by copyright. All rights reserved.'

Accepted Article

Crutsinger, G. M. et al. 2014. Genetic variation within a dominant shrub structures green and brown community assemblages. - Ecology 95:387-398.

de Bello, F. et al. 2010. Towards an assessment of multiple ecosystem processes and services via functional traits. - Biodiversity and Conservation 19:2873-2893.

Díaz, S. et al. 2007. Incorporating plant functional diversity effects in ecosystem service assessments. - Proceedings of the National Academy of Sciences of the United States of America 104:20684-20689.

Dlugosch, K. M. and Parker, I. M. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. - Molecular Ecology 17:431-449.

Fargione, J. et al. 2007. From selection to complementarity: shifts in the causes of biodiversity-productivity relationships in a long-term biodiversity experiment. - Proceedings of the Royal Society B-Biological Sciences 274:871-876.

Fargione, J. E. and Tilman, D. 2005. Diversity decreases invasion via both sampling and complementarity effects. - Ecology Letters 8:604-611.

Feld, C. K. et al. 2009. Indicators of biodiversity and ecosystem services: a synthesis across ecosystems and spatial scales. - Oikos 118:1862-1871.

Fitzpatrick, M. C. and Keller, S. R. 2015. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. - Ecology Letters 18:1-16.

Genung, M. A. et al. 2012. Welcome to the neighbourhood: interspecific genotype by genotype interactions in Solidago influence above- and belowground biomass and associated communities. - Ecology Letters 15:65-73.

Genung, M. A. et al. 2013. Belowground interactions shift the relative importance of direct and indirect genetic effects. - Ecology and Evolution 3:1692-1701.

Ginzburg, L. R. and Jensen, C. X. J. 2004. Rules of thumb for judging ecological theories. - Trends in Ecology and Evolution 19:121-126.

Gorman, C. E. et al. 2014. Shifts in Species Interactions Due to the Evolution of Functional Differences between Endemics and Non-Endemics: An Endemic Syndrome Hypothesis. - PLoS One 9:e111190.

Hartley, M. J. 2002. Rationale and methods for conserving biodiversity in plantation forests. - Forest Ecology and Management 155:81-95.

Hiltunen, T. and Becks, L. 2014. Consumer co-evolution as an important component of the eco-evolutionary feedback. - Nature Communications 5:5226.

‘This article is protected by copyright. All rights reserved.’

Accepted Article

Hobbs, R. J. and Huenneke, L. F. 1992. Disturbance, diversity, and invasion - implications for conservations. - Conservation Biology 6:324-337.

Hoehn, P. et al. 2008. Functional group diversity of bee pollinators increases crop yield. - Proceedings of the Royal Society B-Biological Sciences 275:2283-2291.

Hooper, D. U. et al. 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. - Ecological Monographs 75:3-35.

Hughes, A. R. et al. 2008. Ecological consequences of genetic diversity. - Ecology Letters 11:609-623.

Isbell, F. et al. 2013. Nutrient enrichment, biodiversity loss, and consequent declines in ecosystem productivity. - Proceedings of the National Academy of Sciences of the United States of America 110:11911-11916.

Jordan, G. J. et al. 1993. Variation in the Eucalyptus-globulus complex revisited. - Australian Journal of Botany 41:763-785.

Kimbro, D. L. et al. 2013. Biotic resistance in marine environments. - Ecology Letters 16:821-833.

Latta, L. C. et al. 2011. Species and genotype diversity drive community and ecosystem properties in experimental microcosms. - Evolutionary Ecology 25:1107-1125.

Levine, J. M. and D'Antonio, C. M. 1999. Elton revisited: a review of evidence linking diversity and invasibility. - Oikos 87:15-26.

Levine, J. M. et al. 2003. Mechanisms underlying the impacts of exotic plant invasions. - Proceedings of the Royal Society B-Biological Sciences 270:775-781.

Lindholm, A. K. et al. 2005. Invasion success and genetic diversity of introduced populations of guppies Poecilia reticulata in Australia. - Molecular Ecology 14:3671-82.

Nakagawa, S. and Schielzeth, H. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. - Methods in Ecology and Evolution 4:133-142.

