Supplementary Material

Detailed Methods

Modeling framework

As outlined in the main text, our statistical framework is composed of three fundamental steps. Each of the steps could be carried out in a number of ways. We present the model in its full generality highlighting the model presented in the main text as the simplest version of the full model.

a) From tree topology to interactions

We represent the topology of a binary phylogenetic tree with n leaves (i.e., the species in the pool) as a binary matrix V of size $(2n-1) \times n$. The rows of the matrix represent either the extant species or their ancestors (including the common ancestor to all species, i.e., the root of the tree), while the columns represent the extant species. A coefficient $V_{ij} = 1$ if i is an ancestor of j (or j itself), and is zero otherwise. This matrix is called the "basis matrix" of the tree in the literature (Bravo *et al.* 2009), and is closely related to the correlation structure induced by shared evolutionary history (Hansen & Martins 1996): given a vector of $2n-1$ branch lengths α, then $V^T D(\alpha) V$ is the variance-covariance matrix of the tree.

Inspired by classic consumer-resource models, we parameterize the $n \times n$ matrix of interactions between the species as:

$$A = D(\gamma) V^T D(\lambda) V D(\theta) \quad , \tag{1}$$

where γ and θ are positive vectors of length n with entries summing to one (thus representing $n-1$ free parameters each), λ is a positive vector of length $2n-1$, $D(\zeta)$ is a diagonal matrix with the elements of ζ on the diagonal, and V^T denotes the transpose of matrix V. We therefore have $4n-3$ free parameters in total. This is the full, most complex model we consider, referred to as model 4. The model presented in the main text (model 1) has the same formulation except that in its simplest version both $D(\gamma)$ and $D(\theta)$ are taken to be $\frac{1}{n} I$, i.e., a scaled version of the identity matrix, which reduces the number of parameters to $2n-1$. We also consider models with intermediate complexity in which either $D(\gamma)$ (model 2) or $D(\theta)$ (model 3) are taken as the scaled identity matrix, resulting in $3n-2$ parameters on each of these alternative versions. For every tree with n species that we are fitting, regardless of the model, we fit $2n-1$ branch length parameters, which include the branches related to the n tips, representing the independent evolutionary history of each species, and all $n-1$ internal branches. Eq.1 expresses overlap in resource use between species, with more closely related species sharing more similar resource requirements, and competing more strongly as a result. Further details on these parameterizations and their derivation from consumer-resource models are reported in the section "Derivation of the model from consumer-resource dynamics".

Provided with a tree topology describing the evolutionary history of the n species in the pool (defining the associated matrix V), and empirical abundances for a pool of species grown in different combinations, we aim to find the maximum likelihood estimates of the parameters γ, λ, and θ, keeping in mind that for the simplest model formulation in the main text we are only estimating λ and for model 2 we also estimate γ, for model 3 we also estimate θ and for the full model we estimate all parameters. To achieve this goal, we first relate interactions to predicted abundances, and then predicted abundances to the experimental data.

b) From interactions to predictions

We want to relate species interactions, encoded in the phylogenetically-structured matrix A (Eq. 1) with the abundances species would attain in different assemblages. Take k to be the set of species in a given assemblage. We assume that the abundance of species i in assemblage k, denoted $x_i^{(k)}$, is linearly related to the abundance of the other species in the assemblage, with each effect mediated by the relevant pairwise interaction strength:

$$x_i^{(k)} = \frac{1}{A_{ii}} \left(1 - \sum_{j \in k; j \neq i} A_{ij} x_j^{(k)} \right) \quad .$$

Thus, species i grown alone would reach $1/A_{ii}$ (its carrying capacity), while growth together with j would yield a lower value whenever $A_{ij} > 0$, i.e., when species i and j are competitors.

Writing $A^{(k)}$ for the sub-matrix of A obtained by retaining only the columns and rows of species in assemblage k, we can write this system of equations in more compact form as:

$$A^{(k)} x^{(k)} = 1^{(k)} \quad ,$$

where $1^{(k)}$ is a vector of ones with as many elements as there are species in k. Thus, provided with a matrix A we can compute the value of $x^{(k)}$, for any k:

$$x^{(k)} = \left(A^{(k)}\right)^{-1} 1^{(k)} \quad . \tag{2}$$

There are two possible outcomes: a) all the elements of $x^{(k)}$ are positive, which we interpret as predicted coexistence of the species in assemblage k, with expected abundances specified by the vector $x^{(k)}$; or b) one or more of the elements are negative, which we interpret as the impossibility of coexistence for this set of species. Versions of this basic framework have been proposed numerous times in the ecological literature (Ansari *et al.* 2021; Fort 2018; Maynard *et al.* 2020; Xiao *et al.* 2017), and here we closely follow the approach of Maynard *et al.* (2020) and its extension in Skwara *et al.* (2023).

Provided with a matrix A defined in Eq. 1, we can predict the coexistence and expected abundances of species in all possible assemblages using Eq. 2. We then compare these predictions to the empirical data.

c) From predictions to likelihoods

Each data set comprises a set of measurements $\tilde{x}_i^{(k,r)}$, specifying the biomass of species i when growing in assemblage k (i.e., defined by the extant species in the plot) and replicate r. The data sets contain a large number of assemblages, possibly with replicated experiments (i.e., distinct plots in which the same final assemblage is observed). We take these measurements to be noisy estimates of the "true", expected biomass of species i in assemblage k:

$$\tilde{x}_i^{(k,r)} \overset{iid}{\sim} \text{Gamma}\left(\alpha, \frac{\alpha}{x_i^{(k)}}\right) \quad , \tag{3}$$

where we assume that the observations have been sampled independently from a Gamma distribution with parameters α (scale) and $\alpha/x_i^{(k)}$ (rate), such that the expected value is $\mathbb{E}(\tilde{x}_i^{(k,r)}) = x_i^{(k)}$ and the variance is $\mathbb{V}(\tilde{x}_i^{(k,r)}) = \frac{1}{\alpha}\left(x_i^{(k)}\right)^2$. The parameter α is assumed to be specific to the data set. Thus, for a given data set all measurements are taken from distributions with a constant coefficient of variation, $\frac{1}{\sqrt{\alpha}}$. Whenever $\alpha > 1$, the mode of the distribution is larger than zero, and the distribution displays a central tendency.

The choice of this particular Gamma distribution is convenient (e.g., it has support on the positive real line, a constant coefficient of variation, and adds a single free parameter); alternatively, one could assume a Log-Normal distribution, or any other appropriate alternative. We do not expect the choice of a particular distribution to affect our qualitative conclusions.

Numerical search for maximum likelihood estimates

In summary, provided with the phylogenetic tree for a species pool, the parameters of the model (the vectors γ, θ and λ, as well as the parameter α controlling the shape of the distribution), and the empirical data, comprising measurements of the biomass of the species across different (possibly replicated) assemblages, we can: a) use the parameters and the tree to build a matrix of interactions (Eq. 1), b) compute the predicted abundances for all assemblages (Eq. 2), and c) compute a likelihood by multiplying the probability of observing each measurement under the fitted model Eq. 3.

This bridge from parameters to likelihoods allows us to search numerically for the maximum likelihood estimates of the parameters. Briefly, a) we initialize the parameters such that A is the sum of the identity matrix and a random matrix with small norm (practically, ensuring coexistence in all possible assemblages); b) search for maximum likelihood parameters via a heuristic optimization routine that combines simulated annealing, hill-climbing (where we slightly alter parameters and accept the move if the likelihood is increased), and numerical optimization via the Nelder-Mead and BFGS algorithms implemented in the *optim* function in *R* (R Core Team 2021); c) each step in b) is repeated 500 times until (presumptive) convergence to the maximum likelihood estimates. Parameterizations that predict a lack of coexistence for assemblages that were observed to coexist experimentally are discarded. Parameterizations that yield a likelihood lower than found for the corresponding star tree are also discarded. Finding optimal parameters is computationally very difficult (there can be many local optima in the parameter space), and very expensive (because each prediction requires the inversion of a matrix, following Eq. 2). We implemented the approach using a combination of *Stan* (Stan Development Team 2023) and *R* (R Core Team 2021) code, started each parameter search five times with a random initialization, and retained the best likelihood for each data set/model combination. The code, raw and organized data, and results are available at `github.com/StefanoAllesina/lemos-costa-2023`.

Testing for phylogenetic structure

In the preceding sections, we showed that, given empirical data and a matrix V corresponding to a phylogenetic tree, we can find the maximum likelihood estimates of the parameters in Eq. 1, and thus the maximum likelihood associated with the tree, $\mathcal{L}_V$. To test whether phylogeny can explain species interactions, and consequently the observed abundances in different assemblages, we repeat the same fitting procedure for a random tree V', yielding $\mathcal{L}_{V'}$. By repeating this procedure many times, we can estimate the probability that a random tree yields a higher likelihood than the "true" tree associated with V. As such we take $p \approx \frac{1}{m}\sum_i^m \mathcal{H}(\mathcal{L}_V \leq \mathcal{L}_{V_i'})$, where m is the number of random trees, $\mathcal{H}(\cdot)$ the Heaviside step function and V_i' is the matrix associated with the i^{th} random tree.

Finally, we need to define what we mean by a random tree. We consider two possibilities: a) a random tree with the same number of leaves as the original constructed according to the Yule model, a classic speciation model in which every branch has the same probability of speciating at every time; b) a tree with the same topology as the original, but where we shuffle species' identities. Therefore, for each data set we present randomizations for four nested models (all variations on Eq. 1, with the simplest model detailed in the main text) and two distinct ways of constructing the random trees. For each randomization method, we sample tree topologies until we have found one thousand distinct trees; whenever duplicate trees are encountered in this process, we fit both and retain the best likelihood.

Data and data organization

The methods of Maynard *et al.* (2020) were devised to analyze experiments carried out in controlled conditions with a small species pool, for which a substantial proportion of all possible combinations of species were co-cultured. We anaylzed data sets from biodiversity and ecosystem functioning experiments which are often conducted in experimental plots, using a pool of a dozen or more species. Because the number of possible assemblages one can form with n species is 2^n, and only dozens or hundreds of experimental plots are typically feasible, the number of sampled assemblages in these experiments is necessarily small compared to the theoretical limit (e.g. 2^{10} is already more than a thousand distinct assemblages for $n = 10$). We therefore attempted to select data sets in which a) only species that appear in multiple plots are considered to be part of the pool; b) unidentified biomass or non-target species do not dominate the plots; c) we have sufficient replication or diversity of assemblages to confidently fit the parameters of the model.

To this end, we consider data collected as part of the Biodiversity II experiment (Tilman *et al.* (2001), data from 2001 to 2018, inclusive, downloaded from the Cedar Creek website), the Wageningen diversity experiment (Van Ruijven & Berendse (2010), provided by Dr. Jasper van Ruijven), and an experiment carried out at the Toronto's Koffler Scientific Reserve by Marc W. Cadotte (data accompanying Cadotte (2013)).

The data from Cadotte (2013) contains plots in which a subset of 14 species were planted. Some did not

grow and went extinct during the experiment. We consider a species that did not reach 2.5% of its maximum observed biomass to be extinct in the plot. We then considered species that appeared in at least 5 distinct assemblages; if a plot contained species that did not meet this requirement, the whole plot was discarded; the procedure was then repeated until we obtained a subset of the original data in which all species appeared in at least 5 distinct assemblages. This left us with 8 species, forming 27 distinct assemblages, for a total of 87 measurements (including replicated assemblages). Code to implement this procedure, along with the raw and elaborated data are in the aforementioned repository (`github.com/StefanoAllesina/lemos-costa-2023`).

The same general procedure was applied to the data from Van Ruijven & Berendse (2010), with each species also required to be present in 5 distinct assemblages. After the selection procedure, all eight species were retained, and we obtained nine separate data sets, one for each yearly sampling from 2001 to 2009, inclusive.

The data from Tilman *et al.* (2001) are the most difficult to process, because aside from the original target pool of eighteen species, many other species are found in the plots, along with a substantial amount of litter or unidentified biomass. We therefore included in the "species pool" a larger number of species, excluded plots in which the total biomass for species not in the pool surpassed 10% of the total biomass, and then applied the procedure used for the other data sets. In this way, we were able to extract nine data sets (for distinct years) containing between 13 and 20 species and a substantial number of assemblages.

Summary statistics for each data set, including number of species, number of unique assemblages and number of measurements are reported in Table 2. The name of the species in each data set across all years (when applicable) are in Table 1.

To obtain a phylogeny for each species pool, we normalized species names using the World Flora Database (*http://www.worldfloraonline.org/*) and then constructed the phylogenetic tree associated with each set of species using the *V.PhyloMaker* package in *R* (Jin & Qian 2019).

Table 1: Species names and family for each data set

source	species	family
Cadotte	Andropogon gerardii	Poaceae
Cadotte	Desmodium canadense	Fabaceae
Cadotte	Elymus canadensis	Poaceae
Cadotte	Elymus trachycaulus	Poaceae
Cadotte	Monarda fistulosa	Lamiaceae
Cadotte	Rudbeckia hirta	Asteraceae
Cadotte	Schizachyrium scoparium	Poaceae
Cadotte	Solidago altissima	Asteraceae
Biodiversity II	Achillea millefolium	Asteraceae
Biodiversity II	Agrostis rudis	Poaceae
Biodiversity II	Amorpha canescens	Fabaceae
Biodiversity II	Andropogon gerardi	Poaceae
Biodiversity II	Asclepias tuberosa	Apocynaceae
Biodiversity II	Bouteloua aristidoides	Poaceae
Biodiversity II	Elymus repens	Poaceae
Biodiversity II	Koeleria pyramidata	Poaceae
Biodiversity II	Lespedeza capitata	Fabaceae

source	species	family
Biodiversity II	Liatris aspera	Asteraceae
Biodiversity II	Lupinus perennis	Fabaceae
Biodiversity II	Monarda fistulosa	Lamiaceae
Biodiversity II	Panicum oligosanthes	Poaceae
Biodiversity II	Panicum sumatrense	Poaceae
Biodiversity II	Petalostemum villosum	Fabaceae
Biodiversity II	Poa pratensis	Poaceae
Biodiversity II	Rumex acetosella	Polygonaceae
Biodiversity II	Schizachyrium scoparium	Poaceae
Biodiversity II	Solidago nemoralis	Asteraceae
Biodiversity II	Sorghastrum nutans	Poaceae
Biodiversity II	Sporobolus cryptandrus	Poaceae
Wageningen	Agrostis capillaris	Poaceae
Wageningen	Anthoxanthum odoratum	Poaceae
Wageningen	Centaurea jacea	Asteraceae
Wageningen	Festuca rubra	Poaceae
Wageningen	Holcus lanatus	Poaceae
Wageningen	Leucanthemum vulgare	Asteraceae
Wageningen	Plantago lanceolata	Plantaginaceae
Wageningen	Rumex acetosa	Polygonaceae

Derivation of the model from consumer-resource dynamics

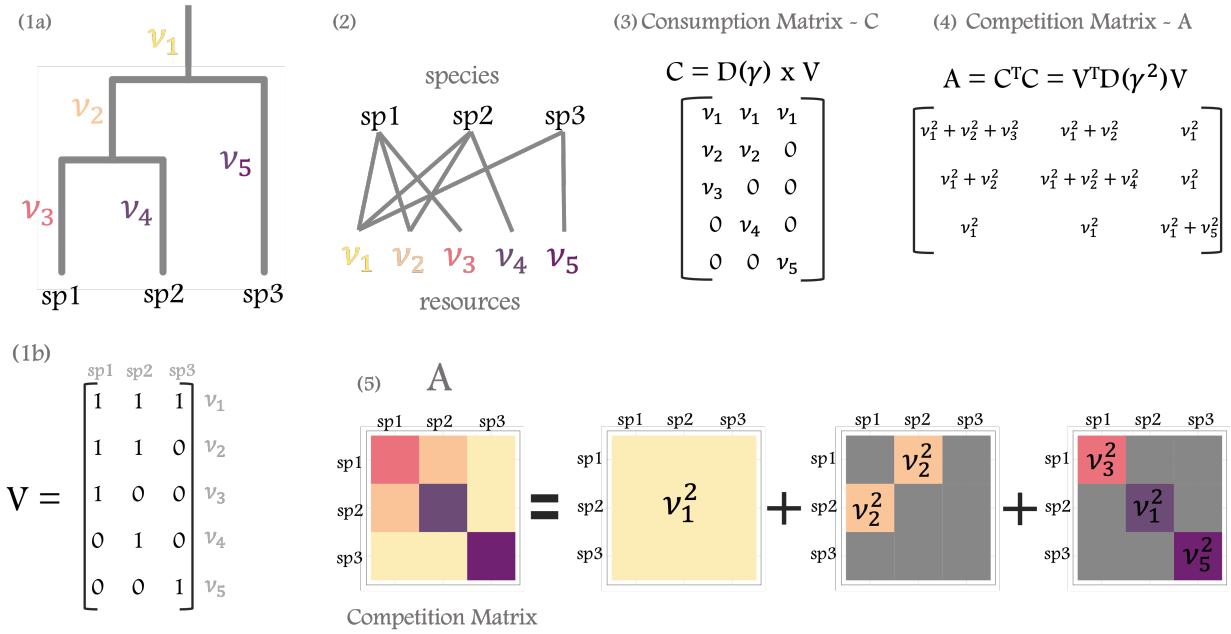

Figure 5: Consumer-resource framework for the model presented in the main text. 1) Each clade (including all the descendent species of a common ancestor) has access to a "private" resource, as encoded in matrix V (the rows are the resources, the columns the species, and a coefficient is one if the species can access the resource). 2) Network representation of the species (nodes sp_i) and the resources (nodes ν_i) they consume (links) 3) Each resource has a value, stored in vector ν; in this case the consumption matrix is simply $C = D(\nu)V$. 4) The competition matrix A is obtained through the consumption matrix and is related to resources' values and species' shared evolutionary history enconded in the matrix V. The resulting matrix has a nested block structure, with a block per clade. 5) An additional interpretation of the competition matrix can be derived by considering the A matrix as the sum of the competition coefficients (ν_j^2) for each clade a species belongs to

We consider n consumer species with a specified phylogeny. We imagine a situation in which each clade in the tree has access to a clade-specific resource. Clade-specific resources are an idealization of the resources needed by the clade to support themselves and can be considered as the pool of all resources used by the species in the clade (including "private" resources which are the requirements of individual species. Thus we can build a matrix V with $2n-1$ rows (the resources, one per clade, which also includes a resource accessible to all species) and n columns (the consumers). The matrix has a coefficient of $V_{ij} = 1$ whenever consumer j has access to resource i, and zero elsewhere. Next, we model consumption, by assuming that each resource has a certain value (taken to be the same irrespective of the consumer), and that each consumer has a given uptake rate (taken to be the same for each resource):

$$C = D(\upsilon)V\,D(\alpha) = \upsilon\alpha^T \circ V \quad ,$$

where $D(\upsilon)$ is a diagonal matrix with the value of each resource on the diagonal, $D(\alpha)$ a diagonal matrix of uptake rates and $\circ$ is the Hadamard product. We obtain a system of $2n-1+n$ differential equations, which follow the Generalized Lotka-Volterra model (MacArthur consumer-resource model MacArthur (1970)):

$$\begin{cases} \frac{dx_i}{dt} = x_i\left(-\delta_i + \sum_{j=1}^{2n-1} C_{ji}y_j\right) \\ \frac{dy_j}{dt} = y_j\left(\rho_j - \kappa_j y_j - \sum_{i=1}^{n} C_{ji}x_i\right) \end{cases} ,$$

in which $x_i = x_i(t)$ is the density of consumer i, $y_j = y_j(t)$ is the density of resource j, and all parameters are positive. δ_i is the intrinsic death rate of consumer i, ρ_j is the growth rate of resource j and κ_j is its self-limitation. If we assume that resource dynamics quickly equilibrate for any consumer densities (practically, if $\rho_j \gg \sum_{i=1}^{n} C_{ji} x_i$), we can perform a "separation of time-scales" as in MacArthur (1970), and solve the equation for the equilibrium of y_j:

$$y_j^\star = \frac{1}{\kappa_j}\left(\rho_j - \sum_{i=1}^{n} C_{ji} x_i\right) \quad .$$

Plugging these values into the equations for the consumers yields:

$$\begin{aligned}
\frac{dx_i}{dt} &= x_i\left(-\delta_i + \sum_j C_{ij}\frac{1}{\kappa_j}\left(\rho_j - \sum_i C_{ji} x_i\right)\right) \\
&= x_i\left(-\delta_i + \sum_j C_{ij}\frac{\rho_j}{\kappa_j} - \sum_j \sum_k C_{ji}\frac{1}{\kappa_j} C_{jk} x_k\right) \quad , \\
&= x_i\left(\zeta_i - \sum_k B_{ik} x_k\right)
\end{aligned}$$

where we have defined $\zeta_i = (\sum_j C_{ij}\rho_j/\kappa_j) - \delta_i$ and $B = C^T D(\kappa)^{-1} C$. Finally, we factor ζ_i outside the parenthesis, obtaining:

$$\frac{dx_i}{dt} = x_i\zeta_i\left(1 - \sum_k \frac{1}{\zeta_i} B_{ik} x_k\right) \quad .$$

The matrix of species' interactions is therefore:

$$\begin{aligned}
A &= D(\zeta)^{-1} B = D(\zeta)^{-1} C^T D(\kappa)^{-1} C \\
&= D(\zeta)^{-1} D(\alpha) V^T D(\nu) D(\kappa)^{-1} D(\nu) V D(\alpha) \\
&= D(\gamma) V^T D(\lambda) V D(\theta) \quad ,
\end{aligned}$$

where $\theta = \alpha$, $\gamma = \frac{\alpha}{\zeta}$ and $\lambda = \frac{\nu^2}{\kappa}$. This is exactly the Eq. 1 in the main text, which we use for model 4. Further assuming that $\zeta = 1/\alpha$ we obtain model 2; choosing $\alpha = 1$ yields model 3 and, finally, using both simplifications we obtain model 1. Figure 5 illustrates the parameterization for the case of the simplest model: $A = V^T D(\lambda) V$. We note that the variance-covariance matrix of a tree with branch lengths λ would take same form; as such, an alternative parameterization is thus to take competition coefficients to be the result of evolution via Brownian motion of many traits (Ives & Helmus 2011; Serván *et al.* 2023). Additionally the matrix A is proportional to the expected variance-covariance matrix of the residual error in a phylogenetic generalized least squares regression model (Revell 2010).

Tree distances and likelihoods

To better understand the significantly good fit we observe for the true phylogenetic tree topology for most data sets, we investigated whether randomized trees that were more similar to the true tree also had higher likelihoods on average. To quantify the similarity between randomized trees and the true topology for each data set, we calculated the distance between them using an information-based generalized Robinson-Foulds metric, which we implemented using the TreeDist package in R (Smith 2020). This metric measures the similarity between two tree topologies by quantifying how much information the splits in the two trees hold in common. We also considered several alternative tree distance metrics, all of which produced similar results (not shown).

We then calculated the Pearson correlation between log-likelihood and distance (from the original tree) in each set of randomized trees (across all 19 data sets and 2 randomization methods). The original tree was always excluded from this analysis. The correlations are shown in Figure 6, plotted against the p-value found for the true tree against the corresponding randomization ensemble. We also regressed log-likelihood against distance in a linear model, to check whether there was a significant effect of distance.

Because many of the distributions of tree distances contained high leverage points, we used bootstrap resampling to calculate confidence intervals for the correlation coefficients and counted the fraction of bootstrap samples where distance was a significant predictor of log-likelihood (at the $\alpha = 0.05$ level).

Overall, we found consistent negative relationships between the distance of each randomized tree from the true topology and its log-likelihood. These relationships were generally weak but consistently significant for 17 out of 18 cases for the Biodiversity II data sets and 13 out of 18 cases for the Wageningen data sets. For the remaining data sets, including the Cadotte data, we found no evidence for a relationship between distance and log-likelihood. The strongest negative correlations coincided with the data sets where the true tree topology performed best (smallest p-values). However, we also found significant slopes for several data sets where the randomization test did not yield significant evidence for phylogenetic structure. In these cases, we hypothesize that the ability of trees to fit the data does not fall off smoothly with distance to the true phylogeny, but depends on specific topological features, such as the key splits we identify in Figure 4. Thus, because many (i.e. more than 1 in 20) random trees might have these key features, they can provide an equal or better fit compared to the true topology, despite differing from the true topology in other splits (which matter less for structuring interactions).

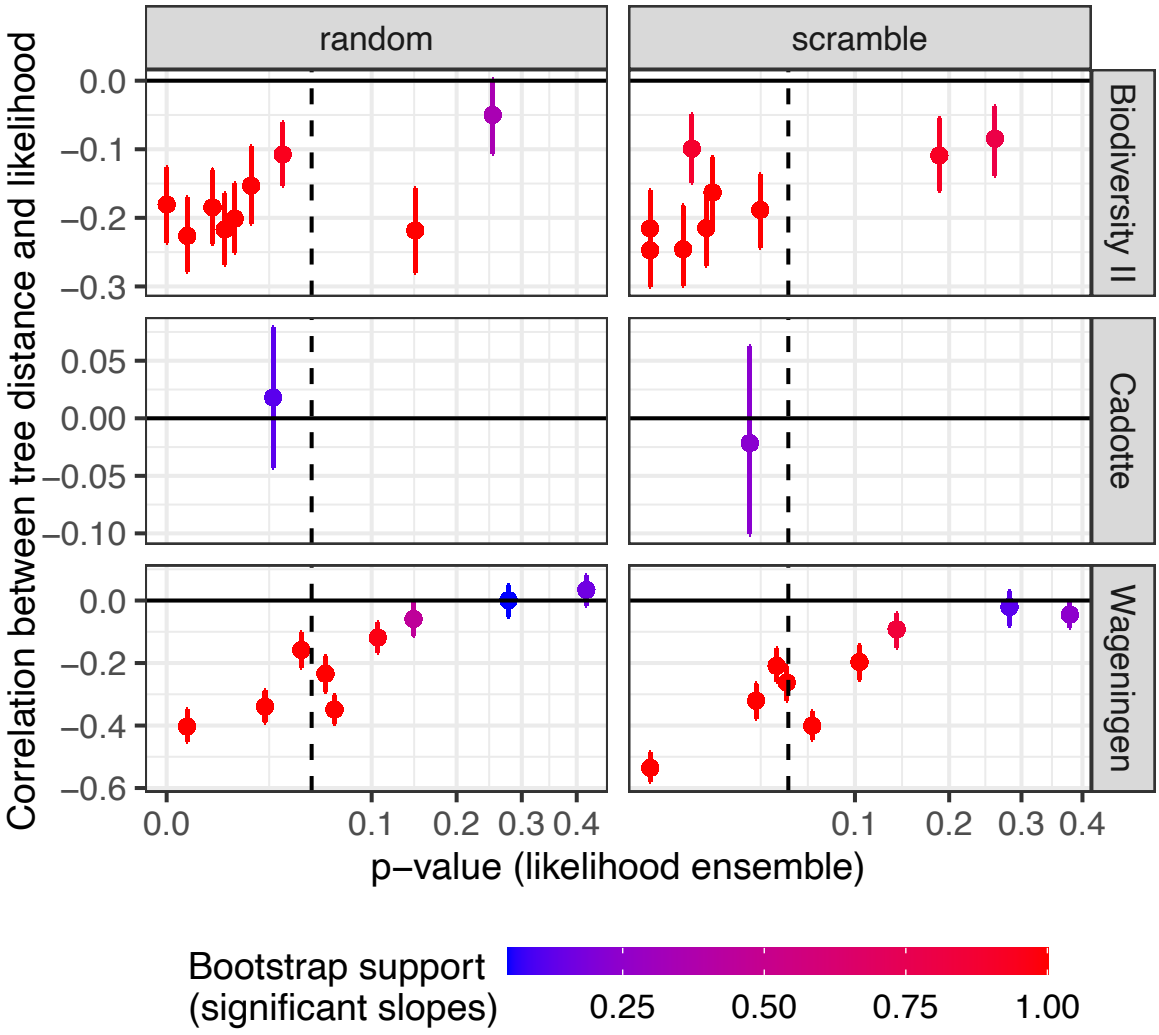

Figure 6: **Relationships between distance from the true tree and log-likelihood for randomized trees**. For each experiment (rows) and randomization method (columns), we plot the correlation between the distance from the true tree and the log-likelihood found in the ensemble of randomized trees (y-axis). For comparison, the p-value for the corresponding true phylogeny is plotted on the x-axis. Each point represents the set of randomizations corresponding to a specific data set (experiment/year combination). Bootstrap 95% confidence intervals are shown, and color indicates the fraction of bootstrap samples in which the slope of log-likelihood vs. distance was significantly different from zero. We find that almost all correlations are less than zero (solid horizontal line), and many data sets have consistently significant relationships between distance and log-likelihood. This is especially true for data sets where the randomization test p-value is less than 0.05 (vertical dashed line). Note that p-values (x-axis) are shown square-root transformed for better visualization.