Newton, A. C. et al. 2010. Cereal landraces for sustainable agriculture. A review. - Agronomy for Sustainable Development 30:237-269.

Pante, E. et al. 2015. Species are hypotheses: avoid connectivity assessments based on pillars of sand. - Molecular Ecology 24:525-44.

Roman, J. and Darling, J. A. 2007. Paradox lost: genetic diversity and the success of aquatic invasions. – Trends in Ecology and Evolution 22:454-64.

Rudman, S. M. et al. 2015. Adaptive genetic variation mediates bottom-up and top-down control in an aquatic ecosystem. - Proceedings of the Royal Society B-Biological Sciences 282:125-132.

'This article is protected by copyright. All rights reserved.'

Accepted Article

Savolainen, O. et al. 2007. Gene flow and local adaptation in trees. - Annual Review of Ecology Evolution and Systematics 38:595-619.

Schöb, C. et al. 2015. Intraspecific genetic diversity and composition modify species-level diversity–productivity relationships. - New Phytologist 205:720-730.

Schoener, T. W. 2011. The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. - Science 331:426-429.

Siefert, A. et al. 2015. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. - Ecology Letters DOI: 10.1111/ele.12508.

Singer, M. C. and Thomas, C. D. 1996. Evolutionary responses of a butterfly metapopulation to human- and climate-caused environmental variation. - American Naturalist 148:S9-S39.

Tester, M. and Langridge, P. 2010. Breeding technologies to increase crop production in a changing world. - Science 327:818-822.

Tilman, D. et al. 1997. The influence of functional diversity and composition on ecosystem processes. - Science 277:1300-1302.

Turcotte, M. M. et al. 2013. Experimental test of an eco-evolutionary dynamic feedback loop between evolution and population density in the green peach aphid. - American Naturalist 181:S46-S57.

Umaña, M. N. et al. 2015. Commonness, rarity, and intraspecific variation in traits and performance in tropical tree seedlings. - Ecology Letters DOI: 10.1111/ele.12527.

Violle, C. et al. 2012. The return of the variance: intraspecific variability in community ecology. - Trends in Ecology & Evolution 27:244-252.

Visser, M. E. 2008. Keeping up with a warming world; assessing the rate of adaptation to climate change. - Proceedings of the Royal Society B-Biological Sciences 275:649-659.

Whitham, T. G. et al. 2006. A framework for community and ecosystem genetics: from genes to ecosystems. - Nature Reviews Genetics 7:510-523.

Zenni, R. D. et al. 2014. Rapid evolution and range expansion of an invasive plant are driven by provenance–environment interactions. - Ecology Letters 17:727-735.

‘This article is protected by copyright. All rights reserved.’

Accepted Article

Table Legend

Table 1. Summary of the effect that increasing simulation parameter values had on the difference in the proportion of variation explained by the species term (ΔR^2) and on the difference in Akaike's Information Criterion (ΔAIC) between non-nested and nested mixed models.

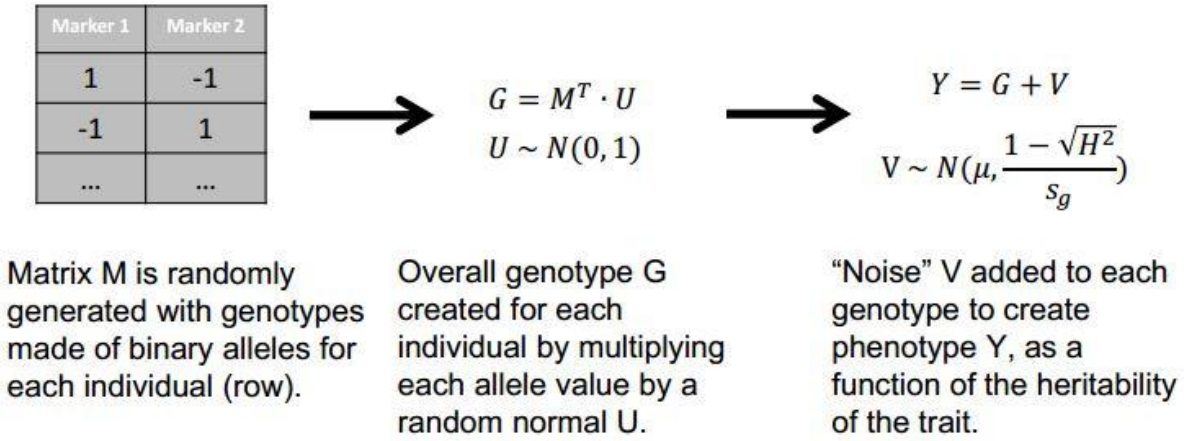
Parameter	**Effect on ΔR^2**	**Effect on ΔAIC**
Population sample size n	negative	positive
Heritability H^2	positive	positive
Genotype variance σ_g	variable	none
Species mean difference μ_1 - μ_2	unimodal	none
Number of loci	negative	unimodal