Parameterizations based on molecular branch lengths

In our framework we fit models of species interactions based on population dynamics and constrained by species shared evolutionary history. Ultimately, our model uses phylogenetic tree structure to find the maximum likelihood estimates of branch lengths which are directly associated with interspecific competition based on data of species' biomasses when grown in different combinations. As highlighted in the main text, we make no assumption on the linearity or shape of the relationship between shared evolutionary history and competition strength. Instead we let the data (ecology) guide the fit, finding the distribution of branch lengths that maximizes the likelihood of the parameters given the data and model parameters while also minimizing predicted biomasses derived from model prediction versus observed biomasses in the experimental communities. It is also reasonable to imagine that the information encoded in the branch lengths of molecular phylogenies could still provide insights into the structure of these experimental ecological communities. If these branch lengths are reflective of trait evolution and most importantly, evolution of traits that are directly related to competitive interactions we would observe a good fit between predictions made by using the molecular phylogeny as the input of the interaction matrix in our modeling framework. To test this, we parameterize our model for species interactions with the molecular phylogeny, i.e, instead of fitting the branch lengths we used the distribution of branch lengths from the molecular phylogeny to get estimates of predicted abundances of species in the experimental communities. We then compared these predictions with the observed abundances. It is worth noticing that we are interested in the correlation between predicted and observed abundances and not necessarily in the absolute values of predicted abundances, as all abundances can be rescaled if we fitted the root. Fitting the root would only rescale the biomasses without changing the observed correlations. Additionally, because the phylogenetic trees used in our analyses were ultrametric, all species are predicted to achieve the same biomass when grown in monoculture, which is already a biologically unrealistic expectation. All data sets showed a poor correlation between the log of predicted and the log of observed biomasses (top panel in Figures 7-24).

Additionally, we compared the fitted branch lengths from our modeling framework with the molecular branch lengths from the phylogenies, showing that also in this case, the correspondence between estimated branch lengths show very little agreement (bottom panels in Figures 7-24). In this comparison we are also interested in the shape of the relationship and not in correspondence between the units, as the units of measurement of the molecular phylogeny are not comparable with the units of measure of the fitted branch lengths from our modeling framework. However, the comparison of the fitted branch length is made with the transformed square root of the molecular branch length as functional dissimilarity has been shown to be related to the square root of molecular distances (Letten & Cornwell 2015) and is an appropriate transformation for molecular branch lengths, but the same result applies to the un-transformed branch lengths.