'This article is protected by copyright. All rights reserved.'

Accepted Article

Figure Legends

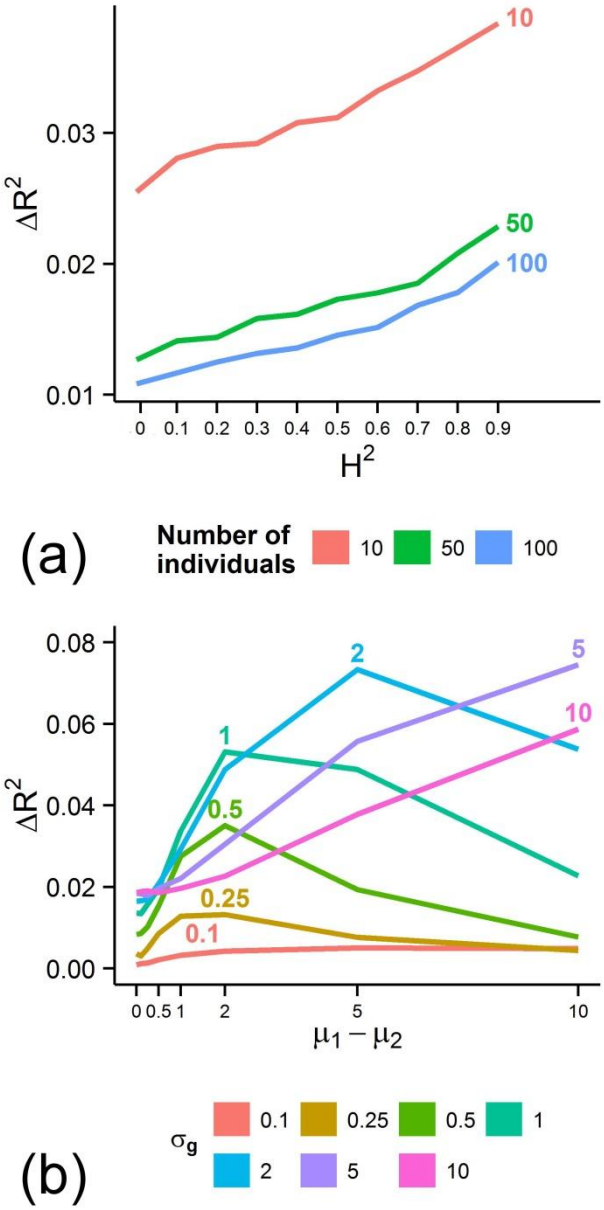
Figure 1. Diagram outlining the procedure we used to generate simulated trait data, including generating sets of alleles for each individual and drawing trait values from hierarchically nested normal distributions.



'This article is protected by copyright. All rights reserved.'

Accepted Article

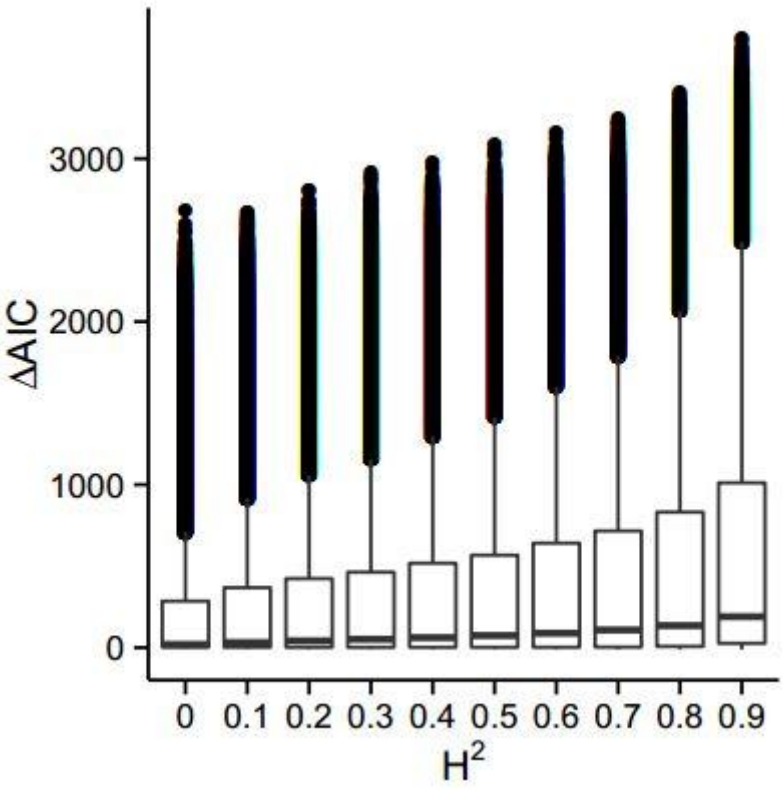
Figure 2. Mean differences between the R^2 value of the non-nested model and the marginal R^2 value of the species effect in the nested mixed model (y-axis) for different combinations of parameters in the simulated data. In panel (a) at top, the heritability parameter is on the x-axis, and each line shows the mean ΔR^2 value at each combination of heritability and sample size (colored labels associated with each line). In panel (b) at bottom, the preset difference between species means is on the x-axis, and each line shows the mean ΔR^2 value at each combination of species mean difference and among-genotype variance (colored labels associated with each line).



'This article is protected by copyright. All rights reserved.'

Accepted Article

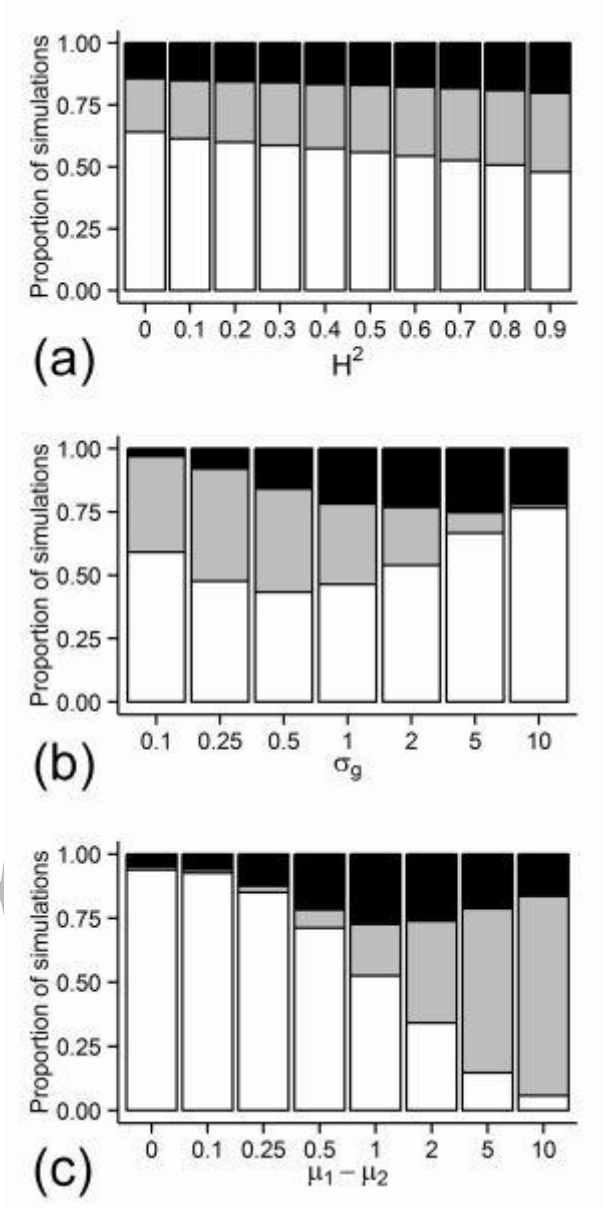
Figure 3. As the heritability parameter (x-axis) of simulated trait data sets increased, there was a corresponding increase in the ΔAIC value (y-axis) comparing the nested mixed model to the non-nested linear model, indicating that the nested mixed model is more superior as heritability increases.