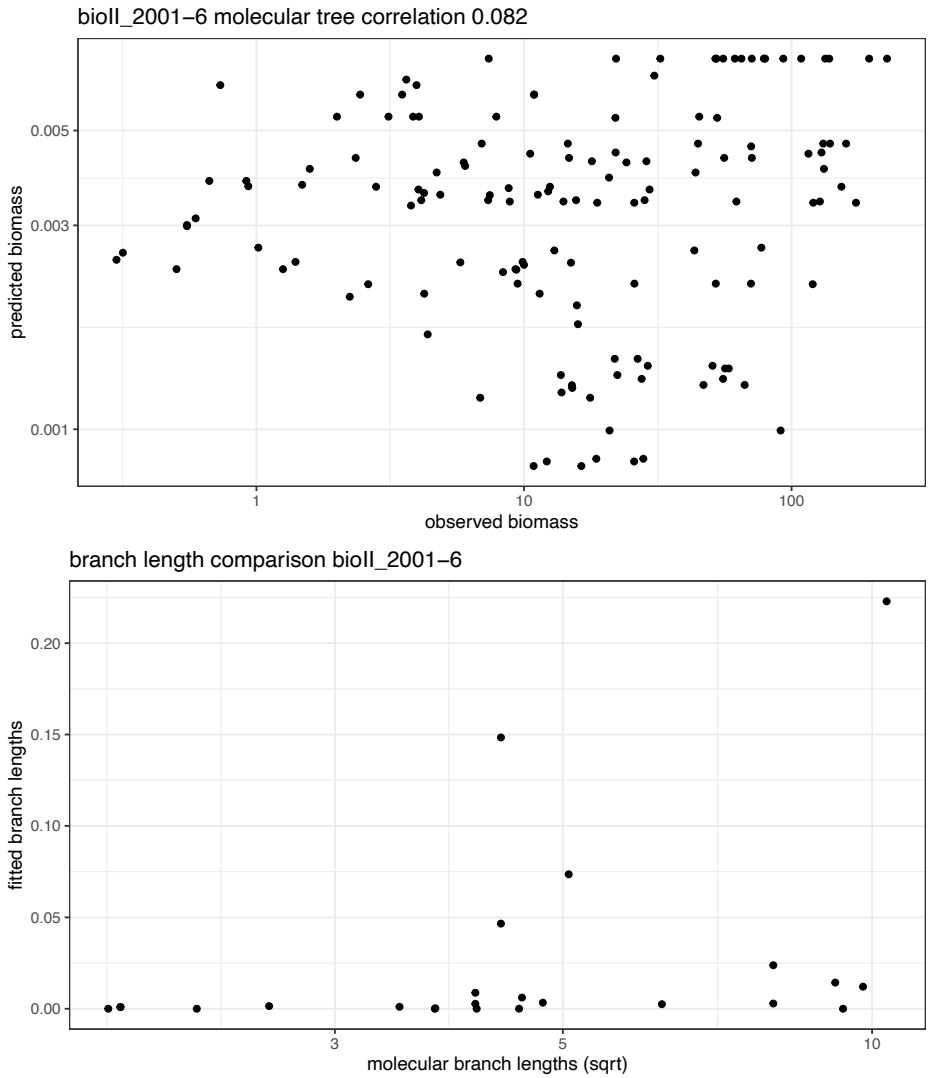

Figure 7: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

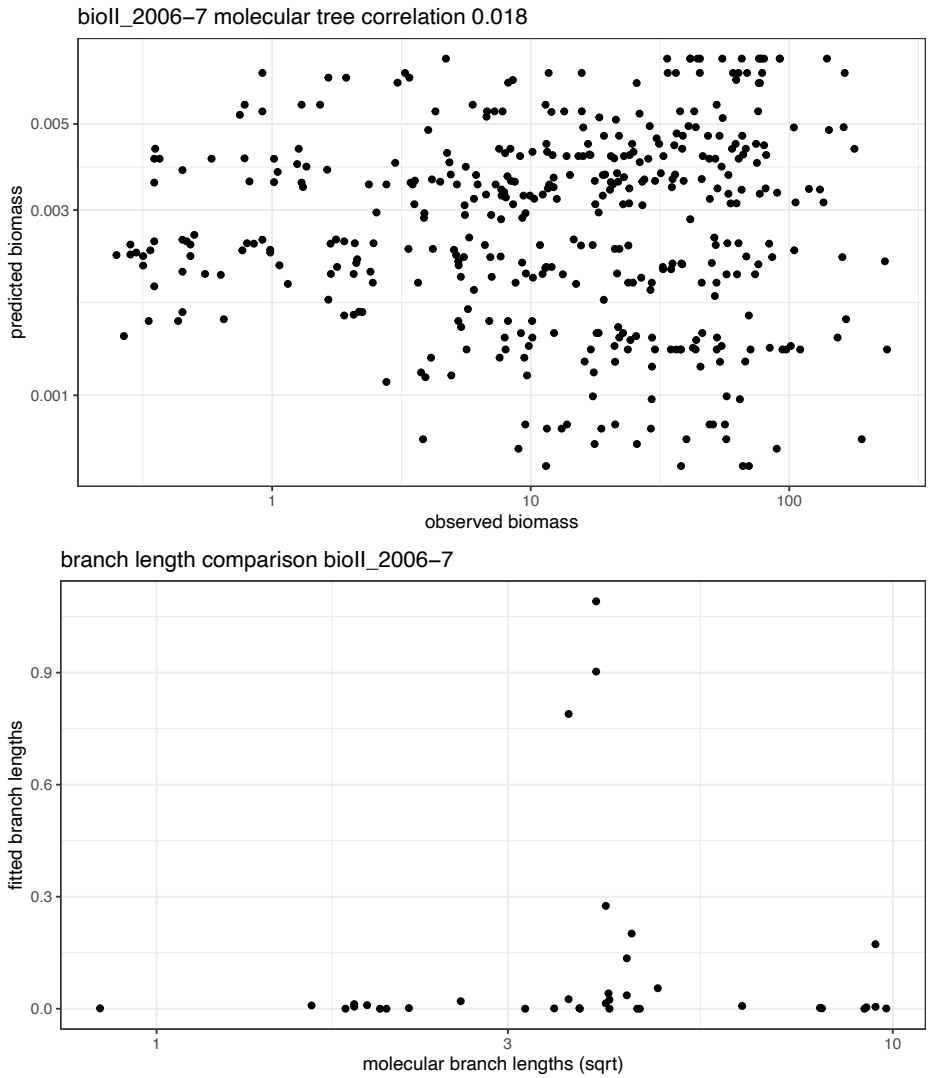

Figure 8: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

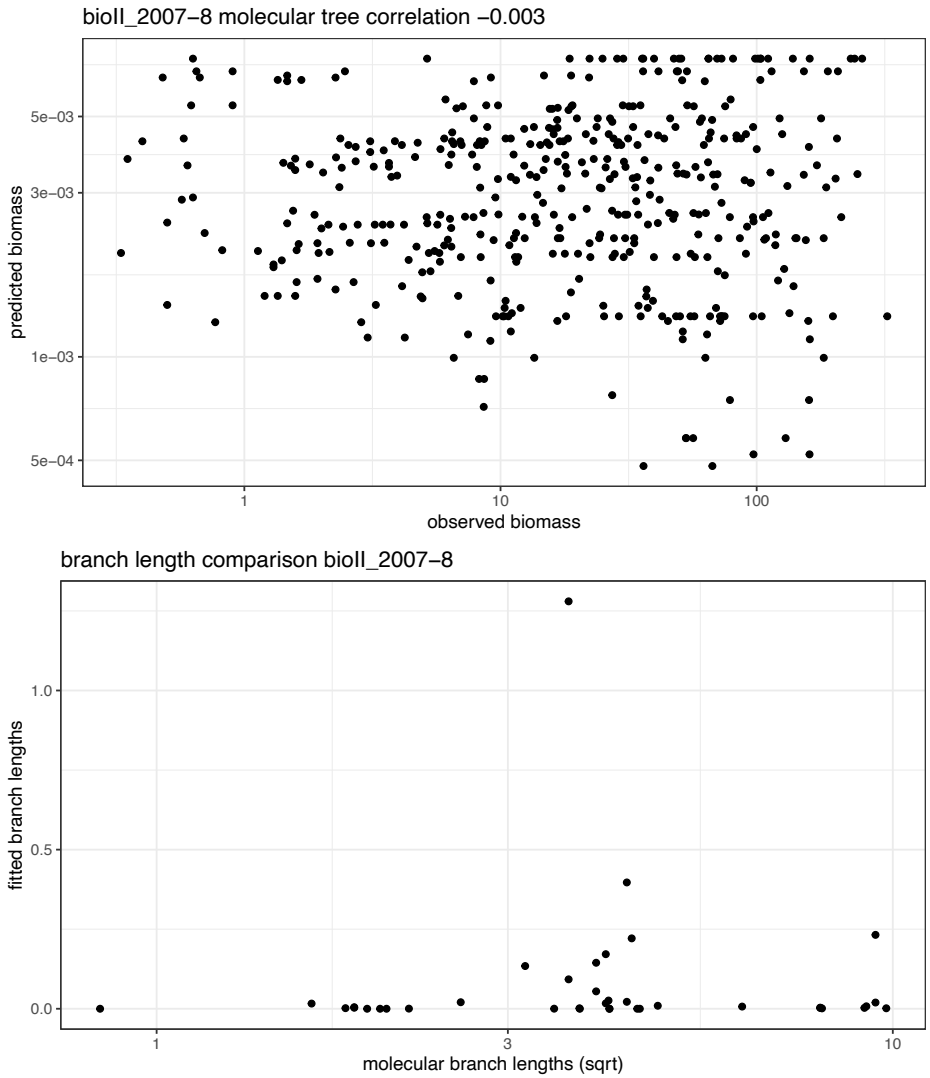

Figure 9: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

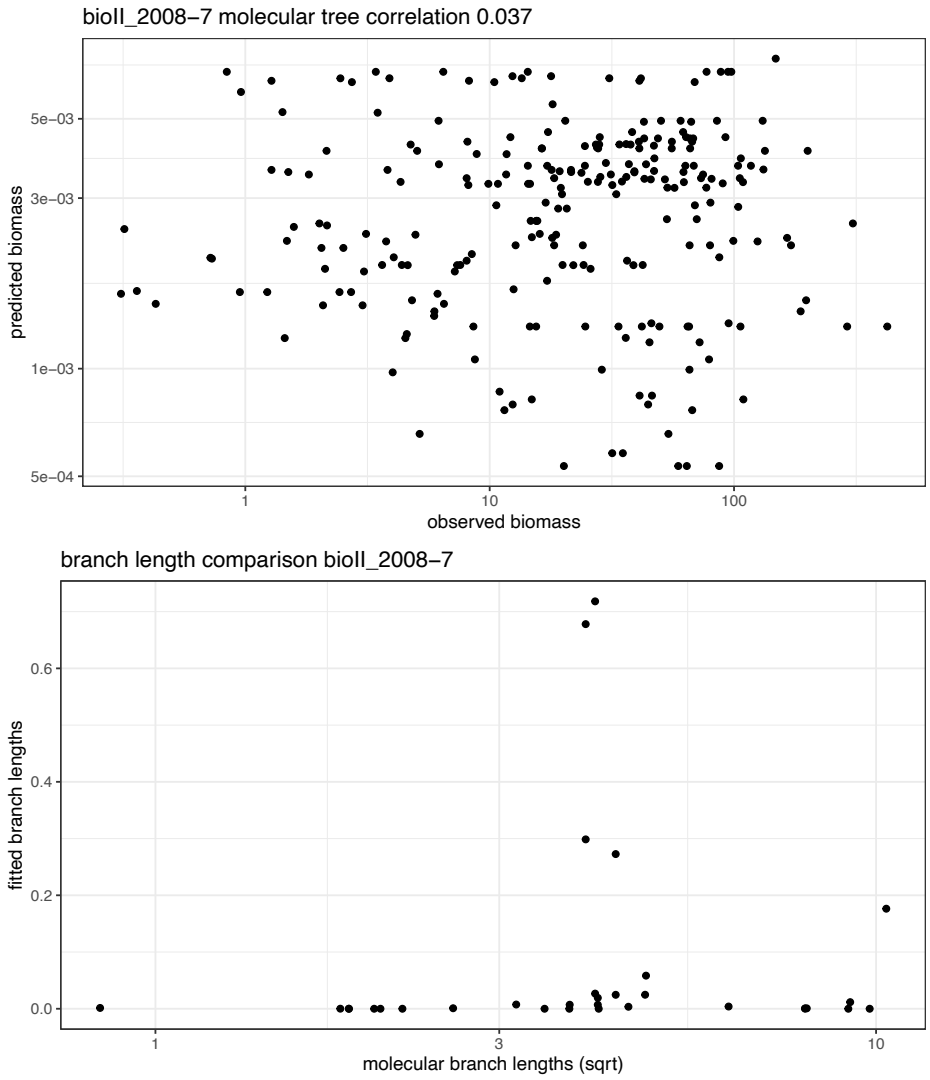

Figure 10: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

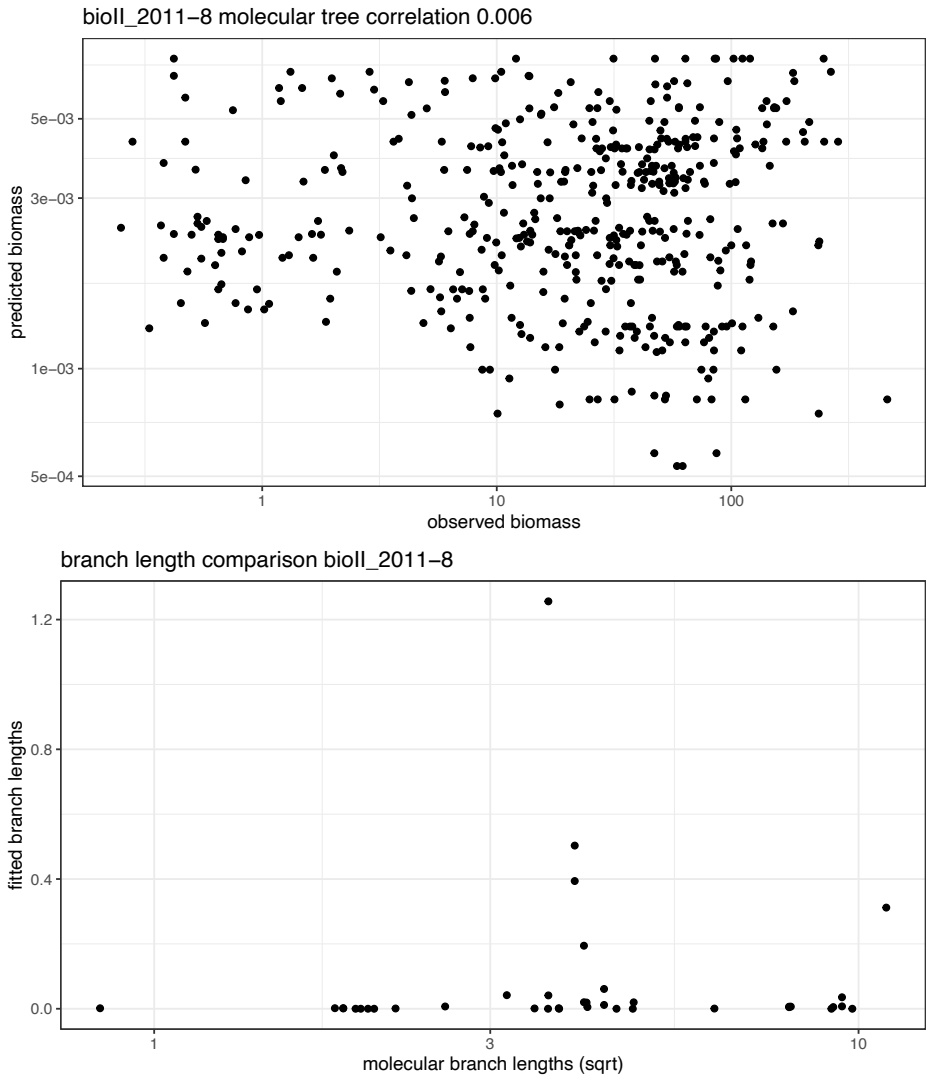

Figure 11: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

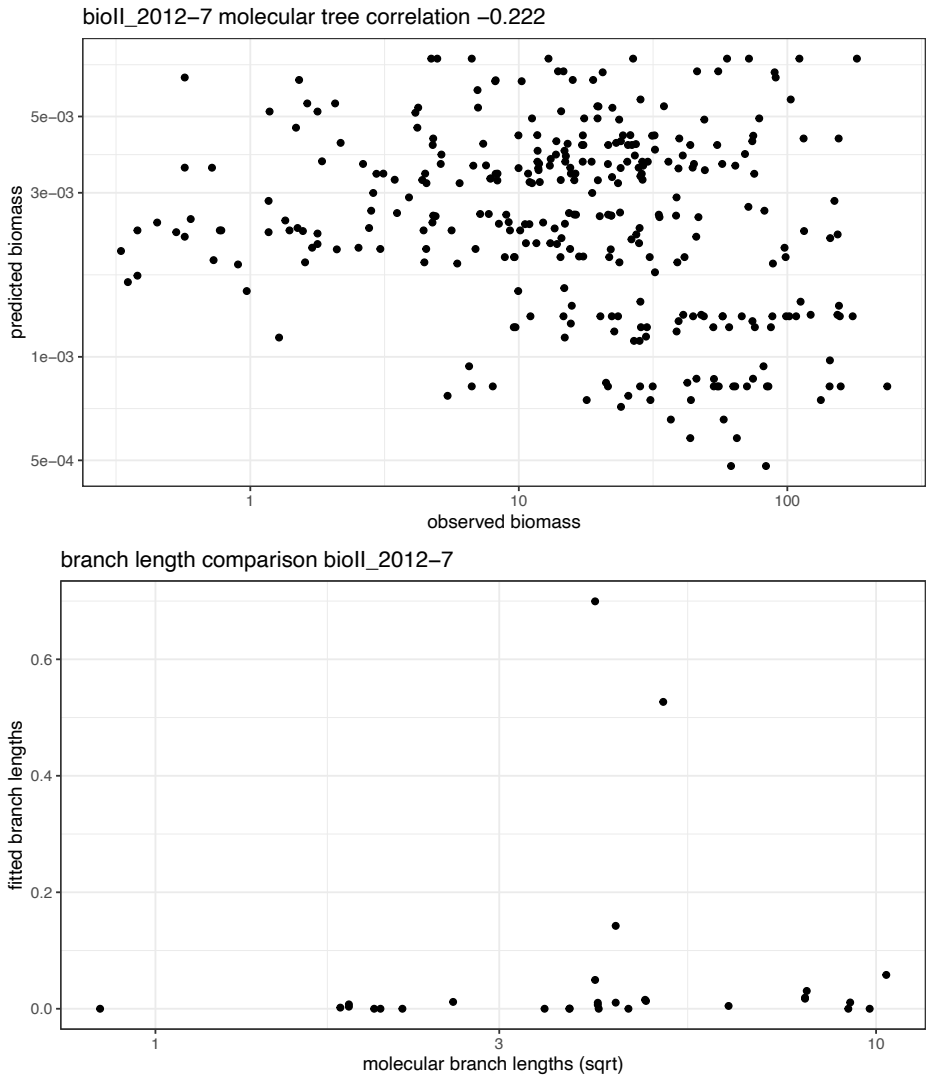

Figure 12: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

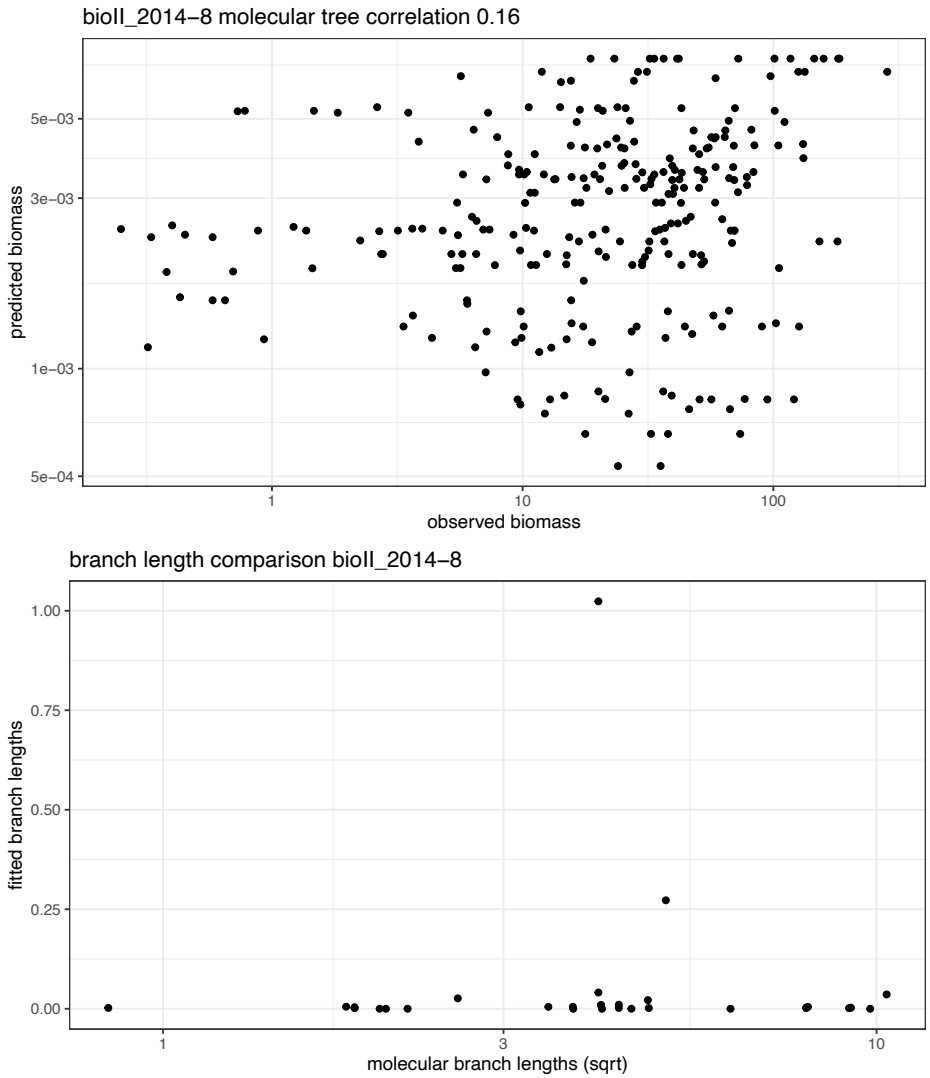

Figure 13: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

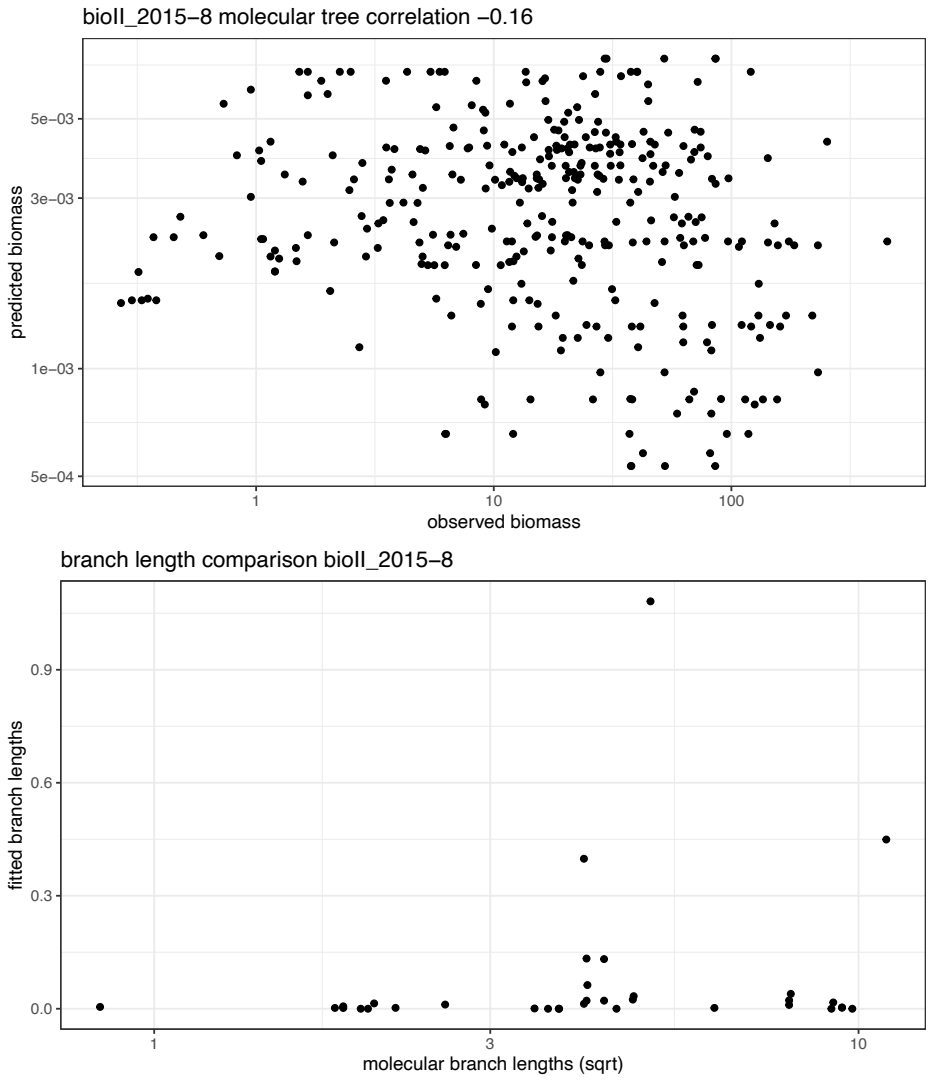

Figure 14: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

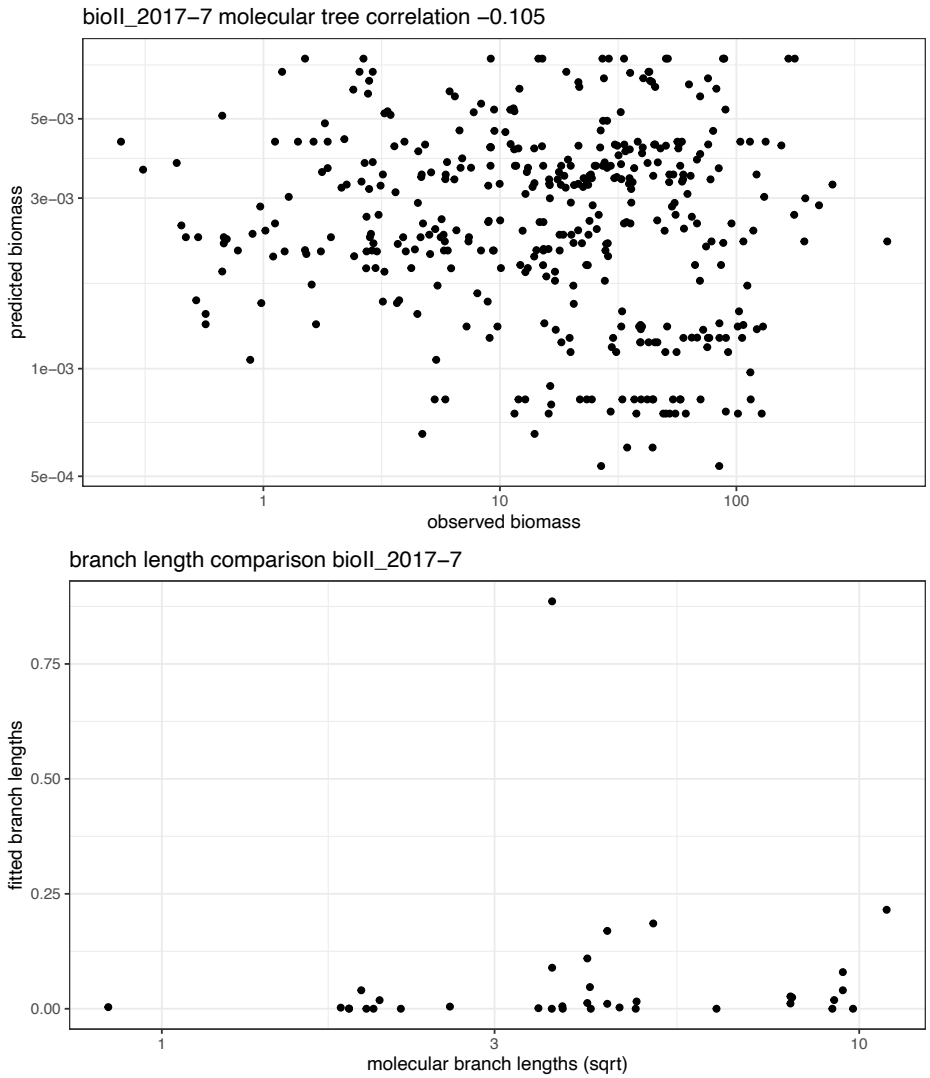

Figure 15: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

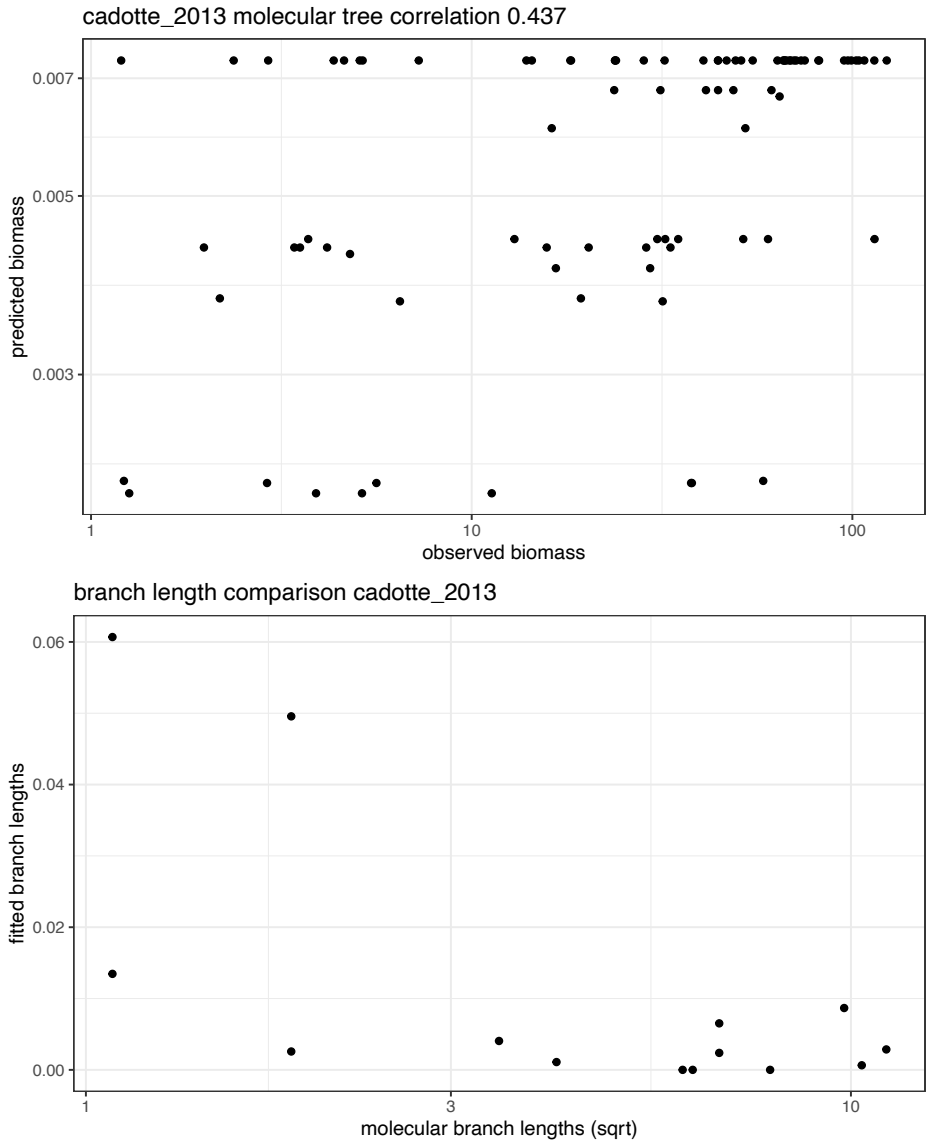

Figure 16: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

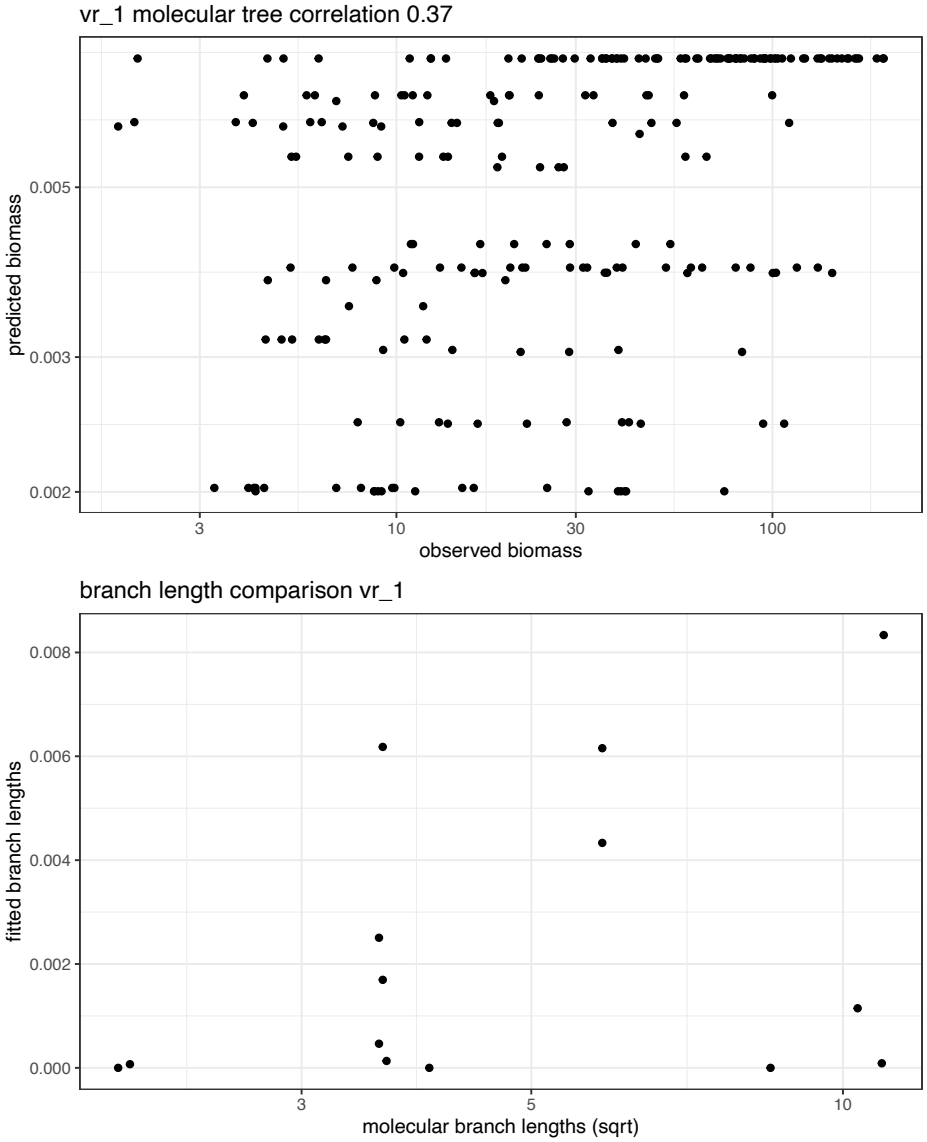

Figure 17: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

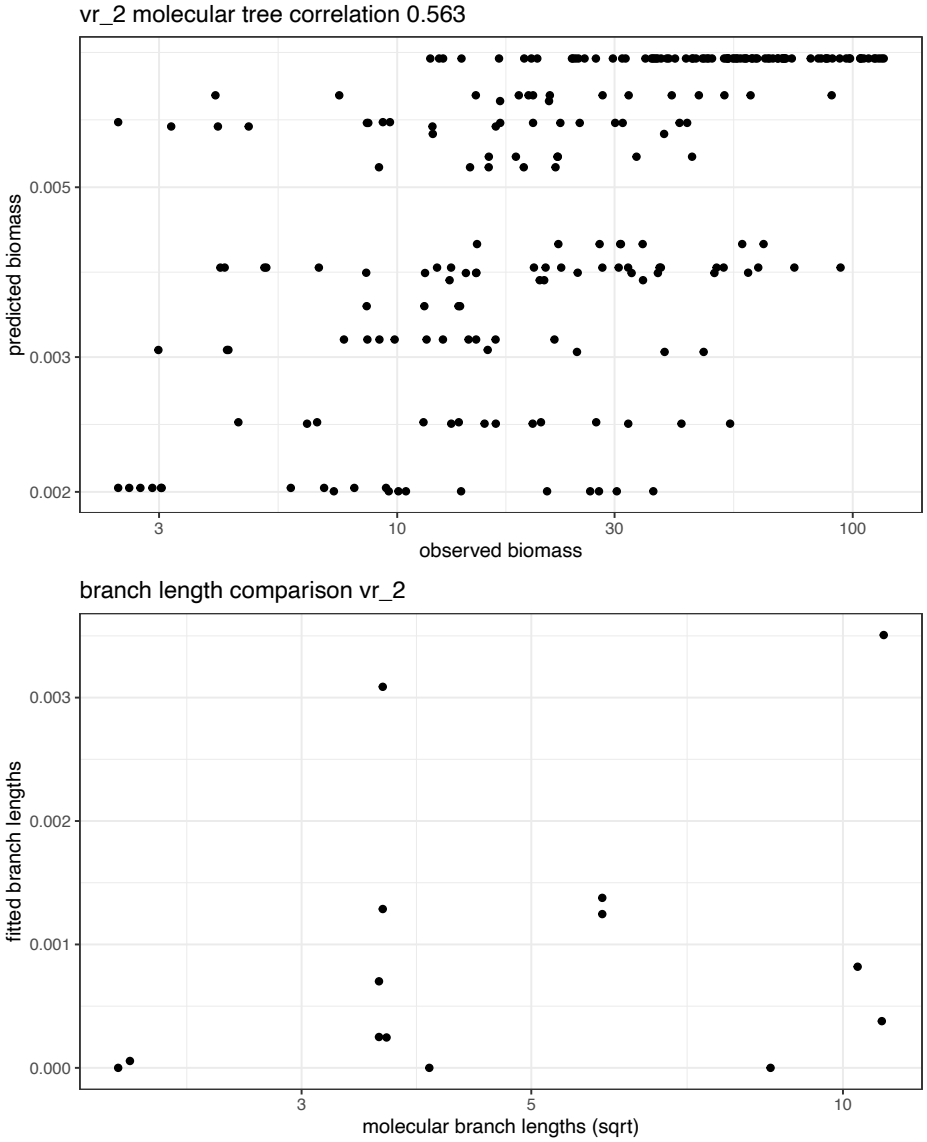

Figure 18: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

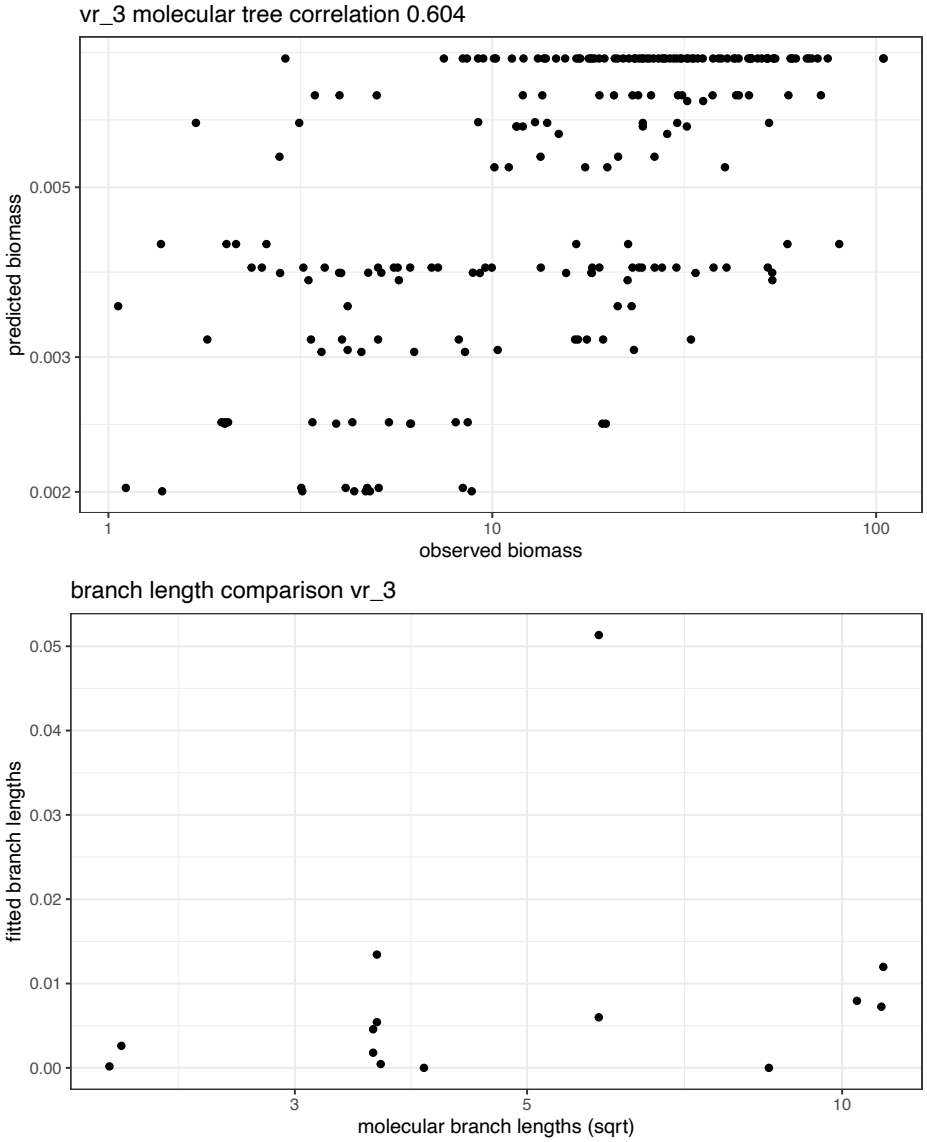

Figure 19: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

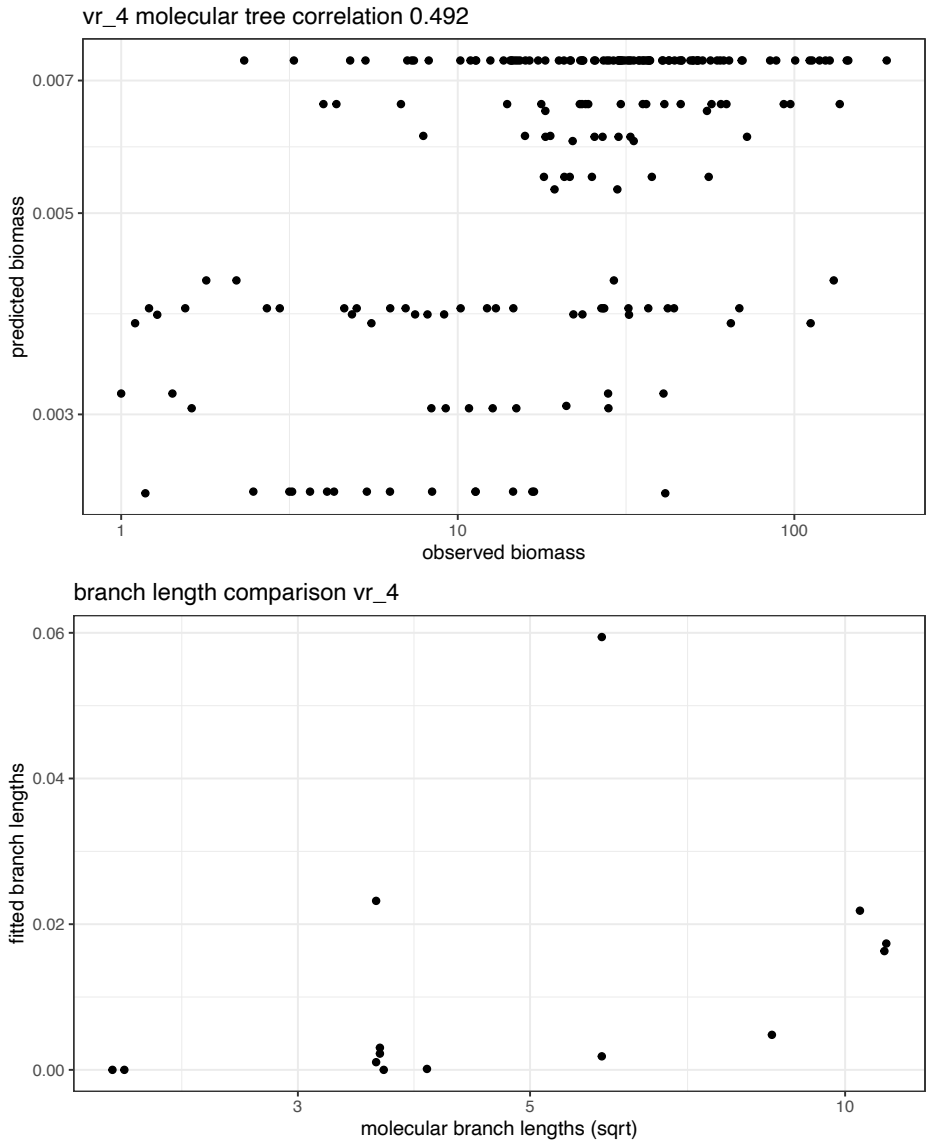

Figure 20: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

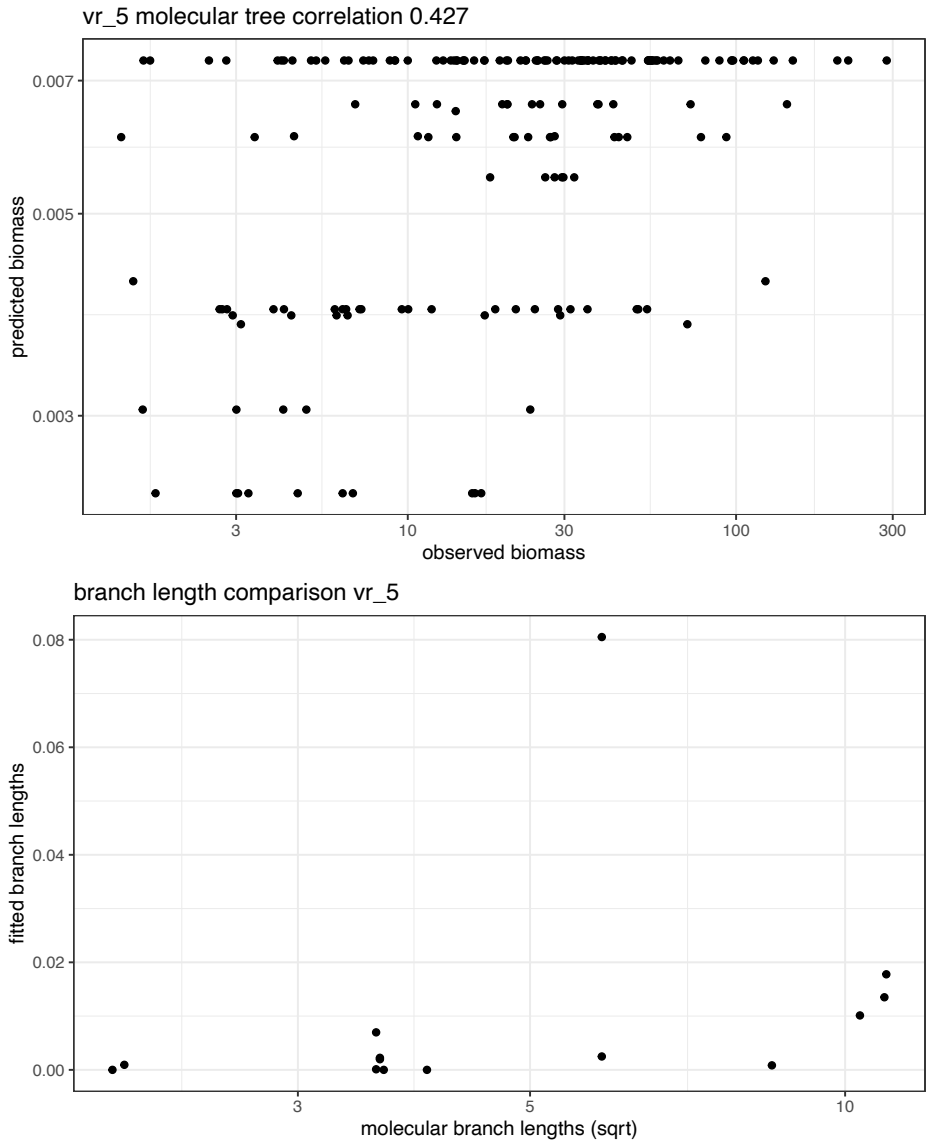

Figure 21: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

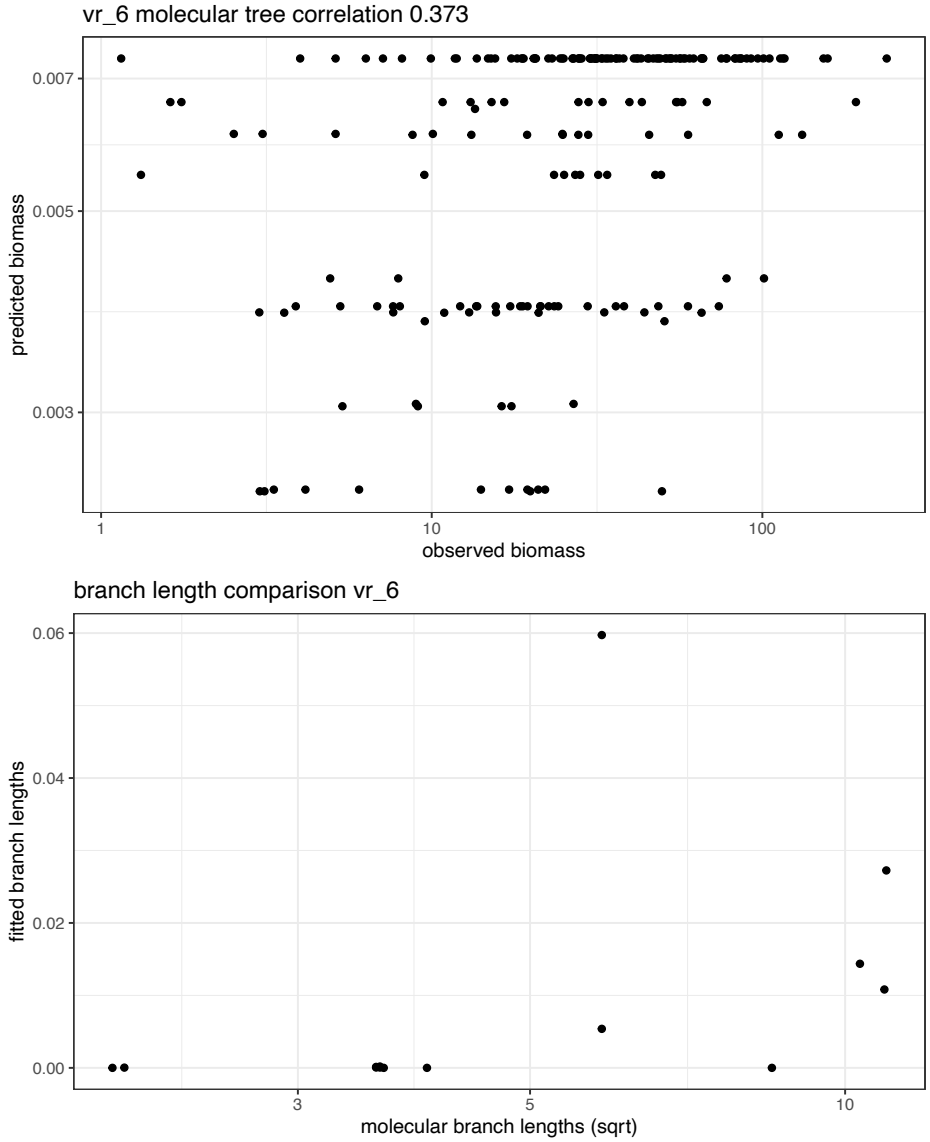

Figure 22: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

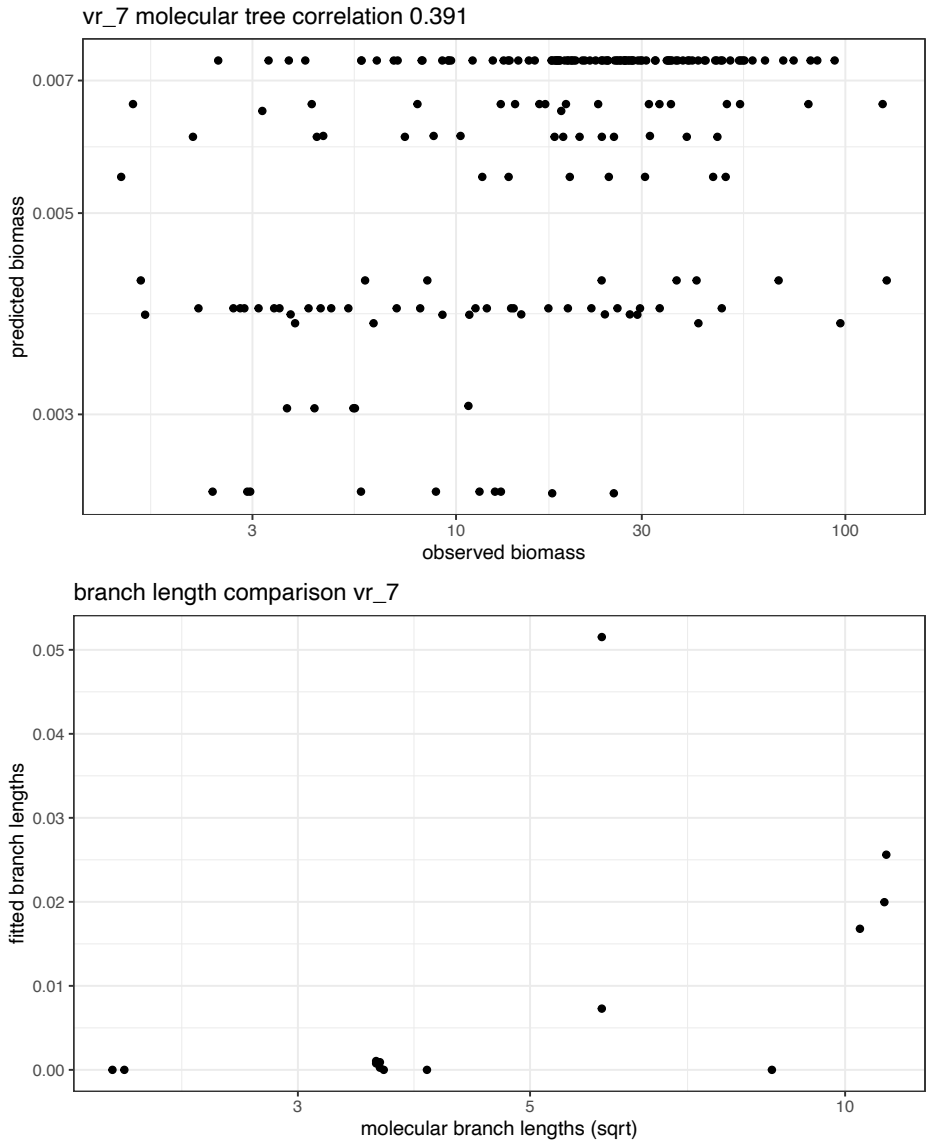

Figure 23: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

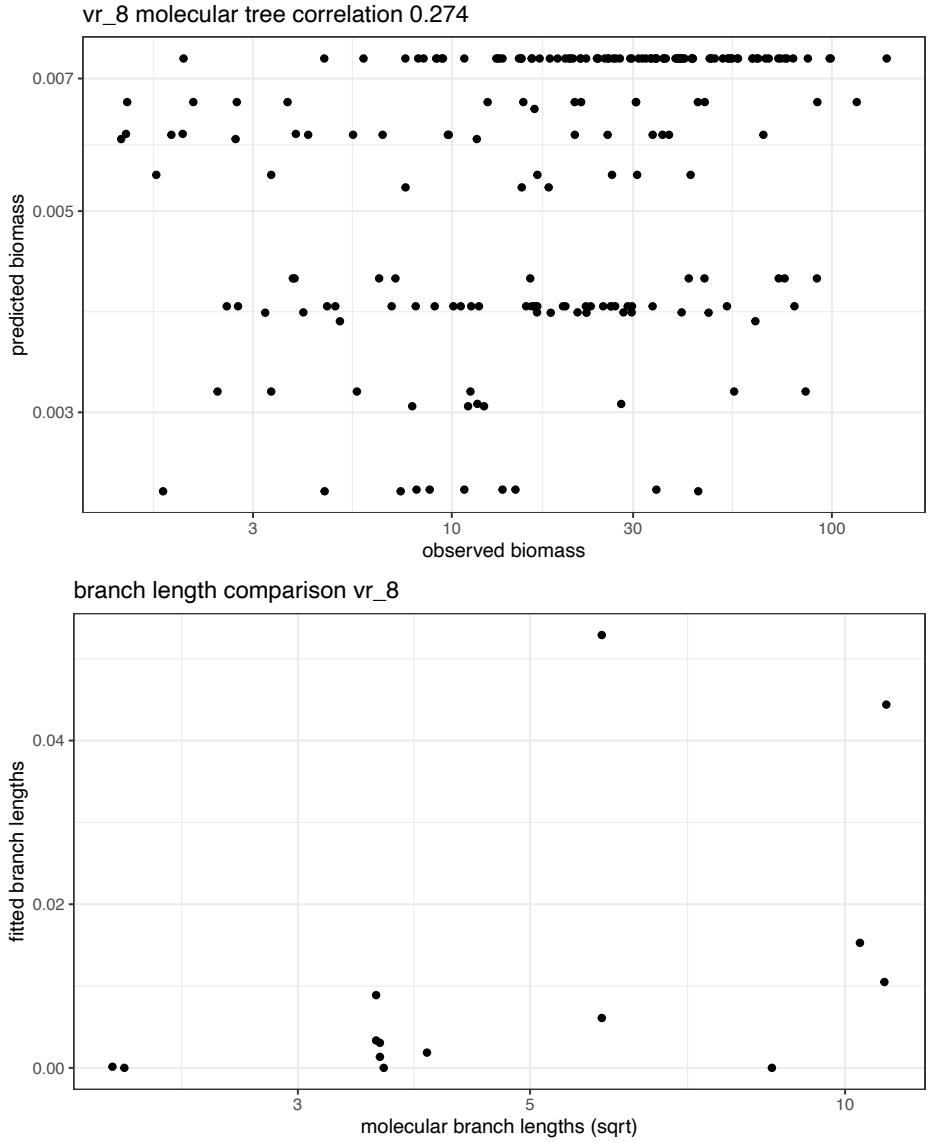

Figure 24: **Model predictions based on molecular branch lengths and comparision between fitted and molecular branch lengths**. Top panel shows predicted biomasses generated by parameterizing the model with the branch lengths from the molecular phylogeny of each data set as a function of the observed biomasses in experimental plots. BOx axes are in log scale and the correlation is estimate between the log of biomasses. The bottom panel shows the relationship between the branch lengths fitted according to our modeling approach as a function of the molecular branch lengths.

Analysis of previous approaches

Previous analysis of BEF data incorporating phylogenetic information to determine the role of phylogeny in structuring such experimental communities share two overarching features: a) they typically consider a single measurement for the function of each community (e.g., total biomass for each experimental plot, or the log-ratio of species abundances in multispecies communities to their abundances in monoculture), and b) they employ a statistical approach based on regressing these measurements against phylogenetic diversity or dissimilarity (Cadotte *et al.* 2008; Flynn *et al.* 2011; Venail *et al.* 2015). Here we show that by abandoning these two features we can perform a strong and principled test for phylogenetic structure in these data, and we find that indeed phylogeny structures species' interactions and thus patterns of coexistence and abundance.

Traditional approaches used to fit data from Biodiversity-Ecosystem Functioning experiments rely on regression models that consider the effects of biodiversity, in terms of species richness and composition, and some measurement of ecosystem function, such as biomass, carbon fixation, and litter decomposition. These models usually consider total biomass of each plot as the response variable, and regress it against the number of species or other properties of the plots; for example, they consider a design matrix of zeros and ones detailing the presence/absence of each species in a given plot. These models can be supplemented by measures of phylogenetic distance between the species, species' functional groups, species-specific effects, pairwise interactions, etc. Because of their underlying structure, these models do not take full advantage of the data, which usually report individual biomasses for each species in each plot—these data are aggregated when computing the total biomass per plot. These models suffer from three main limitations: (i) the use of aggregate data to fit the models when measurements are made at the species-level; (ii) lack of consistency with population dynamical models (but see Parain *et al.* 2019); and (iii) underlying assumptions regarding the error structure of the statistical models, which are not in accordance with the structure of the data or the assumptions of the model. Following the formulation by Connolly *et al.* (2011), a typical regression model might have the following structure:

$$y_i = \beta_0 + \sum_{j=1}^{S} \beta_j P_{ij} + \delta_{AV} \sum_{j<k}^{S} P_{ij} P_{ik} + \kappa C_{D_i} + \epsilon_i \tag{4}$$

where y_i is a measurement of ecosystem function for plot i (e.g., total biomass), S is the total number of species considered in the experiment (total species richness), β_0 is an intercept, β_j is the effect of species j on function in monoculture, P_{ij} is an indicator variable for the presence (or absence) of species j in plot i, δ_{AV} is the average interaction effect, C_{D_i} is the phylogenetic diversity in plot i, computed as $C_{D_i} = \sum_{j<k}^{S} (D_{jk} - \bar{D}) P_j P_k$, where D_{jk} is the phylogenetic distance between species j and k (accounted as the sum of branch lengths between the two species) and $\bar{D}$ is the average distance of all species in the pool), and finally ϵ_i is the error term.

The amount of data needed to fit such a model can be very large. To examine the performance of this kind of regression model, we therefore use simulations where data is generated in agreement with basic biological assumptions and test the model with the simulated data. Here we simulate ecological communities of competing species following the Generalized Lotka-Volterra dynamics, with the matrix of interactions given by the variance-covariance matrix corresponding to a phylogenetic tree. This should be an ideal situation for fitting the model, as interactions are constant and strictly pairwise, with a clear phylogenetic structure. Starting from a pool of $n = 10$ species, we assume all species have the same intrinsic growth rates, so that dynamics are determined solely by phylogenetic relations. We draw a random tree, construct its variance-covariance matrix, assemble all $2^n - 1 = 1023$ communities that can be formed from the species pool and record each species' biomass in each community. The structure of the matrix of interactions can be thought of as emerging from models of trait evolution; for example, if we follow the evolution of many trait values, each changing under Brownian motion, this process gives rise to the trait variance-covariance matrix which we associated with competition coefficients. In this way, differences in trait values are proportional to evolutionary distances (Hansen & Martins 1996; Ives & Helmus 2011). One key aspect

of this parameterization is that all sub-communities that can be formed with a subset of species are feasible and dynamically stable (see Serván *et al.* 2023 for details).

Having produced a large synthetic data set containing more than a thousand "plots", we can assess the quality of fit using a variety of regression models of varying complexity. Model $M0 : y \sim P$ regresses total biomass against the presence/absence of each species (P); model $M1 : y \sim P + C_D$ regresses the total biomass against presence/absence of species (P) and the phylogenetic distance of species in the plot (C_D); $M2 : y \sim P + C_D + P : C_D$ is the same as $M1$ but also considers the interaction between presence and phylogenetic distance; $M3 : y \sim P + \sum P_i P_j$ considers the presence absence of each species and their pairwise interaction; $M4 : y \sim P + \sum P_i P_j + C_D$ is the same as $M3$ adding the phylogenetic distance of each plot; and finally $M5 : y \sim P + \sum P_i P_j + C_D + P : C_D$ is the same as $m4$ including the interaction between presence and phylogenetic distance. Below we show the plots for each of the models, including the predicted versus observed biomasses and diagnostic plots from linear regression.

These results show that a) none of the models can fully capture the simulated data (despite the lack of noise); b) all of the model fits violate some of the main assumptions of linear regressions: residuals are not independent and normally distributed, variances are not constant, and thus Quantile-Quantile plots deviate substantially from the 1:1 line. We conclude that these model structures are not adequate to capture models for population dynamics—generalizing beyond competition, choosing more complex models for population dynamics, or introducing stochasticity and measurement error can only worsen the fit.

The inadequacy of regression-based approaches motivates our alternative methodology for inferring phylogenetic structure in these data.

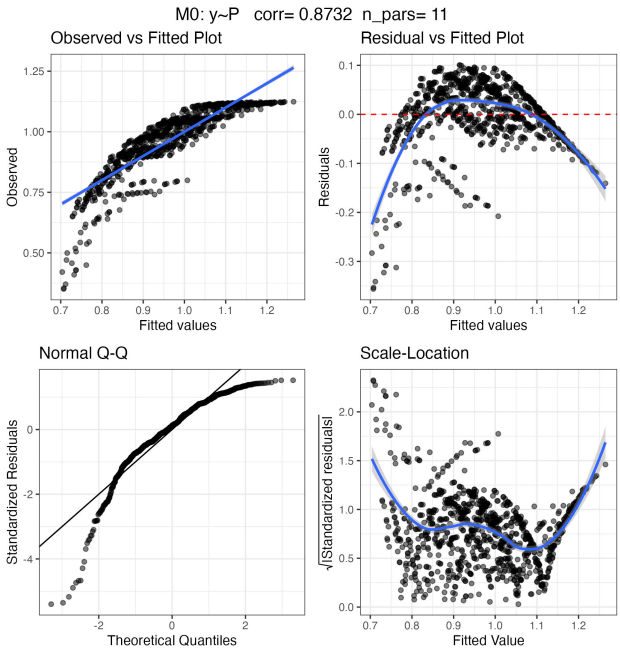

Figure 25: Fitting regression models to simulated data from a phylogenetically-structured Generalized Lotka-Volterra model. Each figure reports the structure of the model (top) as well as scatter-plots showing the relationship between observed and fitted total biomasses per plot; corresponding distribution of standardized residuals (for linear regression, these should be centered at zero and have the same variance irrespective of the fitted values); Quantile-Quantile plots (points should fall on the 1 to 1 line); scale location plots, showing whether residuals change with the input (they should form a horizontal band of points).

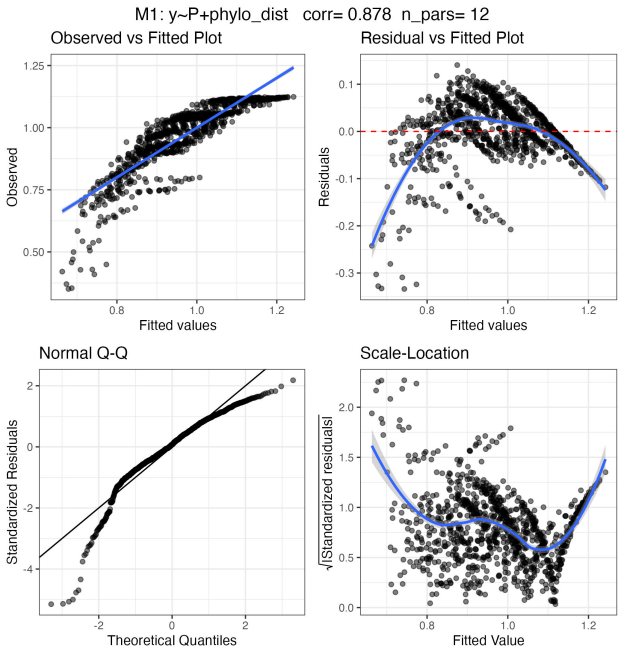

Figure 26: Fitting regression models to simulated data from a phylogenetically-structured Generalized Lotka-Volterra model. Each figure reports the structure of the model (top) as well as scatter-plots showing the relationship between observed and fitted total biomasses per plot; corresponding distribution of standardized residuals (for linear regression, these should be centered at zero and have the same variance irrespective of the fitted values); Quantile-Quantile plots (points should fall on the 1 to 1 line); scale location plots, showing whether residuals change with the input (they should form a horizontal band of points).

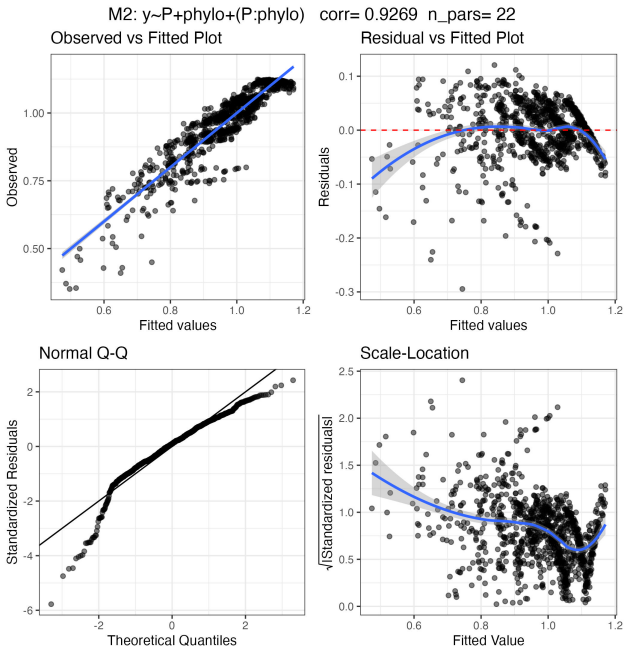

Figure 27: Fitting regression models to simulated data from a phylogenetically-structured Generalized Lotka-Volterra model. Each figure reports the structure of the model (top) as well as scatter-plots showing the relationship between observed and fitted total biomasses per plot; corresponding distribution of standardized residuals (for linear regression, these should be centered at zero and have the same variance irrespective of the fitted values); Quantile-Quantile plots (points should fall on the 1 to 1 line); scale location plots, showing whether residuals change with the input (they should form a horizontal band of points).

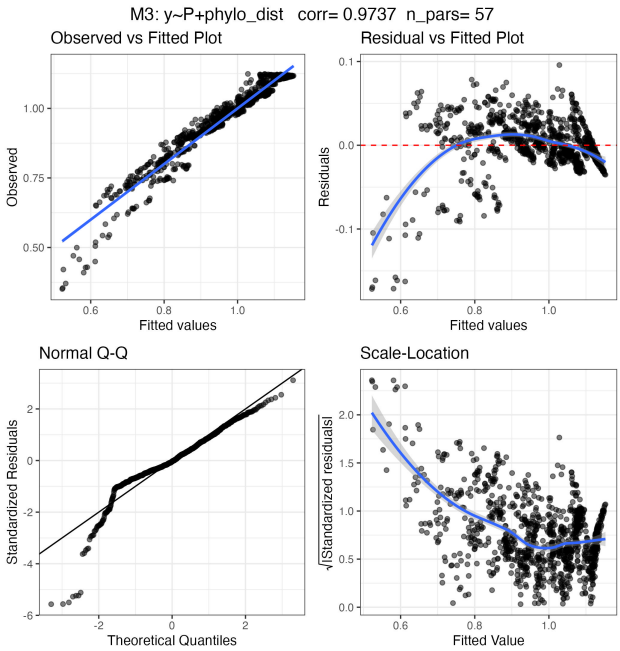

Figure 28: Fitting regression models to simulated data from a phylogenetically-structured Generalized Lotka-Volterra model. Each figure reports the structure of the model (top) as well as scatter-plots showing the relationship between observed and fitted total biomasses per plot; corresponding distribution of standardized residuals (for linear regression, these should be centered at zero and have the same variance irrespective of the fitted values); Quantile-Quantile plots (points should fall on the 1 to 1 line); scale location plots, showing whether residuals change with the input (they should form a horizontal band of points).

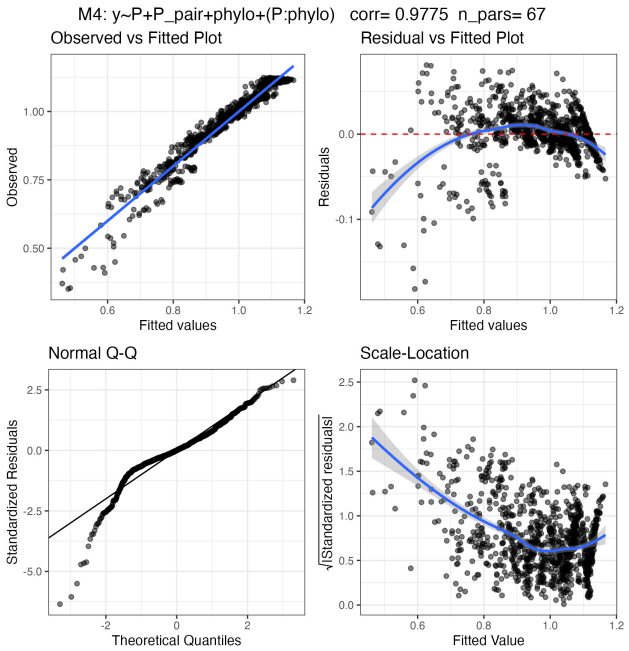

Figure 29: Fitting regression models to simulated data from a phylogenetically-structured Generalized Lotka-Volterra model. Each figure reports the structure of the model (top) as well as scatter-plots showing the relationship between observed and fitted total biomasses per plot; corresponding distribution of standardized residuals (for linear regression, these should be centered at zero and have the same variance irrespective of the fitted values); Quantile-Quantile plots (points should fall on the 1 to 1 line); scale location plots, showing whether residuals change with the input (they should form a horizontal band of points).

Full results

The results for all model formulations and each data set are reported in Table 2. Scatter-plots for the data stemming from the Biodiversity II and the Wageningen experiments are reported in Figure 30 and Figure 32, respectively. Histograms of the likelihoods for the randomized trees are reported in Figure 31 and Figure 33. The fitted trees and associated likelihood boxplots for randomizations stratified by selected clades are reported in Figure 34 - Figure 41 for the Biodiversity II experiment and Figure 42 - Figure 49 for the Wageningen experiment. Finally, the fits for models 2-4 with the associated tree and vectors are also in Figures 50-68.