'This article is protected by copyright. All rights reserved.'

Accepted Article

Figure 4. Discrepancy among models considering $\alpha = 0.05$ significance threshold from simulation results. Bars are shaded based on the proportion of simulation outcomes. White-shaded bars indicate that neither model rejected the null hypothesis that species means are equivalent, grey shading indicates that both rejected this hypothesis, and black indicates that only the non-nested model rejected this hypothesis. The proportion of simulations in which both models reached the same conclusion decreased as the heritability parameter increased (a), was lowest at intermediate among-genotype variance (b), and was lowest at intermediate among-species trait mean differences (c).



'This article is protected by copyright. All rights reserved.'

Accepted Article

Figure 5. Effect of increasing the number of loci, displayed on the x-axis in all three panels, on (a) ΔAIC comparing the non-nested linear model and nested mixed model; (b) ΔR^2 comparing the two models; and (c) discrepancy between the two models considering the $\alpha = 0.05$ significance threshold. As in Figure 4, white-shaded bars indicate that neither model rejected the null hypothesis that species means are equivalent, grey shading indicates that both rejected this hypothesis, and black indicates that only the non-nested model rejected this hypothesis. The ΔAIC value was highest at 3 loci, but the ΔR^2 and the proportion of discrepancies both decreased with increasing numbers of loci.