Table 2: For each data set (Bio II, from Tilman et al. (2001); Cadotte, from Cadotte (2013); Wag, from Van Ruijven and Berendse (2010)), we report the number of species (n_{sp}); the number of unique assemblages (n_{comm}); the total number of measurements (n_{meas}); the model used to fit the data; the corresponding log-likelihood (logLik); the total number of parameters (n_{par}); the p-values for randomized trees either produced by the Yule model (p_{rnd}) or by shuffling the species' identities ($p_{shuffle}$); and the correlation between the logarithms of the observed and predicted biomasses (cor). P-values below 0.05 are bolded.

data set	n_{sp}	n_{comm}	n_{meas}	model	logLik	n_{par}	p_{rnd}	$p_{shuffle}$	cor
Bio II 6/01	13	28	142	1	-515.125	26	**0.011**	**0.006**	0.916
				2	-506.803	38	**0.001**	**0.004**	0.924
				3	-510.384	38	0.606	0.612	0.923
				4	-496.525	50	0.153	0.137	0.936
Bio II 7/06	20	63	416	1	-1 562.862	40	**0.032**	**0.034**	0.848
				2	-1 543.540	59	**0.028**	**0.032**	0.857
				3	-1 548.530	59	**0.013**	**0.028**	0.856
				4	-1 528.100	78	**0.001**	**0.007**	0.866
Bio II 8/07	20	69	445	1	-1 829.444	40	**0.017**	**0.014**	0.829
				2	-1 810.571	59	**0.020**	**0.018**	0.842
				3	-1 816.684	59	**0.033**	0.056	0.840
				4	-1 791.473	78	**0.049**	**0.041**	0.852
Bio II 7/08	17	33	247	1	-1 019.927	34	**0.008**	**0.008**	0.833
				2	-1 005.384	50	**0.045**	**0.049**	0.848
				3	-1 014.225	50	**0.030**	**0.024**	0.842
				4	-989.244	66	**0.014**	**0.008**	0.854
Bio II 8/11	20	59	451	1	-1 872.468	40	0.147	0.188	0.865
				2	-1 851.060	59	**0.037**	**0.033**	0.876
				3	-1 845.121	59	**0.002**	**0.005**	0.880
				4	-1 822.608	78	**0.000**	**0.000**	0.888
Bio II 7/12	16	40	322	1	-1 265.636	32	0.253	0.261	0.814
				2	-1 257.352	47	0.141	0.118	0.821
				3	-1 256.901	47	0.193	0.216	0.822
				4	-1 239.699	62	**0.009**	**0.009**	0.843

data set	n_{sp}	n_{comm}	n_{meas}	model	logLik	n_{par}	p_{rnd}	$p_{shuffle}$	cor
Bio II 8/14	16	39	271	1	-1 100.731	32	**0.000**	**0.001**	0.832
				2	-1 089.324	47	**0.027**	**0.019**	0.844
				3	-1 093.747	47	**0.006**	**0.003**	0.838
				4	-1 077.425	62	**0.012**	**0.013**	0.856
Bio II 8/15	18	43	337	1	-1 344.643	36	**0.005**	**0.012**	0.830
				2	-1 335.333	53	**0.002**	**0.004**	0.835
				3	-1 333.390	53	**0.027**	**0.028**	0.840
				4	-1 315.121	70	**0.004**	**0.009**	0.854
Bio II 7/17	19	53	409	1	-1 599.741	38	**0.001**	**0.001**	0.841
				2	-1 587.881	56	**0.011**	**0.014**	0.850
				3	-1 580.571	56	**0.002**	**0.000**	0.851
				4	-1 562.737	74	**0.001**	**0.000**	0.863
Cadotte	8	27	87	1	-348.372	16	**0.027**	**0.029**	0.878
				2	-344.637	23	0.131	0.132	0.890
				3	-345.628	23	0.512	0.507	0.881
				4	-342.545	30	0.696	0.711	0.891
Wag 00	8	58	235	1	-856.445	16	0.419	0.378	0.948
				2	-850.295	23	0.456	0.435	0.951
				3	-823.108	23	0.246	0.248	0.960
				4	-817.592	30	0.260	0.297	0.962
Wag 01	8	58	231	1	-815.679	16	**0.001**	**0.001**	0.933
				2	-811.568	23	0.103	0.115	0.935
				3	-786.491	23	**0.000**	**0.001**	0.948
				4	-780.650	30	0.071	0.066	0.950
Wag 02	8	56	222	1	-770.436	16	0.067	0.066	0.858
				2	-756.722	23	**0.040**	**0.024**	0.879
				3	-754.152	23	0.068	0.071	0.877
				4	-726.884	30	**0.003**	**0.003**	0.908
Wag 03	8	51	196	1	-754.342	16	**0.043**	**0.043**	0.849
				2	-737.109	23	**0.036**	**0.018**	0.873
				3	-731.931	23	0.076	**0.027**	0.880
				4	-713.488	30	**0.030**	**0.028**	0.903
Wag 04	8	53	184	1	-731.749	16	0.145	0.140	0.796
				2	-721.201	23	0.115	0.088	0.823
				3	-699.034	23	**0.005**	**0.003**	0.853

data set	n_{sp}	n_{comm}	n_{meas}	model	logLik	n_{par}	p_{rnd}	$p_{shuffle}$	cor
				4	-690.901	30	0.214	0.215	0.871
Wag 05	8	53	197	1	-816.254	16	0.060	**0.049**	0.766
				2	-807.928	23	0.125	0.071	0.786
				3	-804.602	23	**0.028**	**0.020**	0.802
				4	-792.838	30	0.380	0.387	0.820
Wag 06	8	54	197	1	-746.003	16	0.106	0.104	0.721
				2	-740.846	23	0.239	0.243	0.741
				3	-720.442	23	**0.028**	**0.028**	0.790
				4	-714.023	30	0.166	0.166	0.803
Wag 07	8	55	206	1	-808.122	16	**0.023**	**0.032**	0.769
				2	-799.968	23	0.148	0.152	0.788
				3	-791.224	23	**0.036**	0.075	0.807
				4	-786.148	30	0.080	0.077	0.809
Wag 08	8	54	210	1	-803.785	16	0.278	0.282	0.780
				2	-792.913	23	0.986	0.961	0.797
				3	-794.166	23	0.236	0.271	0.799
				4	-779.235	30	1.000	1.000	0.822

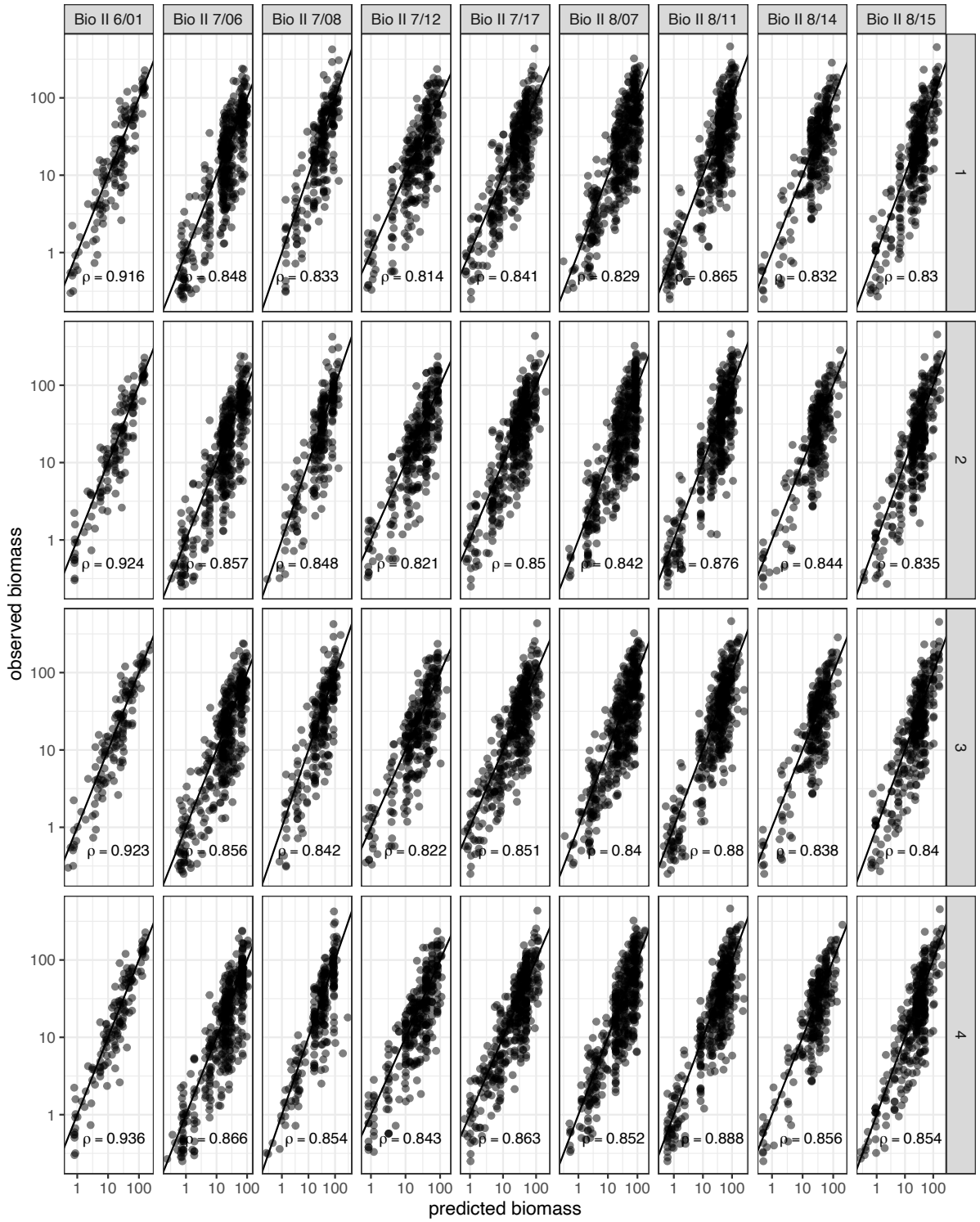

Figure 30: Scatterplot of predicted vs. observed biomass for the Biodiversity II data. Each panel represents a specific sampling date and model. The correlation between the logarithm of the observed biomass and the logarithm of the predicted biomass is reported in each panel.

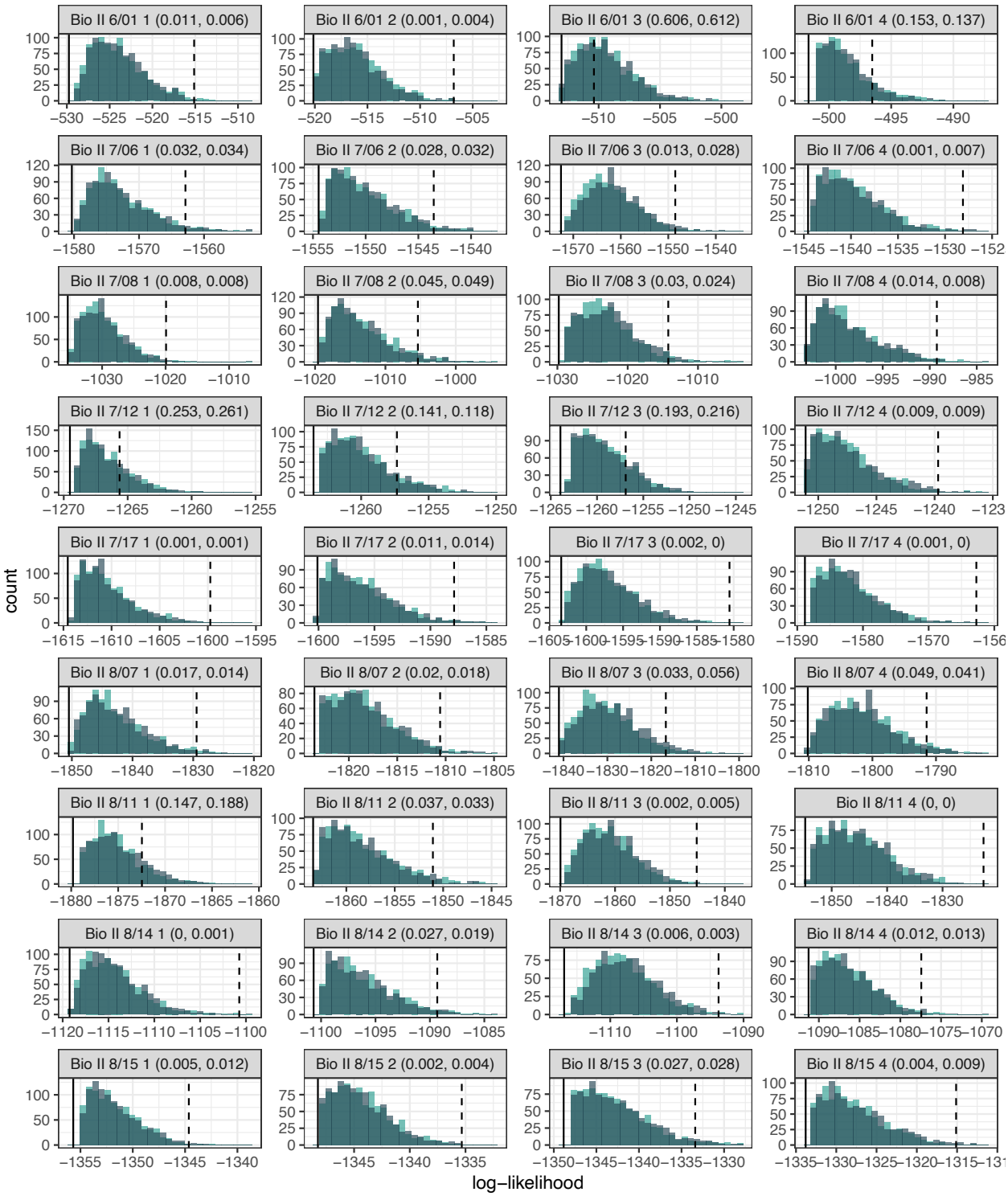

Figure 31: Histograms of the log-likelihoods for models parameterized using random trees and data from the Biodiversity II experiment. The distributions for the random trees generated by the Yule model are in teal, and those for the shuffled trees in grey. Each panel reports the p-values obtained for the Yule and shuffled trees, respectively.

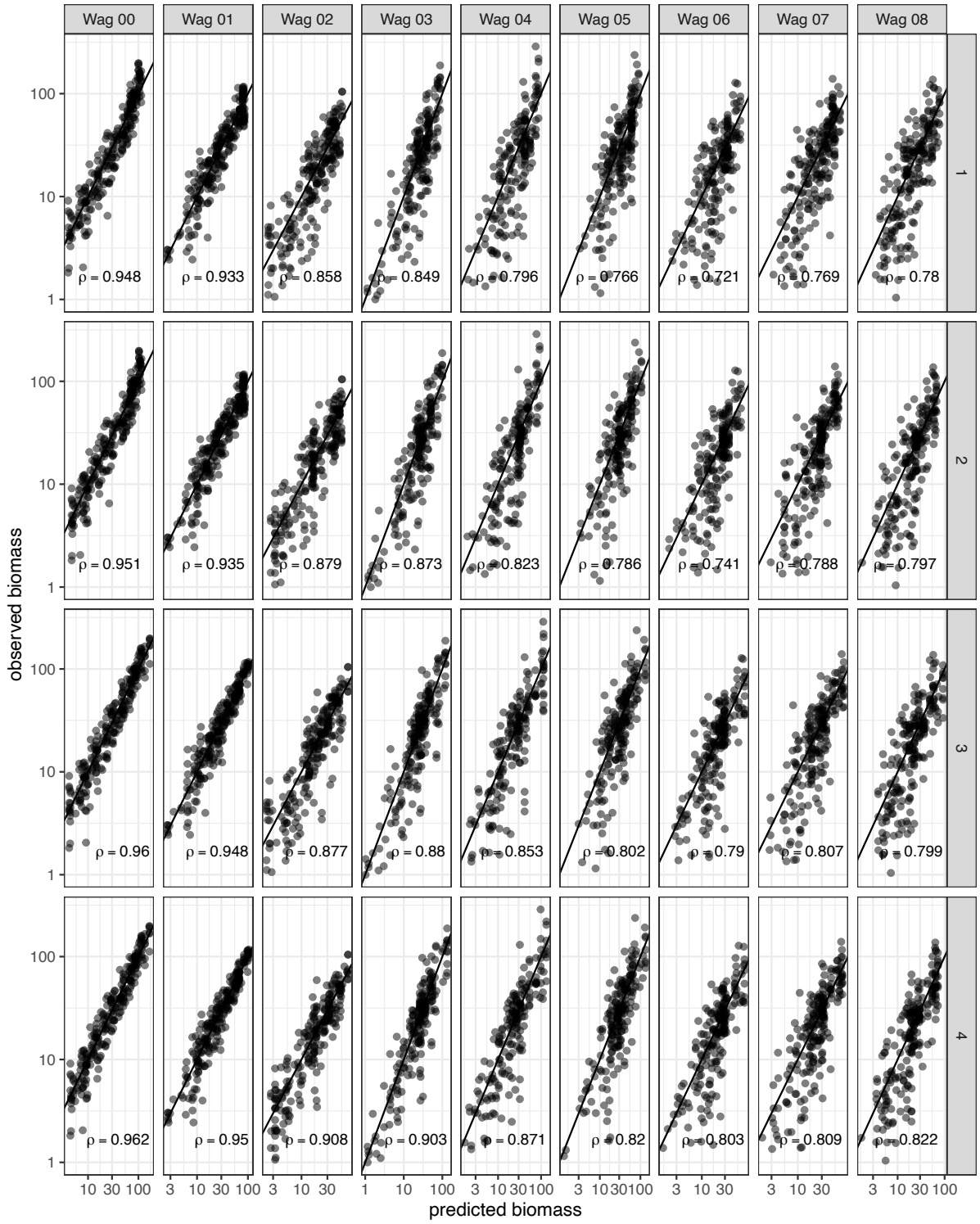

Figure 32: Scatterplot of predicted vs. observed biomass for the Wageningen biodiversity experiment data. Each panel represents a specific sampling date and model. The correlation between the logarithm of the observed biomass and the logarithm of the predicted biomass is reported in each panel.

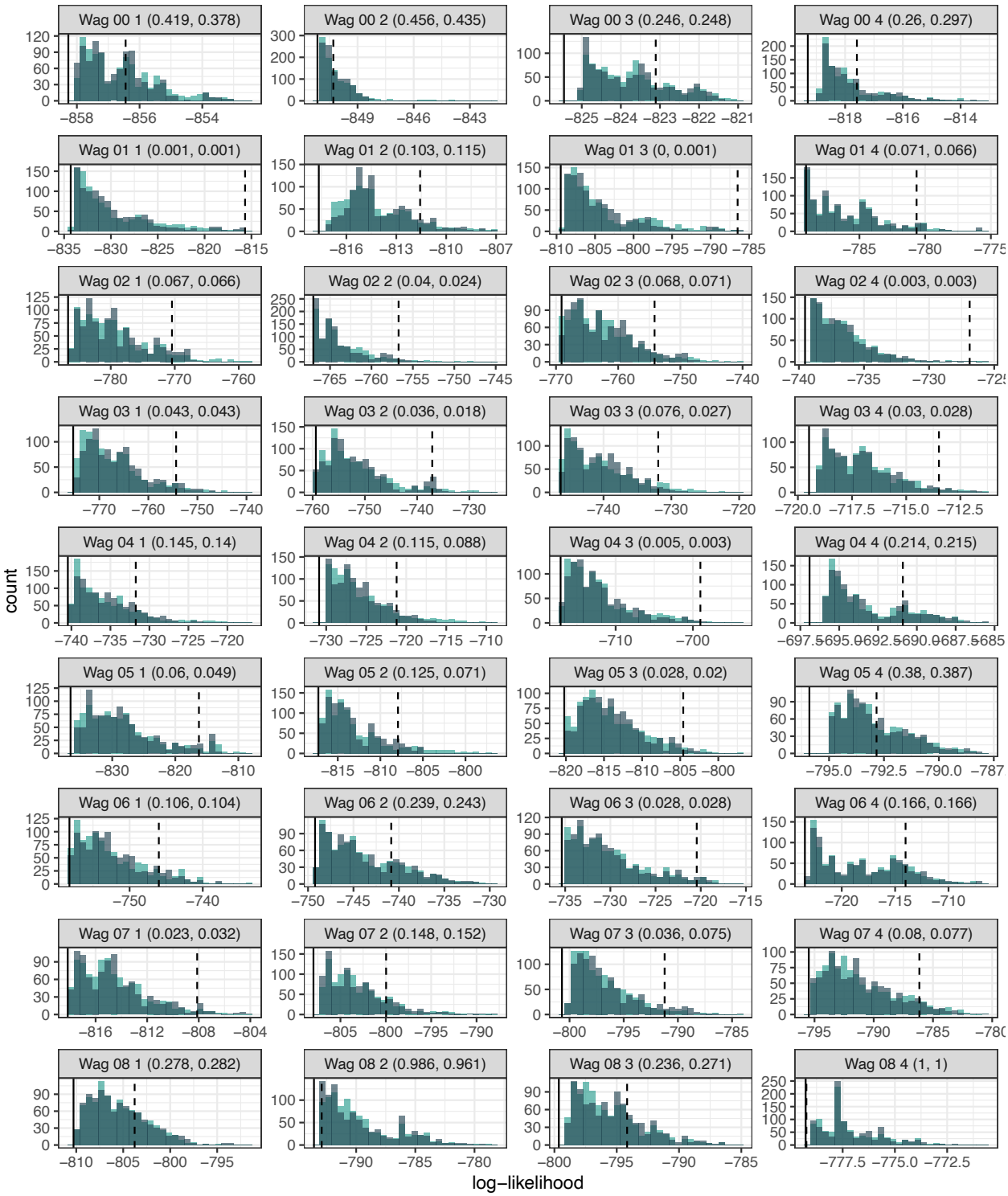

Figure 33: Histograms of the log-likelihoods for models parameterized using random trees and data from the Wageningen diversity experiment. The distributions for the random trees generated by the Yule model are in teal, and those for the shuffled trees in grey. Each panel reports the p-values obtained for the Yule and shuffled trees, respectively.

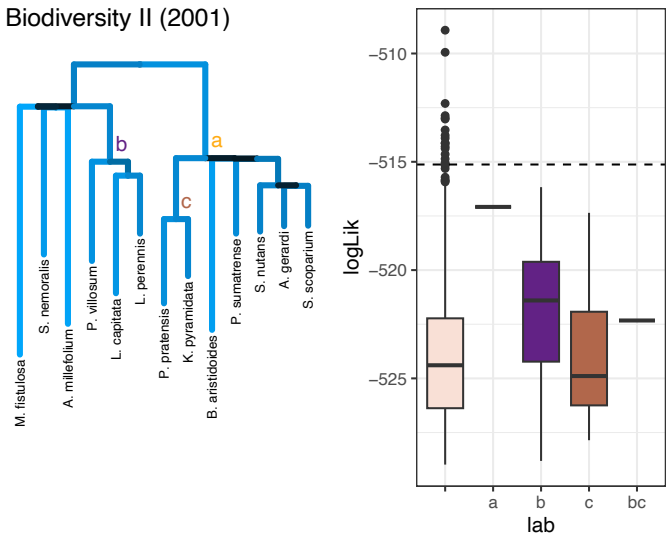

Figure 34: Resulting trees for the Biodiversity II experiments in Cedar Creek, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

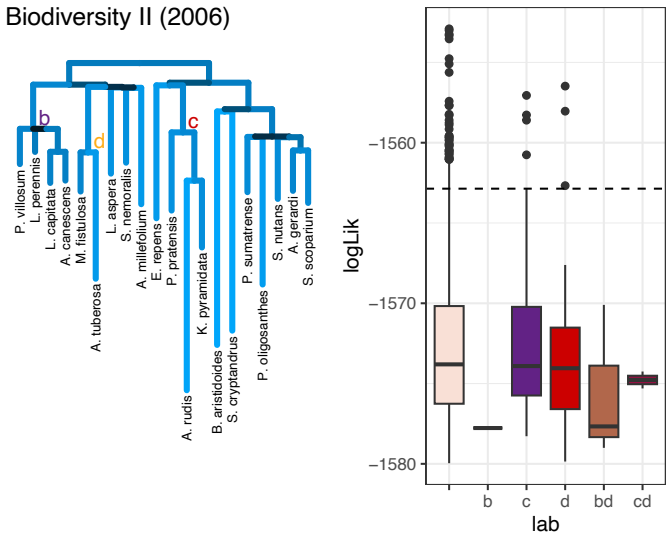

Figure 35: Resulting trees for the Biodiversity II experiments in Cedar Creek, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

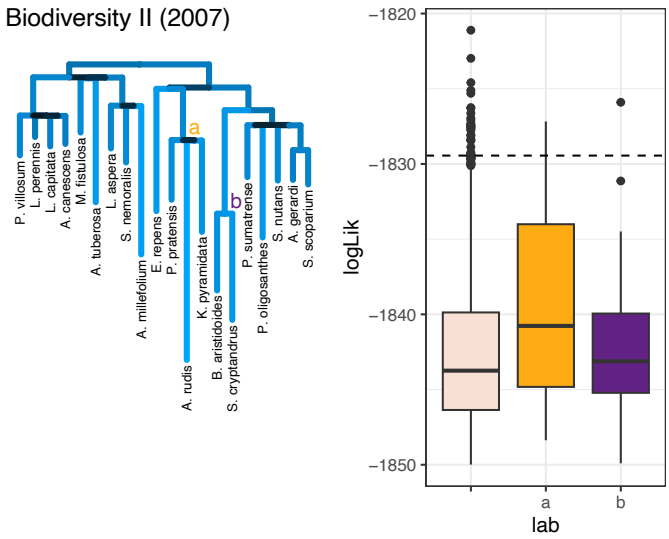

Figure 36: Resulting trees for the Biodiversity II experiments in Cedar Creek, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

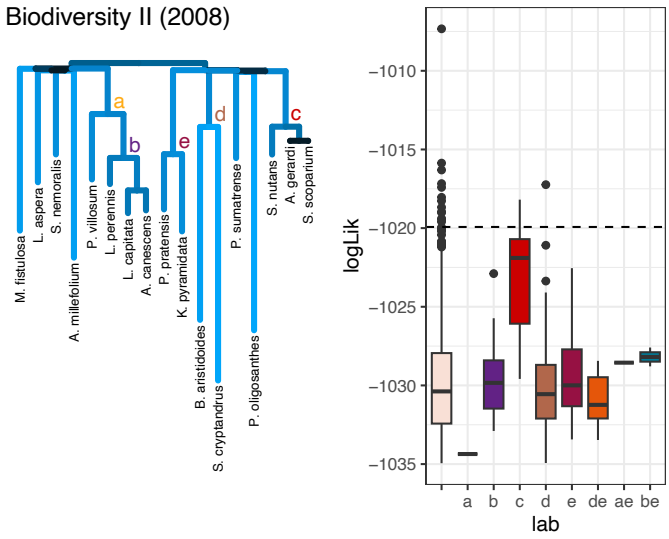

Figure 37: Resulting trees for the Biodiversity II experiments in Cedar Creek, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

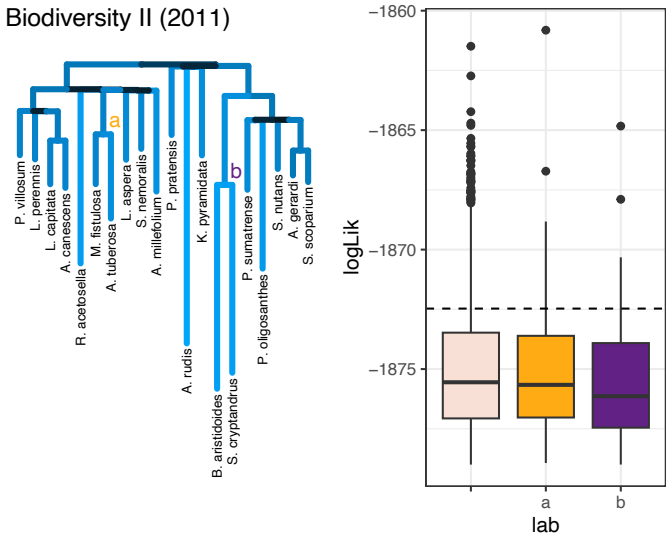

Figure 38: Resulting trees for the Biodiversity II experiments in Cedar Creek, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

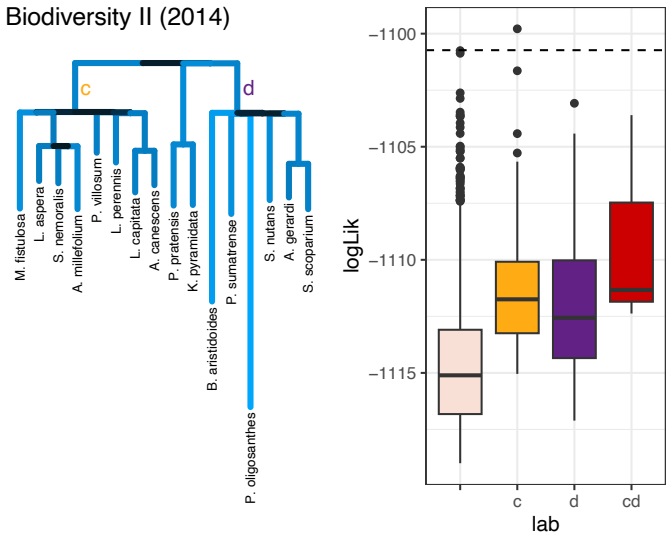

Figure 39: Resulting trees for the Biodiversity II experiments in Cedar Creek, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

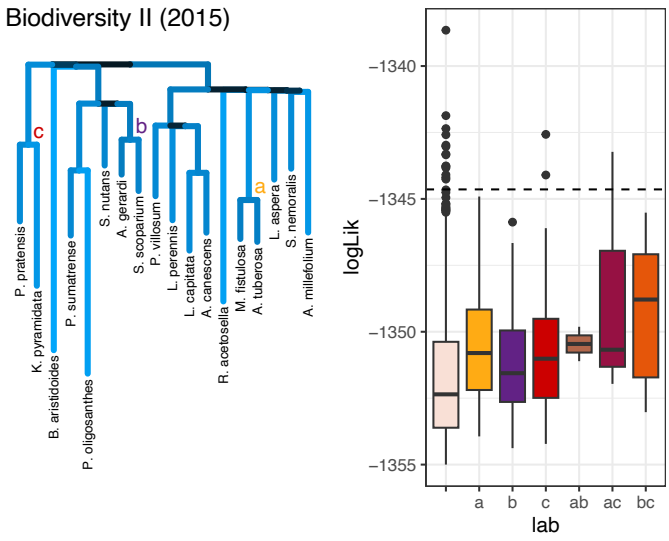

Figure 40: Resulting trees for the Biodiversity II experiments in Cedar Creek, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

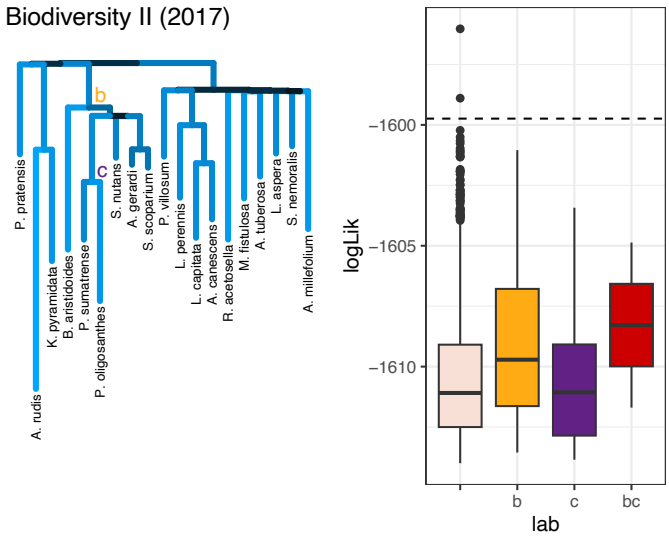

Figure 41: Resulting trees for the Biodiversity II experiments in Cedar Creek, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

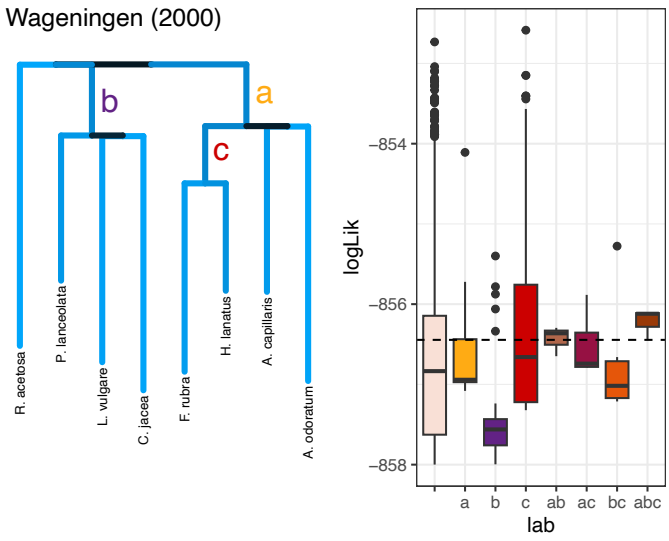

Figure 42: Resulting trees for the Wageningen experiments, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

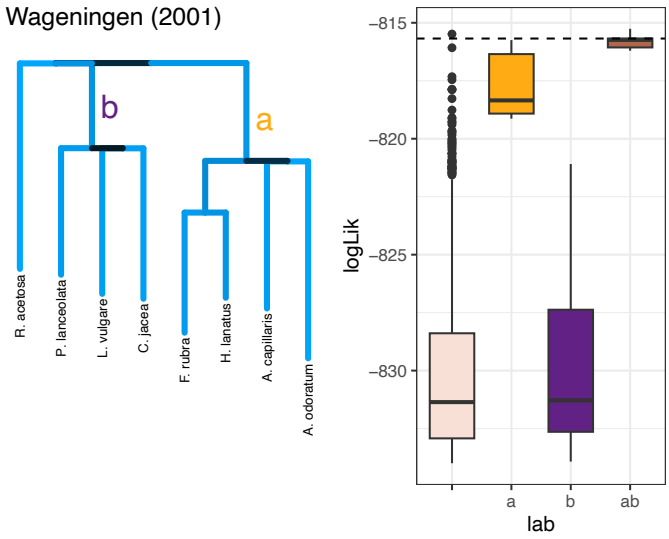

Figure 43: Resulting trees for the Wageningen experiments, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

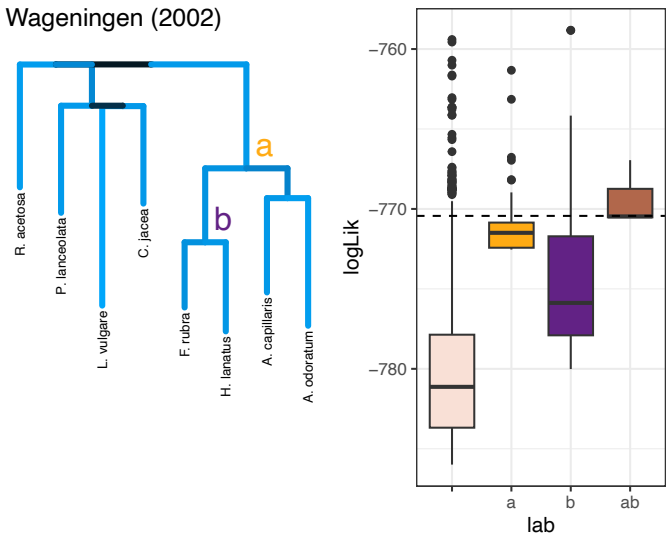

Figure 44: Resulting trees for the Wageningen experiments, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

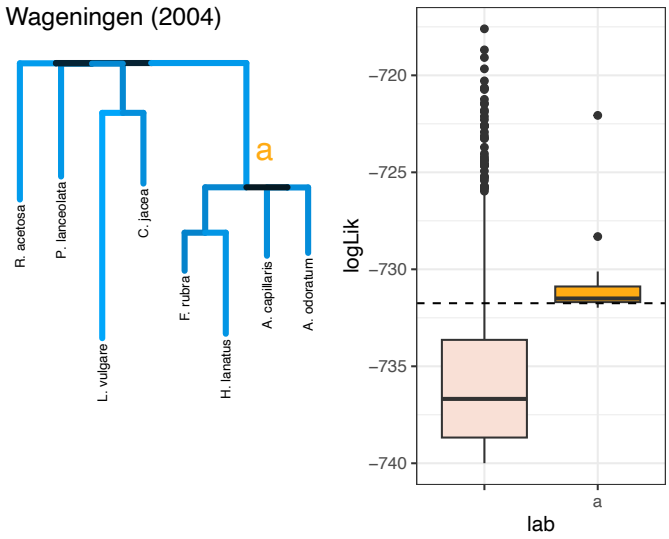

Figure 45: Resulting trees for the Wageningen experiments, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

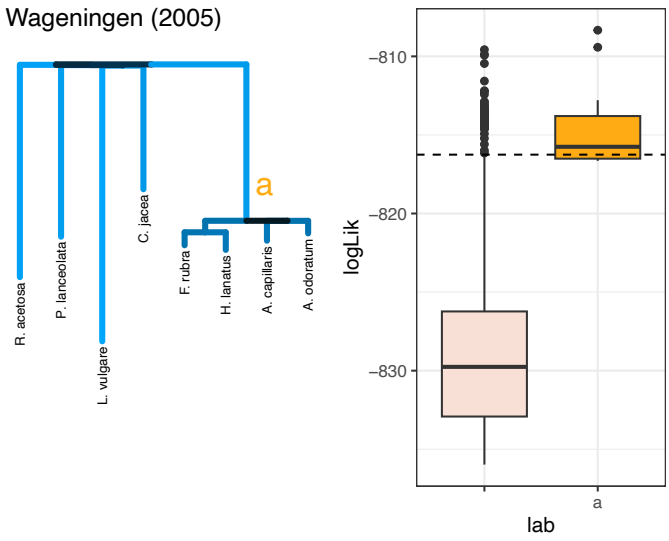

Figure 46: Resulting trees for the Wageningen experiments, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

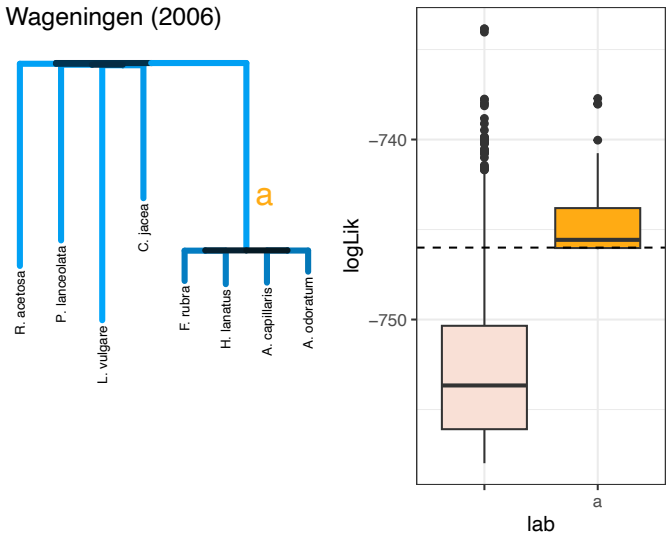

Figure 47: Resulting trees for the Wageningen experiments, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

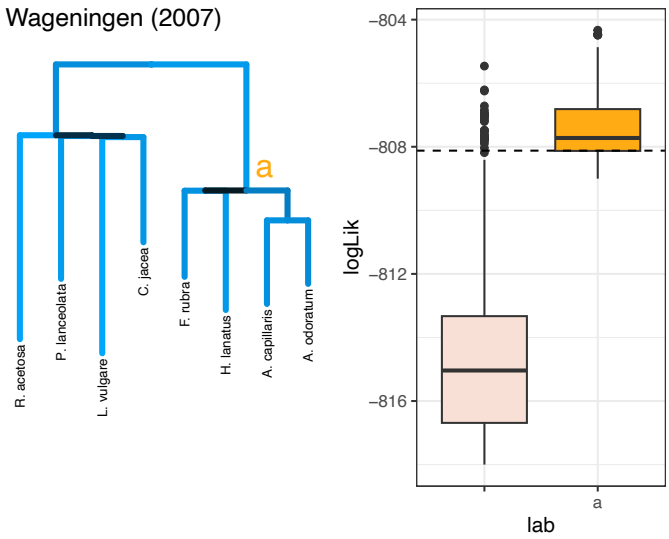

Figure 48: Resulting trees for the Wageningen experiments, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

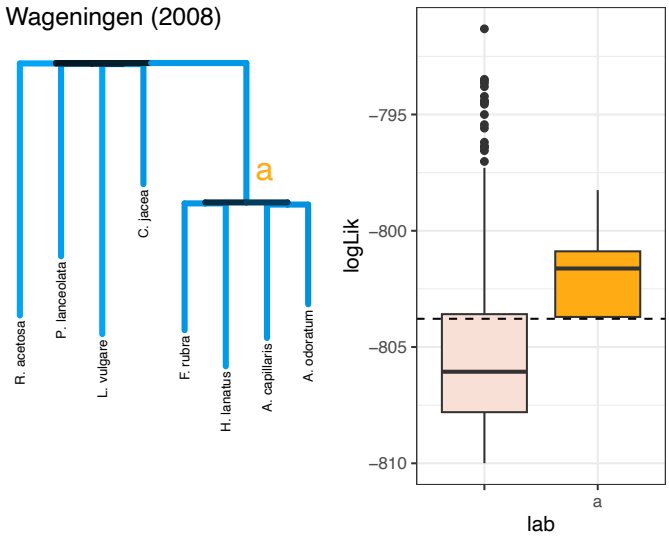

Figure 49: Resulting trees for the Wageningen experiments, fit for phylogenetic tree from model 1 and boxplots of likelihoods for randomizations that kept the selected clades highlighted in the trees. Dashed line represents likelihood of the original tree