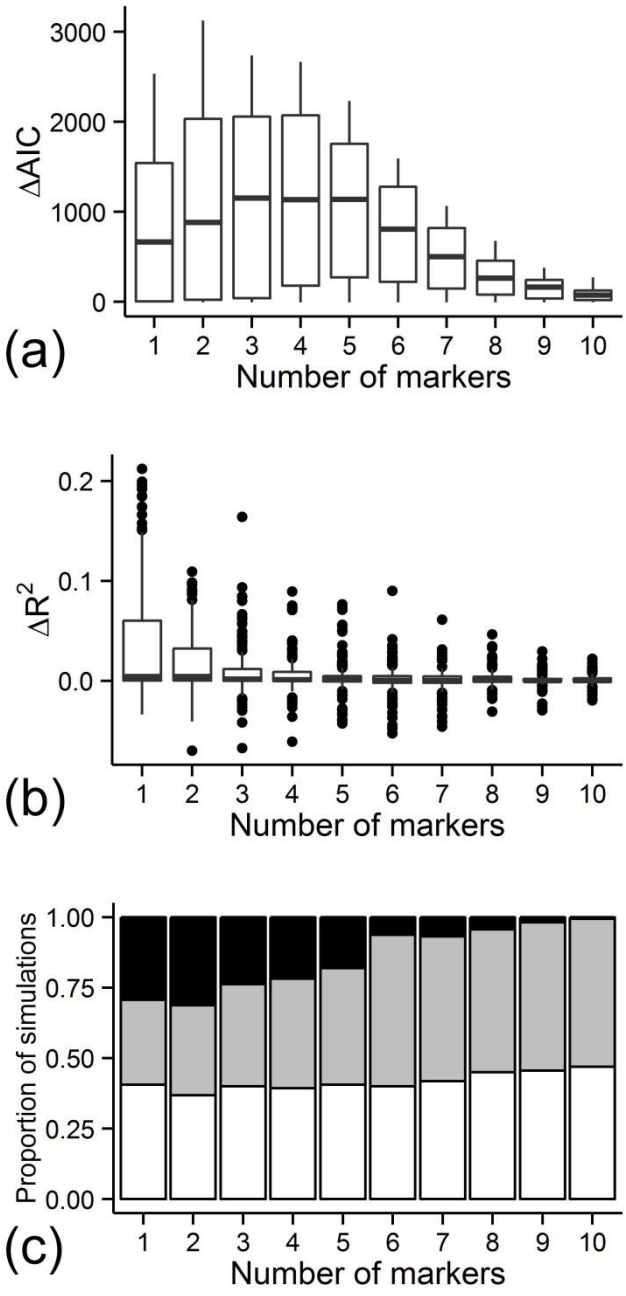
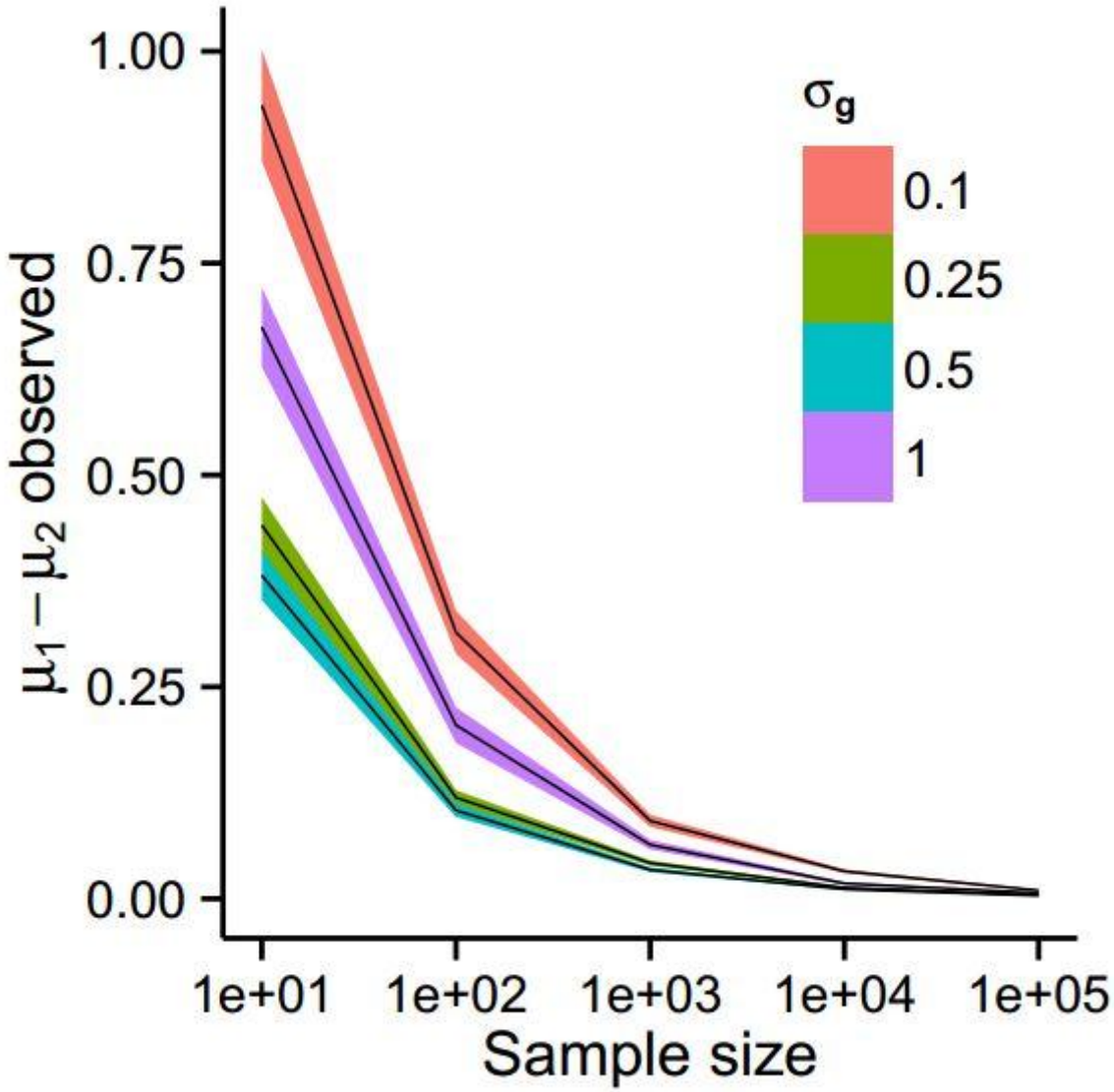


Figure 6. Absolute value of observed difference in species means plotted on the y-axis against the logarithm of sample size on the x-axis. Each of the simulated datasets was drawn from a distribution with

'This article is protected by copyright. All rights reserved.'

Accepted Article

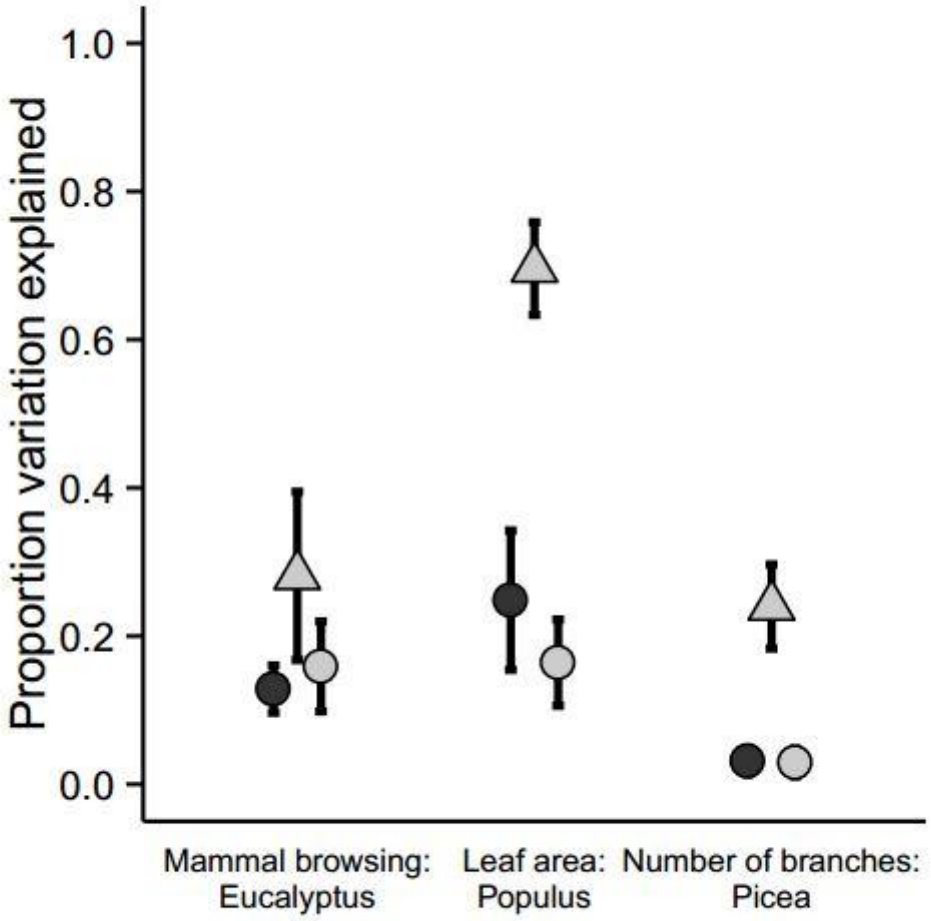
a mean of zero, so the expected $\mu_1 - \mu_2 = 0$. The width of the shaded region around each line is the standard error of the difference between means, and the color of each shaded region represents the value of σ_g, genotype-level standard deviation.



'This article is protected by copyright. All rights reserved.'

Accepted Article

Figure 7. Coefficient of determination (R^2) values calculated from *Eucalyptus* mammal browsing, *Populus* leaf area, and *Picea* branch number data sets, respectively. Black circles represent the non-nested model R^2 values, grey circles represent the marginal R^2 value of the nested mixed models, and grey triangles represent the conditional R^2 values of the nested mixed models. Error bars represent bootstrapped standard errors of the R^2 values.



'This article is protected by copyright. All rights reserved.'

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.