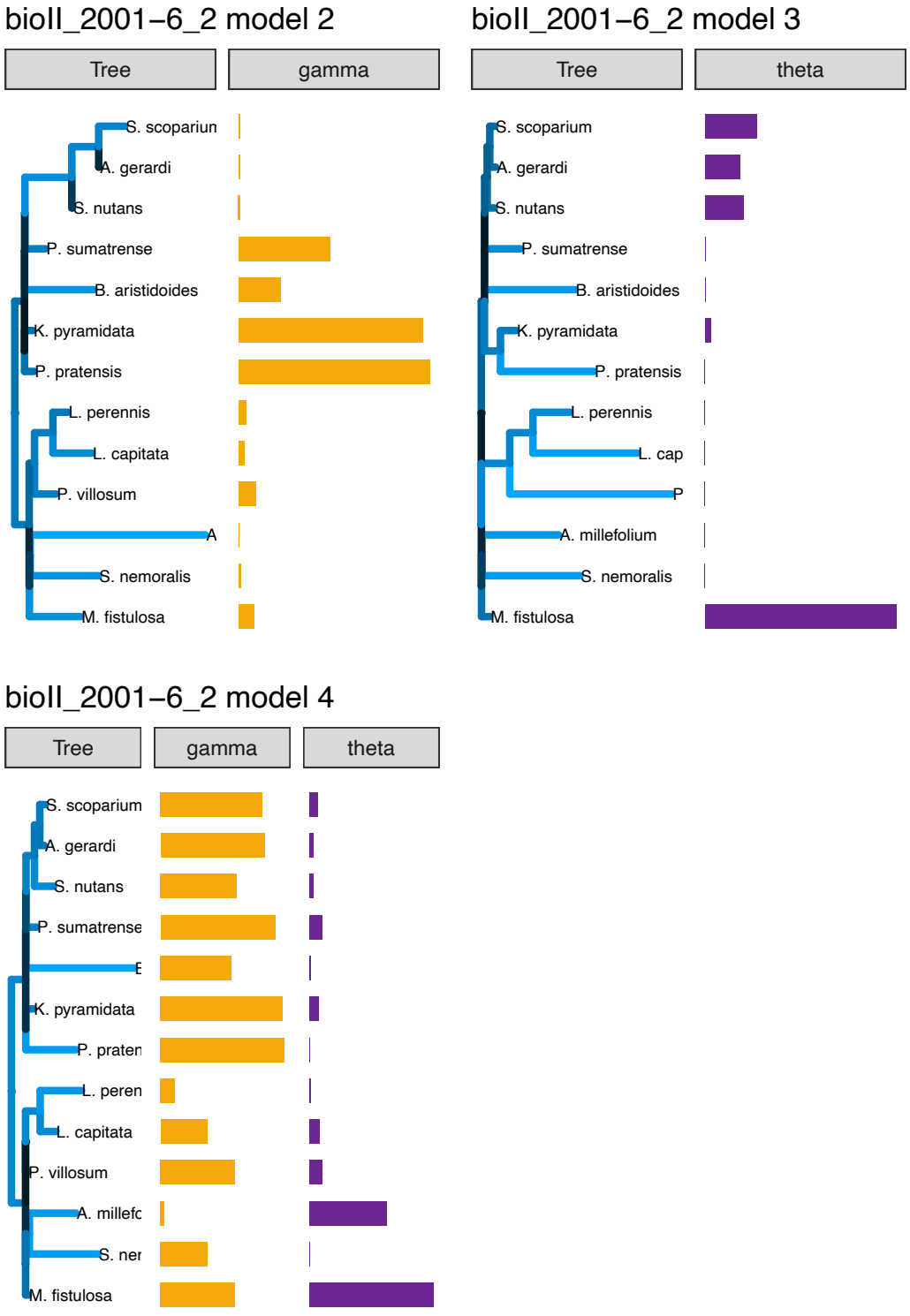

Figure 50: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

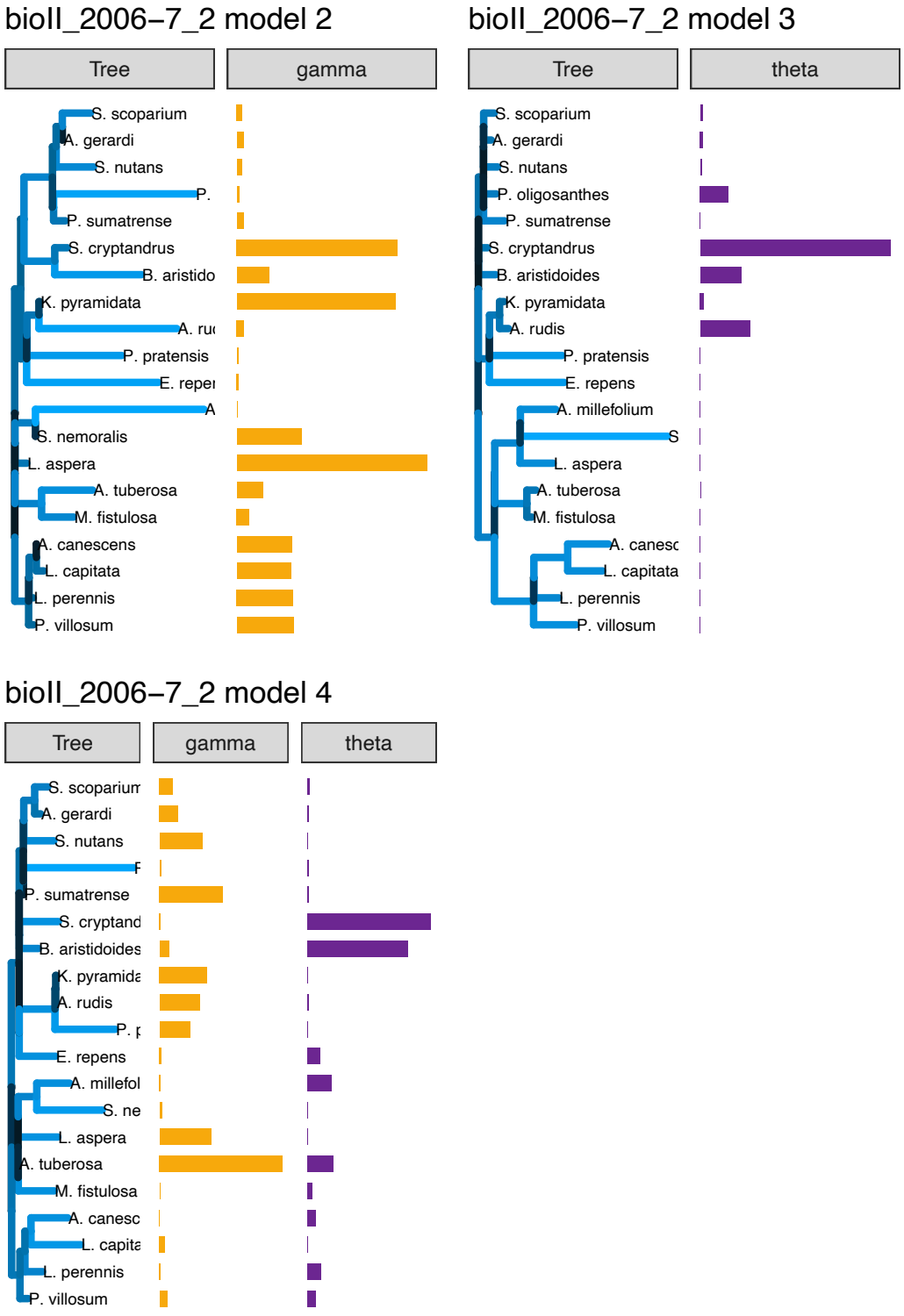

Figure 51: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

Figure 52: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

Figure 53: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

Figure 54: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

Figure 55: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

Figure 56: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

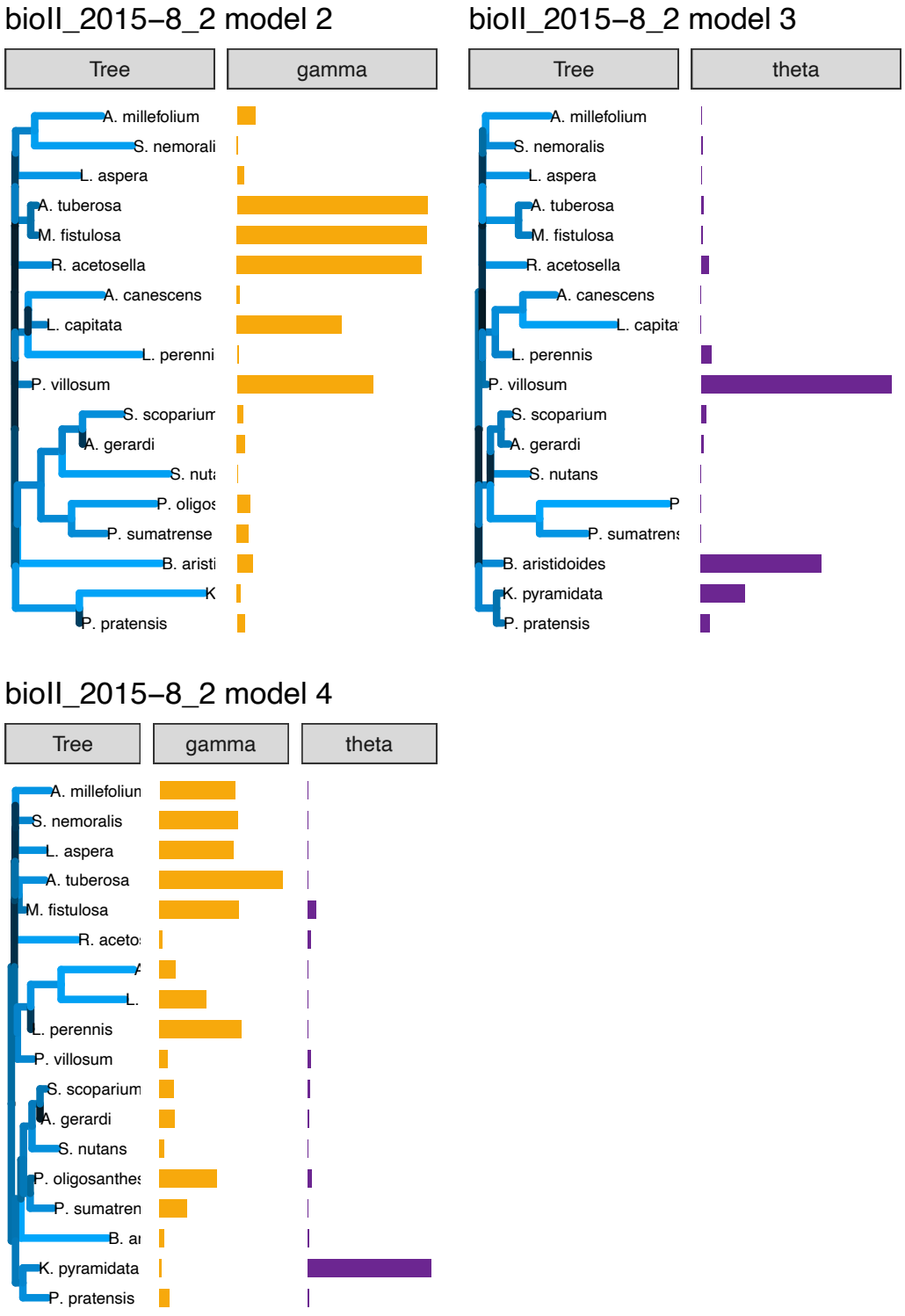

Figure 57: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

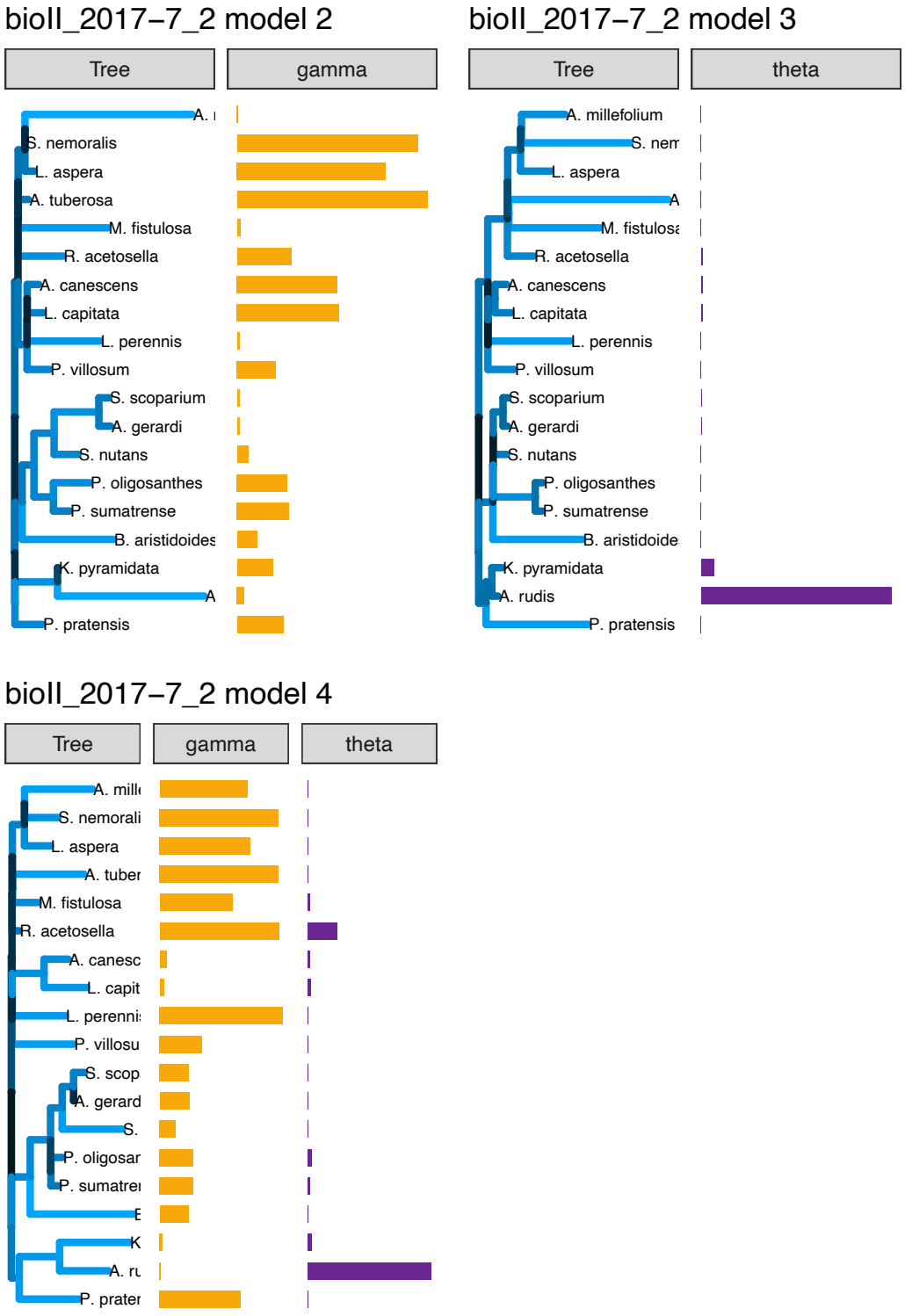

Figure 58: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

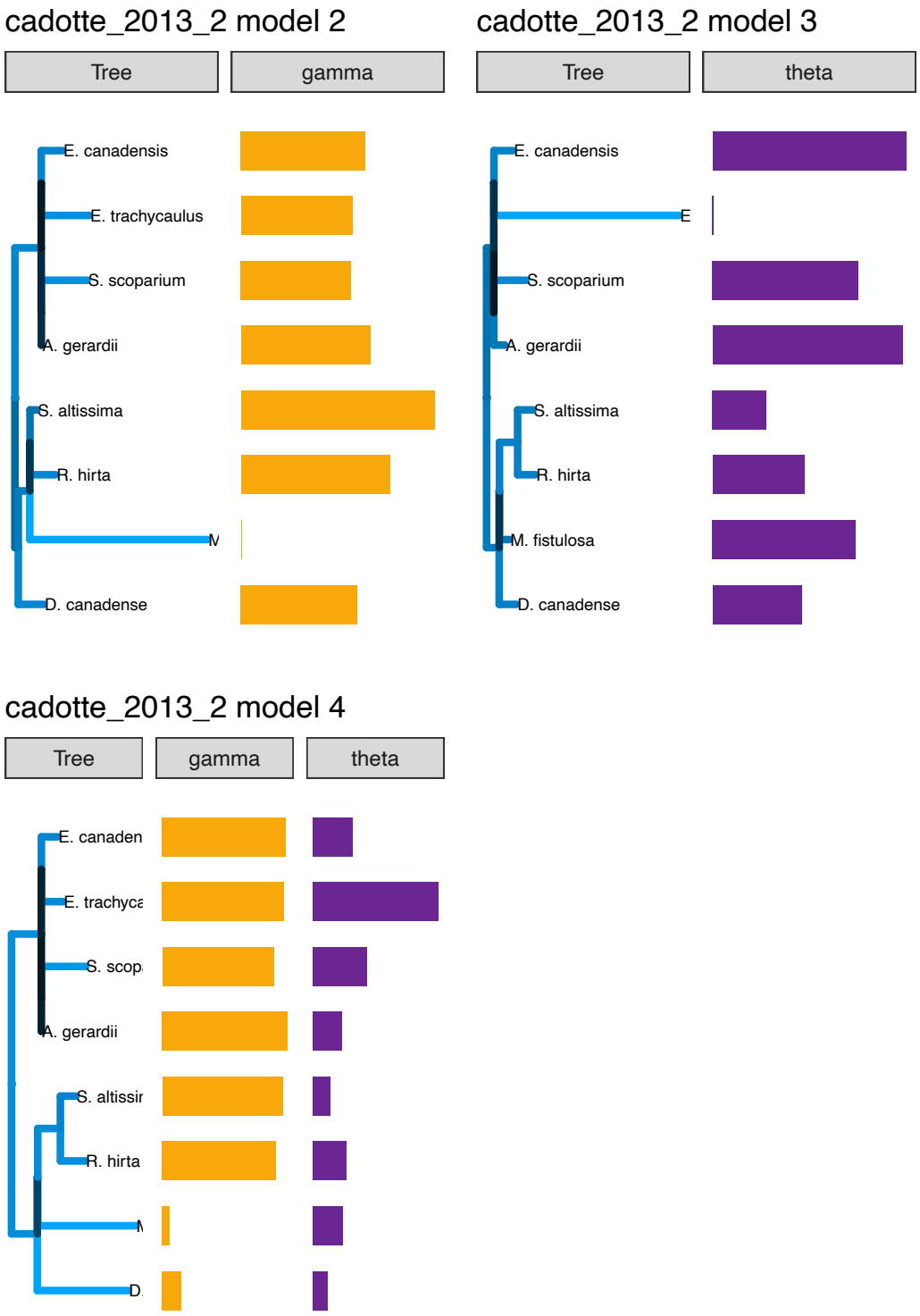

Figure 59: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

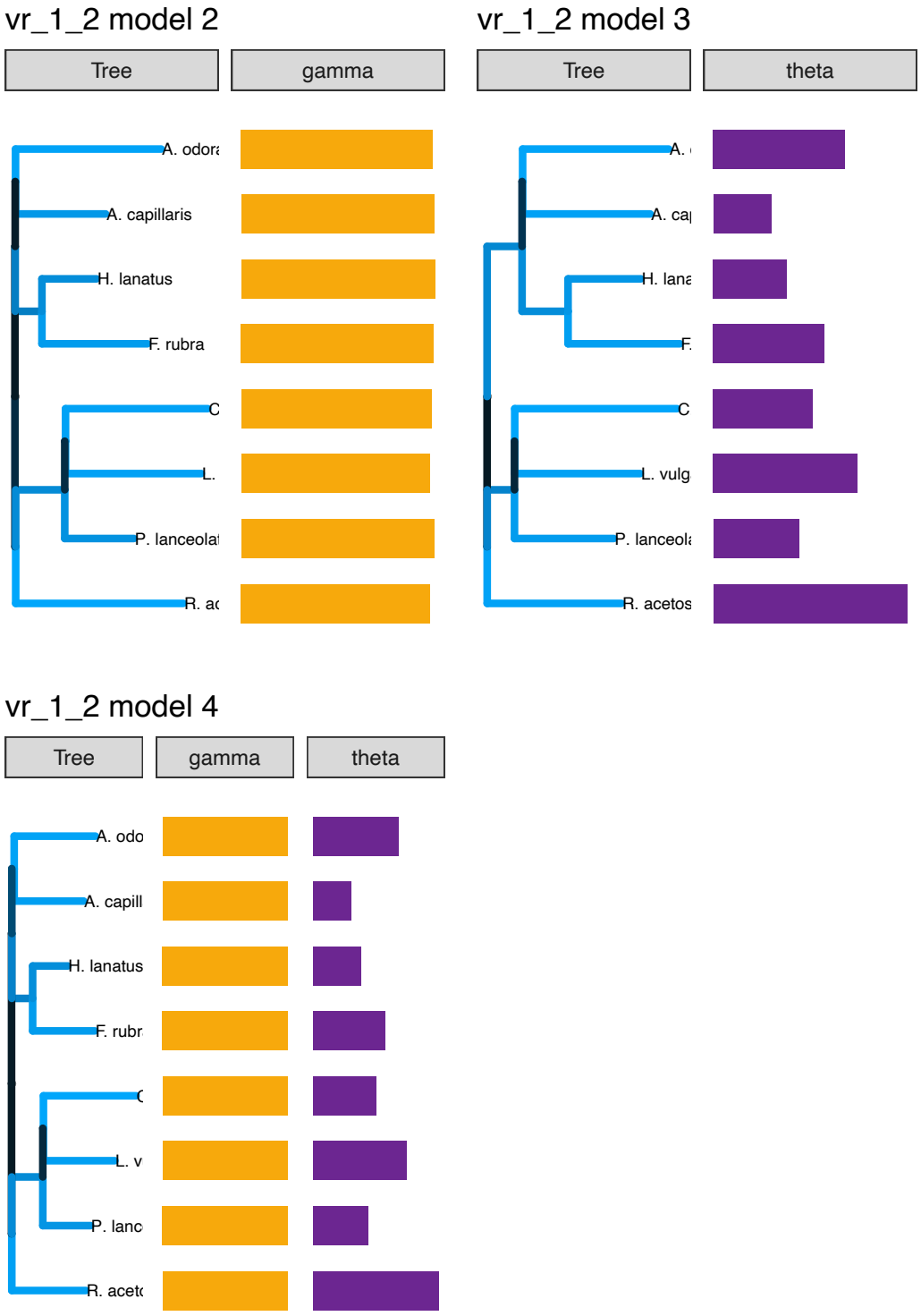

Figure 60: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

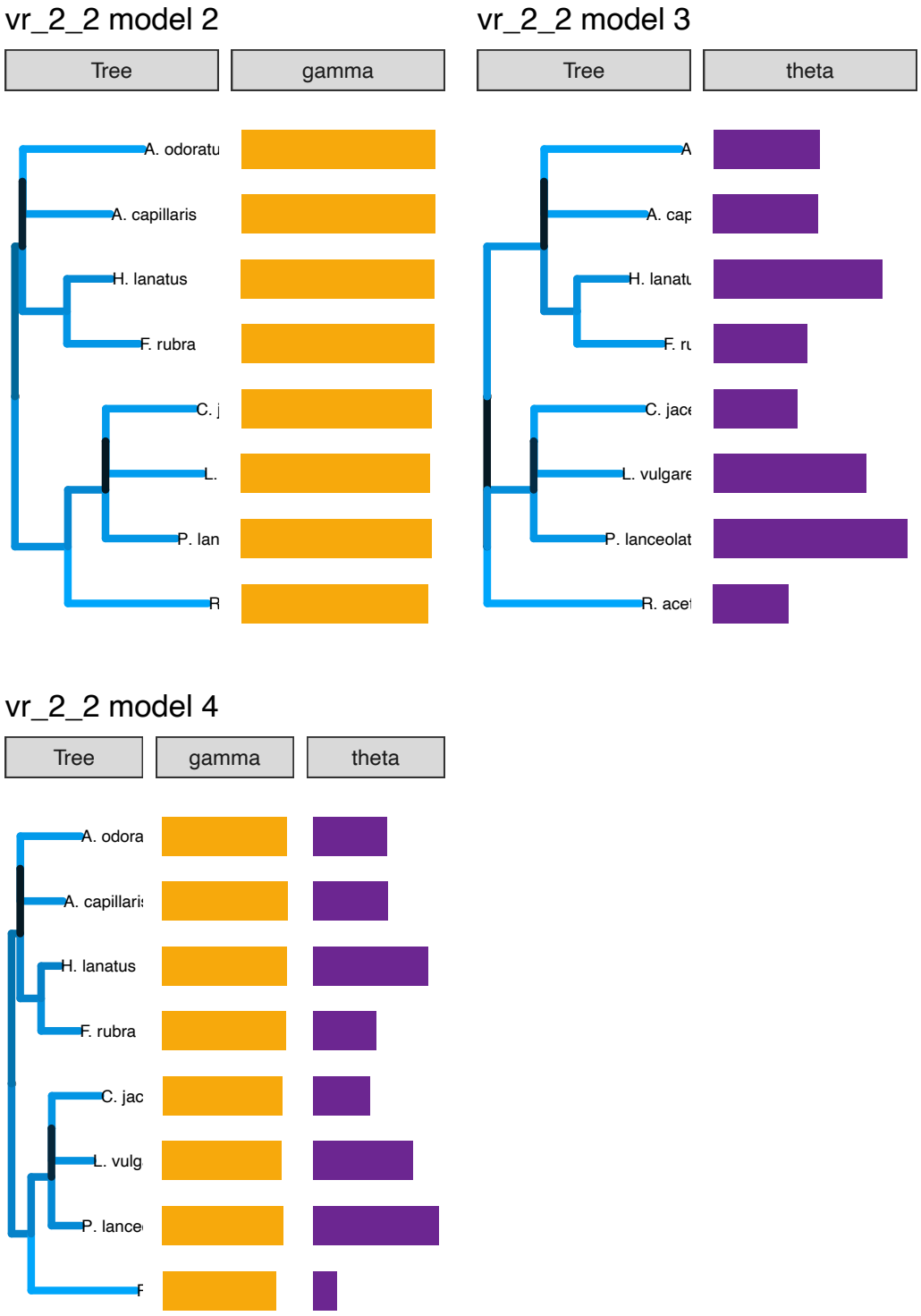

Figure 61: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

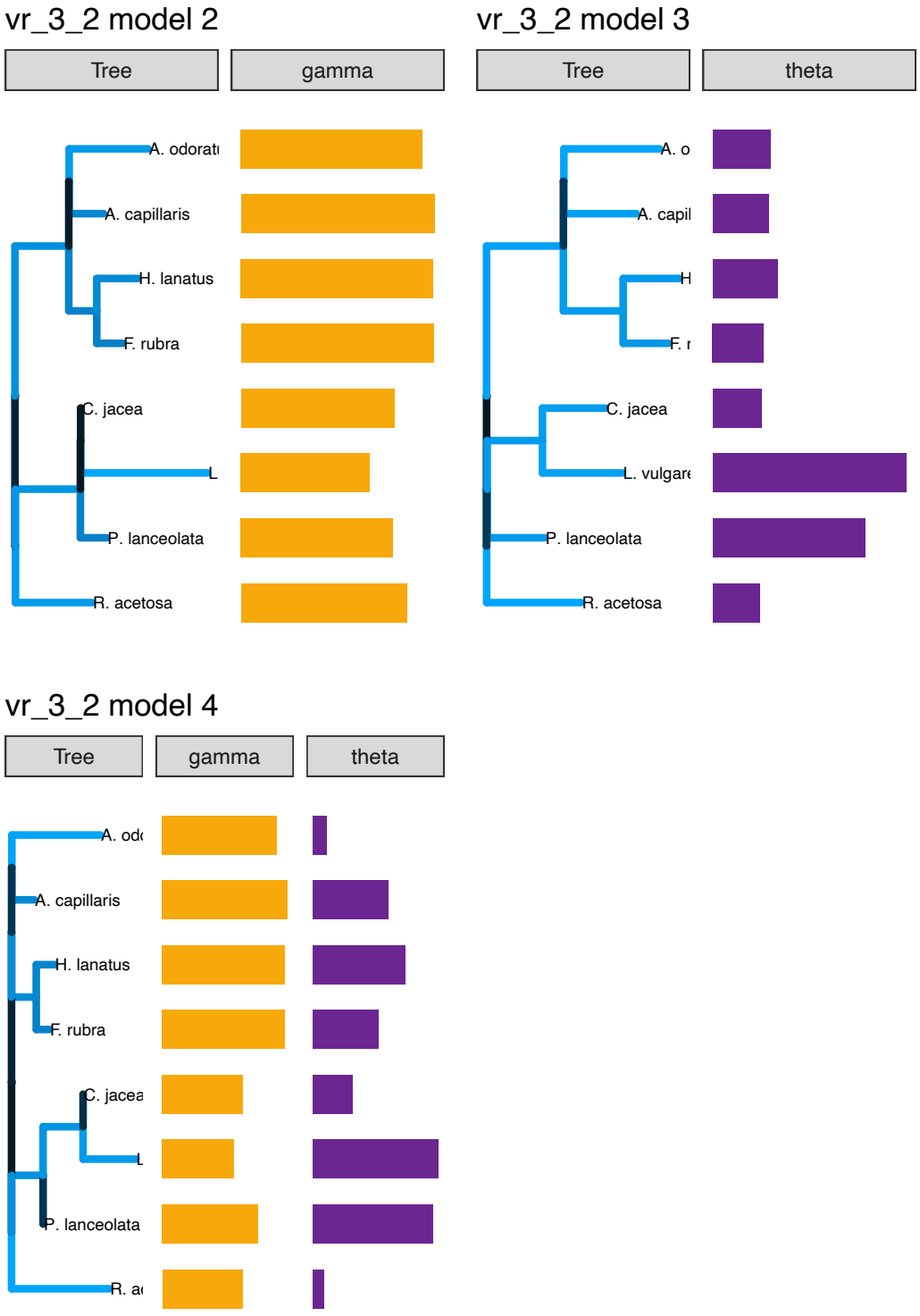

Figure 62: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

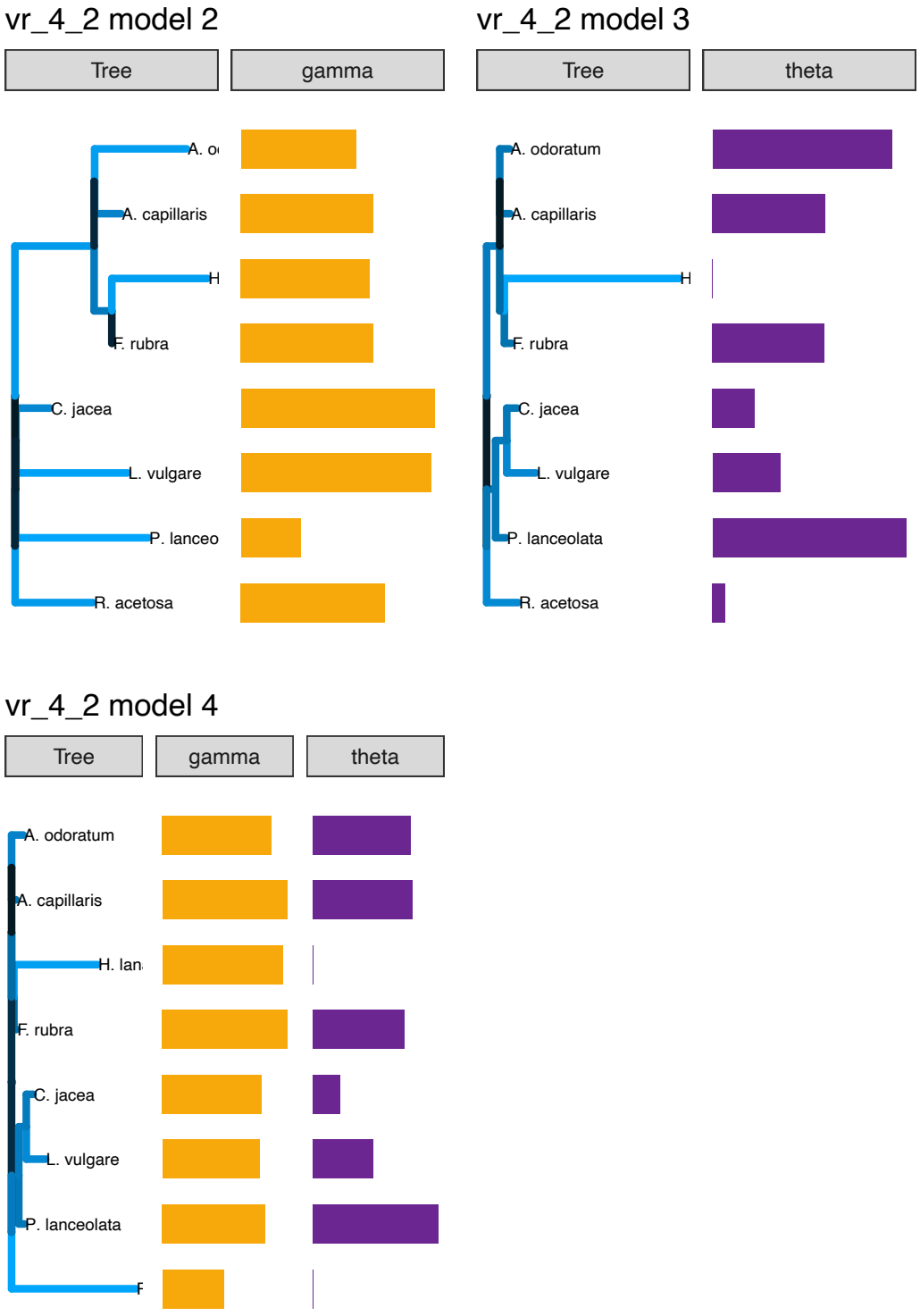

Figure 63: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

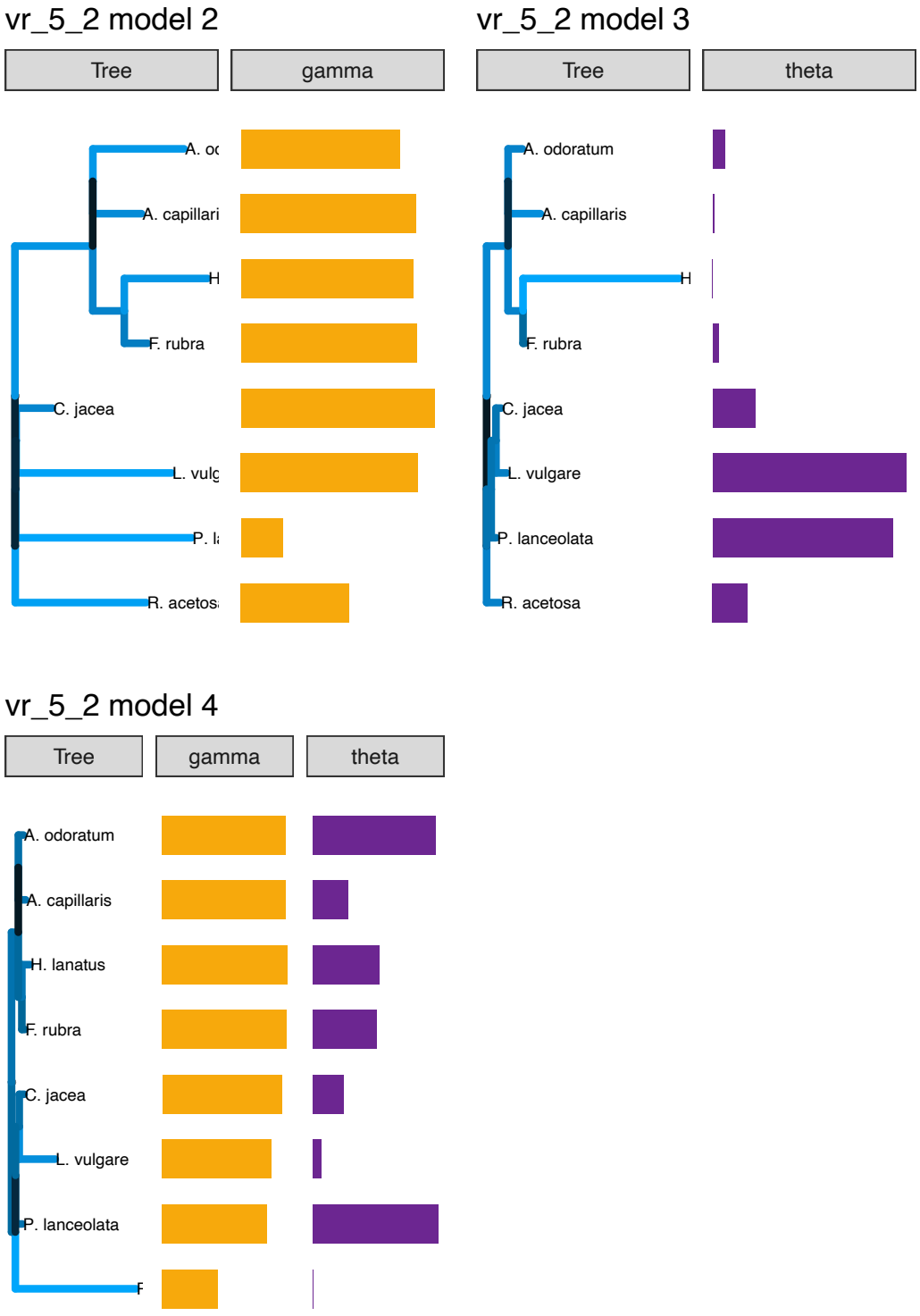

Figure 64: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

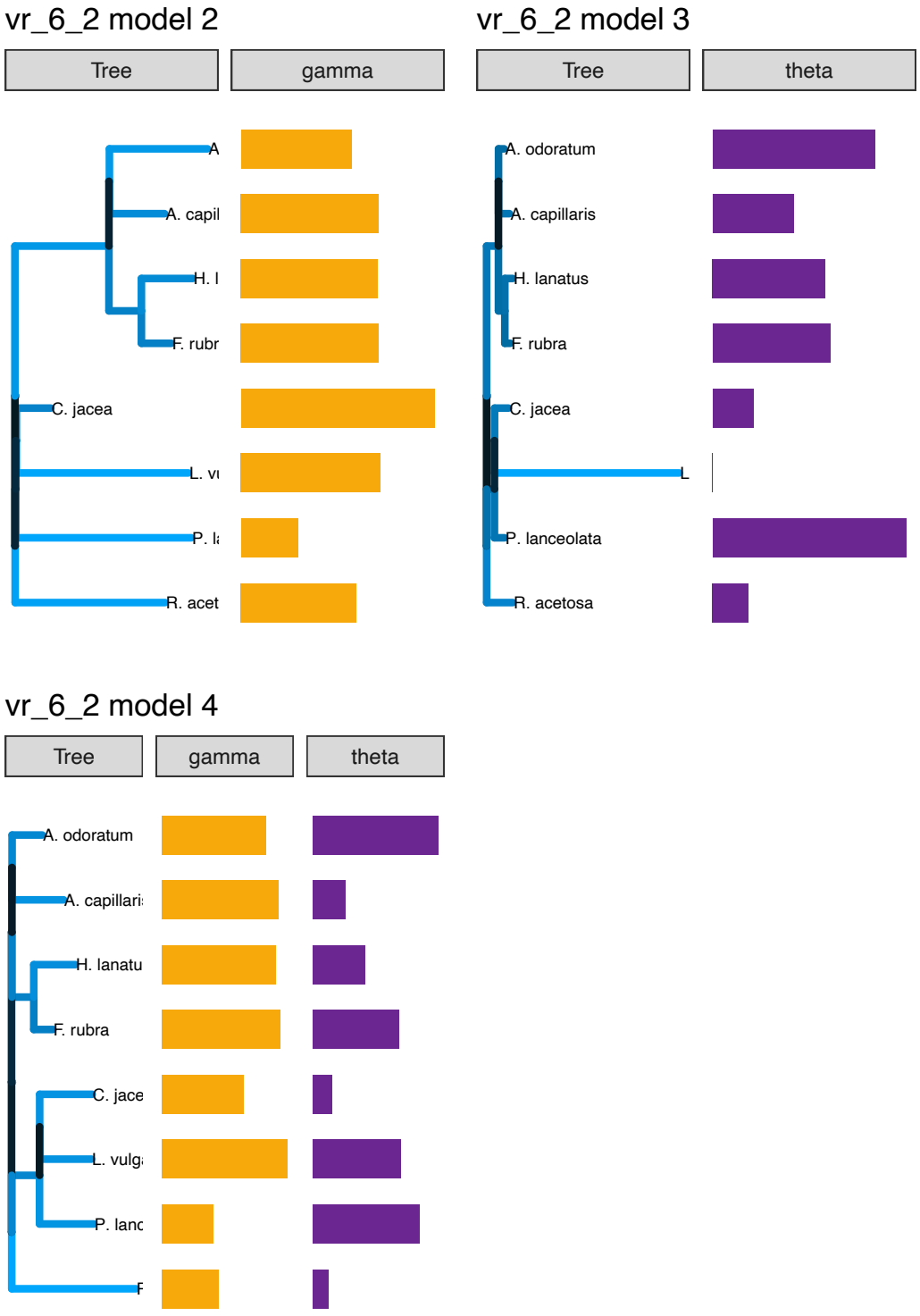

Figure 65: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

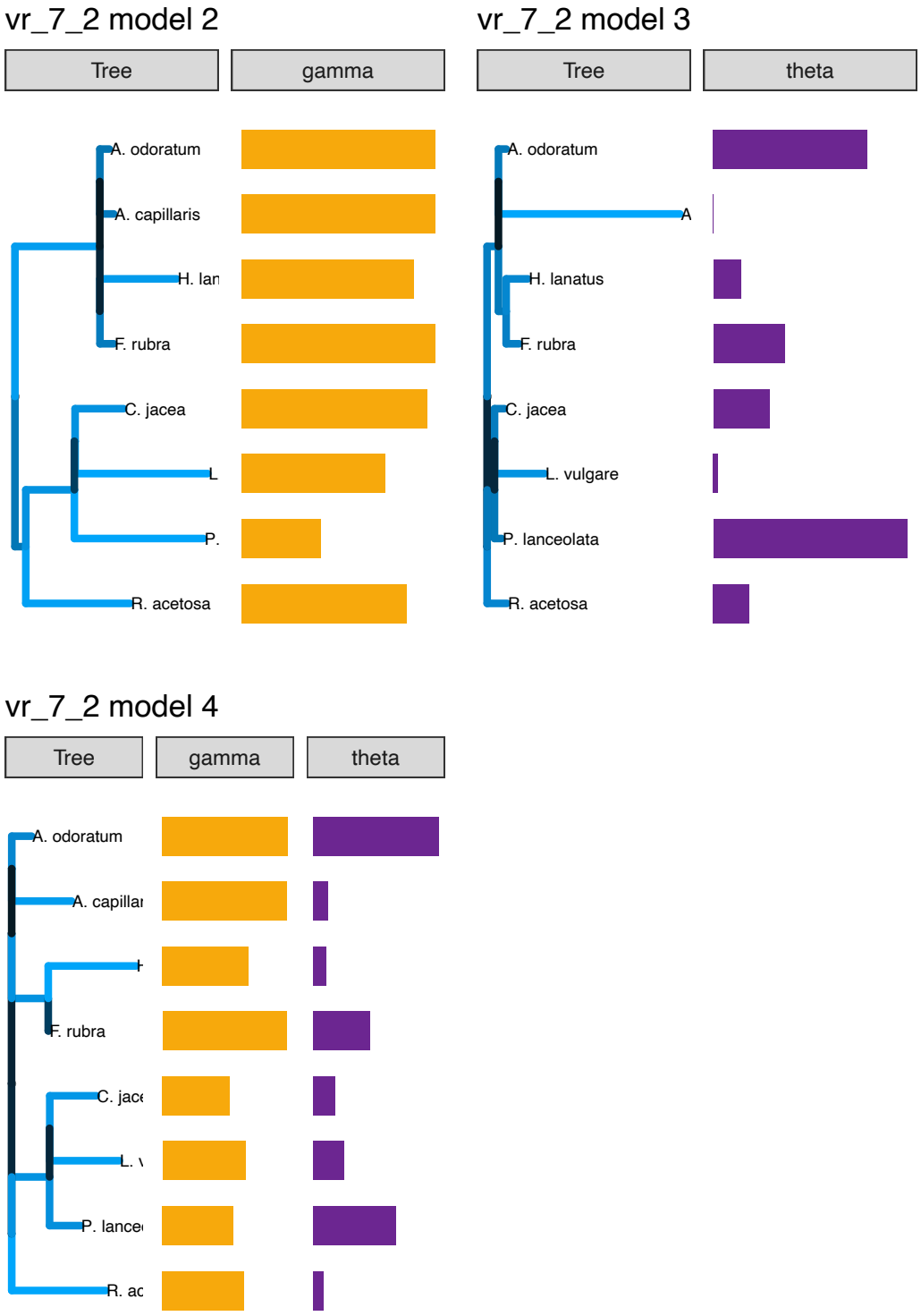

Figure 66: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

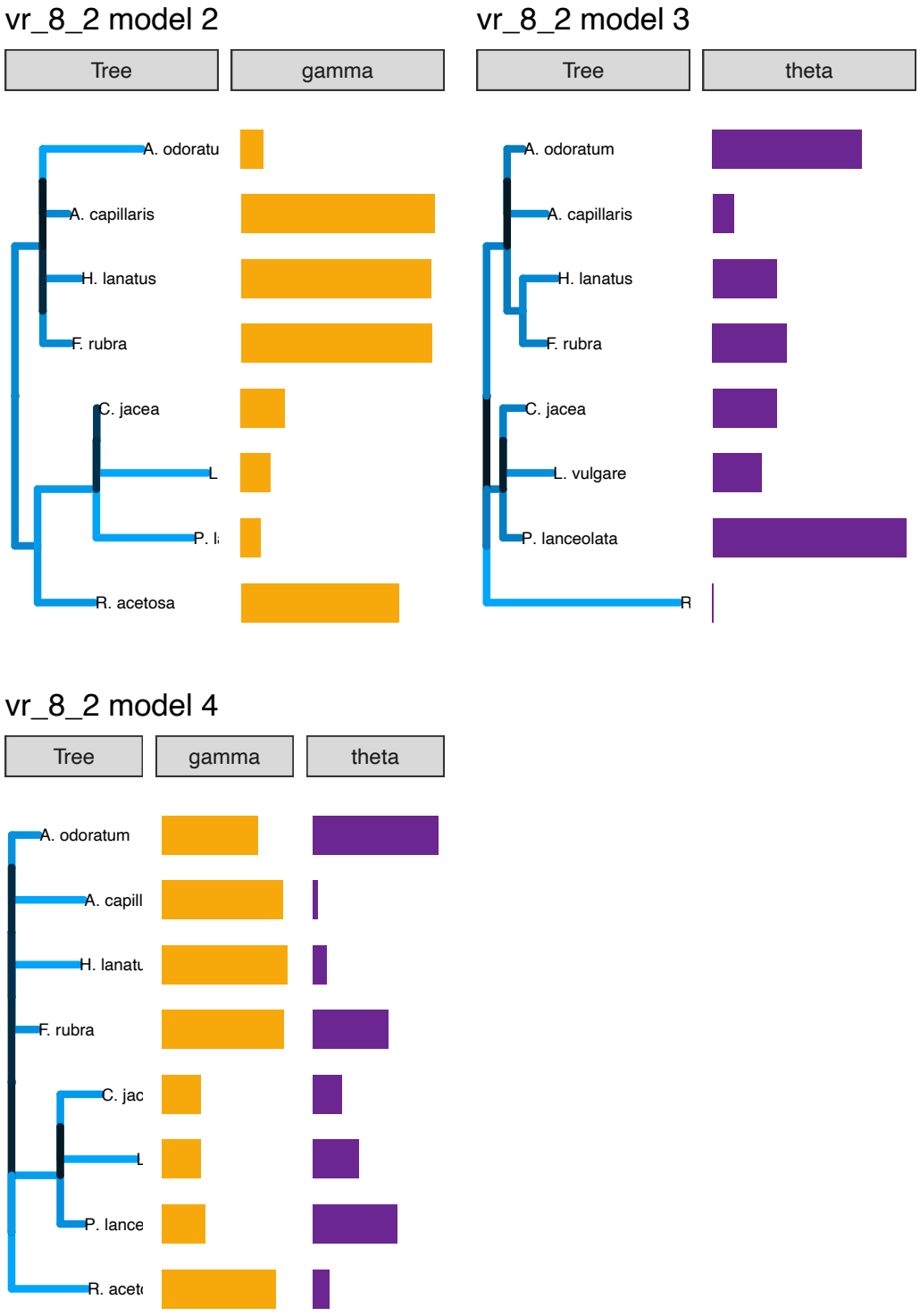

Figure 67: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

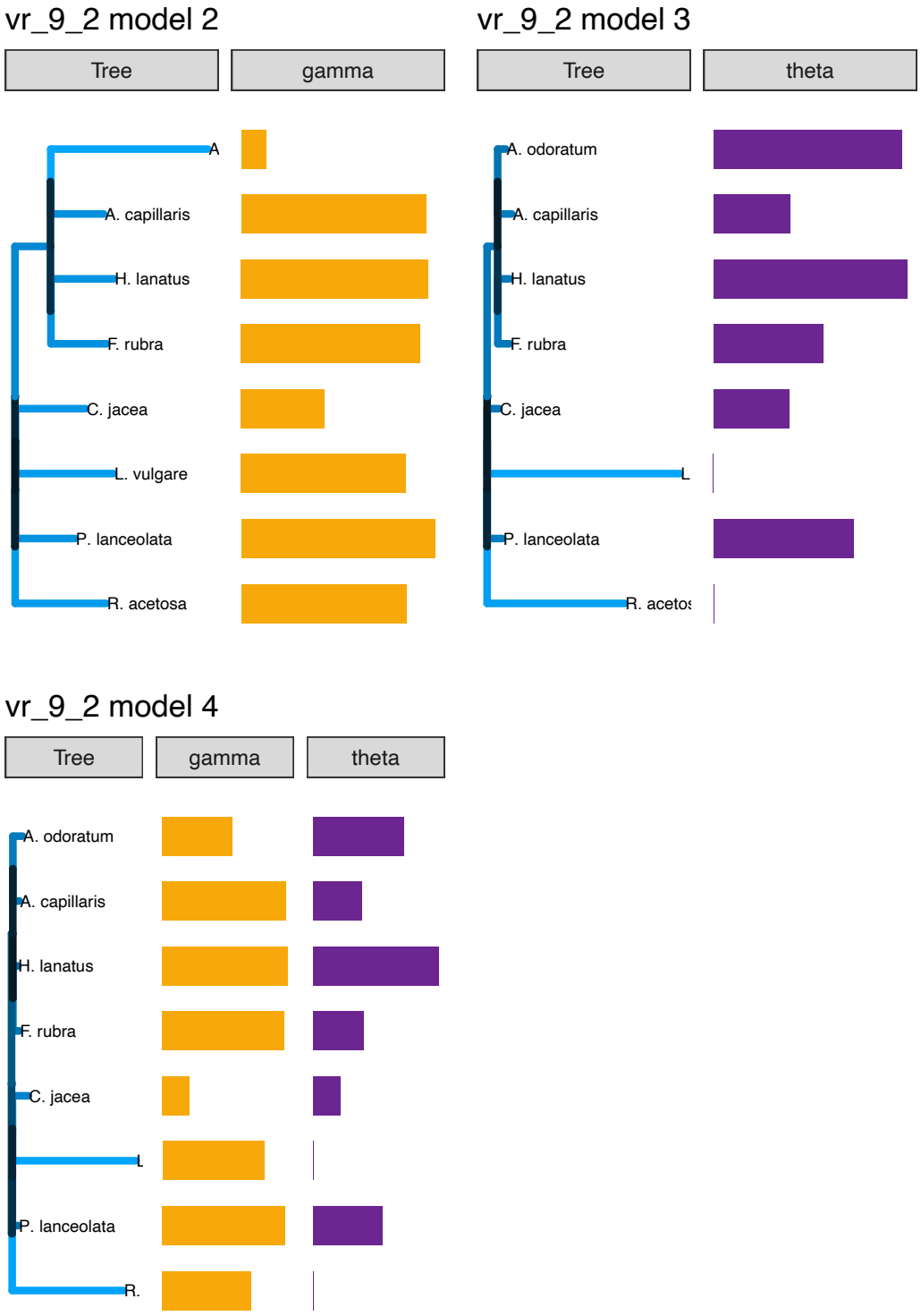

Figure 68: Fit for the phylogenetic tree and associated parameter vectors (Gamma and Theta) for models 2-4. The length and color of the branches correspond to increased strength in competition between the species in the clade and the size of the vector corresponds to the fitted value for the corresponding species in the tree. bioII refers to the Biodiversity II experiment (Tilman et al. (2001)); cadotte is from Cadotte (2013); and vr refers to the Wageningen experiment from Van Ruijven and Berendse (2010).

Complete References

Ackerly, D. (2009). Conservatism and diversification of plant functional traits: Evolutionary rates versus phylogenetic signal. *Proceedings of the National Academy of Sciences*, 106, 19699–19706.

Alexandrou, M.A., Cardinale, B.J., Hall, J.D., Delwiche, C.F., Fritschie, K., Narwani, A., *et al.* (2015). Evolutionary relatedness does not predict competition and co-occurrence in natural or experimental communities of green algae. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20141745.

Ansari, A.F., Reddy, Y.B., Raut, J. & Dixit, N.M. (2021). An efficient and scalable top-down method for predicting structures of microbial communities. *Nature Computational Science*, 1, 619–628.

Baeten, L., Davies, T.J., Verheyen, K., Van Calster, H. & Vellend, M. (2015). Disentangling dispersal from phylogeny in the colonization capacity of forest understorey plants. *Journal of Ecology*, 103, 175–183.

Bennett, J.A., Lamb, E.G., Hall, J.C., Cardinal-McTeague, W.M. & Cahill Jr, J.F. (2013). Increased competition does not lead to increased phylogenetic overdispersion in a native grassland. *Ecology letters*, 16, 1168–1176.

Bergamo, P.J., Wolowski, M., Maruyama, P.K., Vizentin-Bugoni, J. & Sazima, M. (2018). Trait patterns across space and time suggest an interplay of facilitation and competition acting on neotropical hummingbird-pollinated plant communities. *Oikos*, 127, 1690–1700.

Bravo, H.C., Wright, S., Eng, K., Keles, S. & Wahba, G. (2009). Estimating tree-structured covariance matrices via mixed-integer programming. In: *Artificial intelligence and statistics*. PMLR, pp. 41–48.

Burin, G., Campbell, L.C., Renner, S.S., Kiers, E.T. & Chomicki, G. (2024). Mutualisms drive plant trait evolution beyond interaction-related traits. *Ecology Letters*, 27, e14379.

Burns, J.H. & Strauss, S.Y. (2011). More closely related species are more ecologically similar in an experimental test. *Proceedings of the National Academy of Sciences*, 108, 5302–5307.

Cadotte, M.W. (2013). Experimental evidence that evolutionarily diverse assemblages result in higher productivity. *Proceedings of the National Academy of Sciences*, 110, 8996–9000.

Cadotte, M.W. (2015). Phylogenetic diversity and productivity. *Functional Ecology*, 29, 1603–1606.

Cadotte, M.W., Cardinale, B.J. & Oakley, T.H. (2008). Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences*, 105, 17012–17017.

Cadotte, M.W., Davies, T.J. & Peres-Neto, P.R. (2017). Why phylogenies do not always predict ecological differences. *Ecological Monographs*, 87, 535–551.

Cahill Jr, J.F., Kembel, S.W., Lamb, E.G. & Keddy, P.A. (2008). Does phylogenetic relatedness influence the strength of competition among vascular plants? *Perspectives in Plant Ecology, Evolution and Systematics*, 10, 41–50.

Campbell, A.J., Gigante Carvalheiro, L., Gastauer, M., Almeida-Neto, M. & Giannini, T.C. (2019). Pollinator restoration in brazilian ecosystems relies on a small but phylogenetically-diverse set of plant families. *Scientific reports*, 9, 17383.

Cardinale, B.J., Venail, P., Gross, K., Oakley, T.H., Narwani, A., Allan, E., *et al.* (2015). Further re-analyses looking for effects of phylogenetic diversity on community biomass and stability. *Functional Ecology*, 29, 1607–1610.

Cavender-Bares, J., Ackerly, D.D., Baum, D. & Bazzaz, F. (2004). Phylogenetic overdispersion in floridian oak communities. *The American Naturalist*, 163, 823–843.

Cavender-Bares, J., Keen, A. & Miles, B. (2006). Phylogenetic structure of floridian plant communities depends on taxonomic and spatial scale. *Ecology*, 87, S109–S122.

Cavender-Bares, J., Kozak, K.H., Fine, P.V. & Kembel, S.W. (2009). The merging of community ecology and phylogenetic biology. *Ecology letters*, 12, 693–715.

Connolly, J., Bell, T., Bolger, T., Brophy, C., Carnus, T., Finn, J.A., *et al.* (2013). An improved model to predict the effects of changing biodiversity levels on ecosystem function. *Journal of Ecology*, 101, 344–355.

Connolly, J., Cadotte, M.W., Brophy, C., Dooley, Á., Finn, J., Kirwan, L., *et al.* (2011). Phylogenetically diverse grasslands are associated with pairwise interspecific processes that increase biomass. *Ecology*, 92, 1385–1392.

Darwin, C. (1859). *On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life*. John Murray.

Davies, T.J. (2021). Ecophylogenetics redux. *Ecology letters*, 24, 1073–1088.

Davies, T.J., Urban, M.C., Rayfield, B., Cadotte, M.W. & Peres-Neto, P.R. (2016). Deconstructing the relationships between phylogenetic diversity and ecology: A case study on ecosystem functioning. *Ecology*, 97, 2212–2222.

Elton, C. (1946). Competition and the structure of ecological communities. *The Journal of Animal Ecology*, 54–68.

Flynn, D.F., Mirotchnick, N., Jain, M., Palmer, M.I. & Naeem, S. (2011). Functional and phylogenetic diversity as predictors of biodiversity–ecosystem-function relationships. *Ecology*, 92, 1573–1581.

Fort, H. (2018). On predicting species yields in multispecies communities: Quantifying the accuracy of the linear lotka-volterra generalized model. *Ecological modelling*, 387, 154–162.

Frainer, A., McKie, B.G., Amundsen, P.-A., Knudsen, R. & Lafferty, K.D. (2018). Parasitism and the biodiversity-functioning relationship. *Trends in Ecology & Evolution*, 33, 260–268.

Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002). Phylogenetic analysis and comparative data: A test and review of evidence. *The American Naturalist*, 160, 712–726.

Fritschie, K.J., Cardinale, B.J., Alexandrou, M.A. & Oakley, T.H. (2014). Evolutionary history and the strength of species interactions: Testing the phylogenetic limiting similarity hypothesis. *Ecology*, 95, 1407–1417.

Gerhold, P., Cahill Jr, J.F., Winter, M., Bartish, I.V. & Prinzing, A. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, 29, 600–614.

Godoy, O., Kraft, N.J. & Levine, J.M. (2014). Phylogenetic relatedness and the determinants of competitive outcomes. *Ecology letters*, 17, 836–844.

Gómez, J.M., Verdú, M. & Perfectti, F. (2010). Ecological interactions are evolutionarily conserved across the entire tree of life. *Nature*, 465, 918–921.

Graham, C.H., Parra, J.L., Rahbek, C. & McGuire, J.A. (2009). Phylogenetic structure in tropical hummingbird communities. *Proceedings of the National Academy of Sciences*, 106, 19673–19678.

Halty, V., Valdés, M., Tejera, M., Picasso, V. & Fort, H. (2017). Modeling plant interspecific interactions from experiments with perennial crop mixtures to predict optimal combinations. *Ecological Applications*, 27, 2277–2289.

Hansen, T.F. & Martins, E.P. (1996). Translating between microevolutionary process and macroevolutionary patterns: The correlation structure of interspecific data. *Evolution*, 50, 1404–1417.

Huang, M., Liu, X., Cadotte, M.W. & Zhou, S. (2020). Functional and phylogenetic diversity explain different components of diversity effects on biomass production. *Oikos*, 129, 1185–1195.

Ives, A.R. & Helmus, M.R. (2011). Generalized linear mixed models for phylogenetic analyses of community structure. *Ecological Monographs*, 81, 511–525.

Jiang, L., Tan, J. & Pu, Z. (2010). An experimental test of darwin's naturalization hypothesis. *The American Naturalist*, 175, 415–423.

Jin, Y. & Qian, H. (2019). V. PhyloMaker: An r package that can generate very large phylogenies for vascular plants. *Ecography*, 42, 1353–1359.

Jochum, M., Fischer, M., Isbell, F., Roscher, C., Plas, F. van der, Boch, S., *et al.* (2020). The results of biodiversity–ecosystem functioning experiments are realistic. *Nature ecology & evolution*, 4, 1485–1494.

Keck, F. & Kahlert, M. (2019). Community phylogenetic structure reveals the imprint of dispersal-related dynamics and environmental filtering by nutrient availability in freshwater diatoms. *Scientific Reports*, 9, 11590.

Kettenring, K.M., Mercer, K.L., Reinhardt Adams, C. & Hines, J. (2014). Application of genetic diversity–ecosystem function research to ecological restoration. *Journal of applied ecology*, 51, 339–348.

Letten, A.D. & Cornwell, W.K. (2015). Trees, branches and (square) roots: Why evolutionary relatedness is not linearly related to functional distance. *Methods in Ecology & Evolution*, 6.

Liu, X., Swenson, N.G., Zhang, J. & Ma, K. (2013). The environment and space, not phylogeny, determine trait dispersion in a subtropical forest. *Functional Ecology*, 27, 264–272.

Losos, J.B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology letters*, 11, 995–1003.

Lososová, Z., Čeplová, N., Chytrỳ, M., Tichỳ, L., Danihelka, J., Fajmon, K., *et al.* (2016). Is phylogenetic diversity a good proxy for functional diversity of plant communities? A case study from urban habitats. *Journal of Vegetation Science*, 27, 1036–1046.

Louw, N.L., Wolfe, B.E. & Uricchio, L.H. (2024). A phylogenomic perspective on interspecific competition.

Ecology Letters, 27, e14359.
MacArthur, R. (1970). Species packing and competitive equilibrium for many species. *Theoretical population biology*, 1, 1–11.
Maherali, H. & Klironomos, J.N. (2007). Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science*, 316, 1746–1748.
Mayfield, M.M. & Levine, J.M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology letters*, 13, 1085–1093.
Maynard, D.S., Miller, Z.R. & Allesina, S. (2020). Predicting coexistence in experimental ecological communities. *Nature ecology & evolution*, 4, 91–100.
Mazel, F., Mooers, A.O., Riva, G.V.D. & Pennell, M.W. (2017). Conserving phylogenetic diversity can be a poor strategy for conserving functional diversity. *Systematic Biology*, 66, 1019–1027.
McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in ecology & evolution*, 21, 178–185.
Molina-Venegas, R., Fischer, M. & Hemp, A. (2019). Disentangling the fundamental branching patterns of phylogenetic divergence to refine eco-phylogenetic analyses. *Journal of Biogeography*, 46, 2722–2734.
Mouquet, N., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.-F., Chave, J., *et al.* (2012). Ecophylogenetics: Advances and perspectives. *Biological reviews*, 87, 769–785.
Navas, M.-L. & Violle, C. (2009). Plant traits related to competition: How do they shape the functional diversity of communities? *Community Ecology*, 10, 131–137.
Nuismer, S.L. & Harmon, L.J. (2015). Predicting rates of interspecific interaction from phylogenetic trees. *Ecology letters*, 18, 17–27.
Parain, E.C., Rohr, R.P., Gray, S.M. & Bersier, L.-F. (2019). Increased temperature disrupts the biodiversity–ecosystem functioning relationship. *The American Naturalist*, 193, 227–239.
Peay, K.G., Belisle, M. & Fukami, T. (2012). Phylogenetic relatedness predicts priority effects in nectar yeast communities. *Proceedings of the Royal Society B: Biological Sciences*, 279, 749–758.
Petchey, O.L., Beckerman, A.P., Riede, J.O. & Warren, P.H. (2008). Size, foraging, and food web structure. *Proceedings of the National Academy of Sciences*, 105, 4191–4196.
R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
Revell, L.J. (2010). Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution*, 1, 319–329.
Rosas-Guerrero, V., Aguilar, R., Martén-Rodríguez, S., Ashworth, L., Lopezaraiza-Mikel, M., Bastida, J.M., *et al.* (2014). A quantitative review of pollination syndromes: Do floral traits predict effective pollinators? *Ecology letters*, 17, 388–400.
Serván, C.A., Capitán, J.A., Miller, Z.R. & Allesina, S. (2023). Effects of phylogeny on coexistence in model communities. *arXiv preprint arXiv:2310.14392*.
Simberloff, D.S. (1970). Taxonomic diversity of island biotas. *Evolution*, 23–47.
Skwara, A., Lemos-Costa, P., Miller, Z.R. & Allesina, S. (2023). Modelling ecological communities when composition is manipulated experimentally. *Methods in Ecology and Evolution*, 14, 696–707.
Smith, M.R. (2020). Information theoretic generalized robinson–foulds metrics for comparing phylogenetic trees. *Bioinformatics*, 36, 5007–5013.
Srivastava, D.S., Cadotte, M.W., MacDonald, A.A.M., Marushia, R.G. & Mirotchnick, N. (2012). Phylogenetic diversity and the functioning of ecosystems. *Ecology letters*, 15, 637–648.
Stan Development Team. (2023). *Stan modeling language users guide and reference manual, 2.32*.
Steudel, B., Hallmann, C., Lorenz, M., Abrahamczyk, S., Prinz, K., Herrfurth, C., *et al.* (2016). Contrasting biodiversity–ecosystem functioning relationships in phylogenetic and functional diversity. *New Phytologist*, 212, 409–420.
Stouffer, D.B., Rezende, E.L. & Amaral, L.A.N. (2011). The role of body mass in diet contiguity and food-web structure. *Journal of Animal Ecology*, 80, 632–639.
Tilman, D., Reich, P.B., Knops, J., Wedin, D., Mielke, T. & Lehman, C. (2001). Diversity and productivity in a long-term grassland experiment. *Science*, 294, 843–845.
Tucker, C.M., Davies, T.J., Cadotte, M.W. & Pearse, W.D. (2018). On the relationship between phylogenetic diversity and trait diversity. *Ecology*, 99, 1473–1479.
Van Ruijven, J. & Berendse, F. (2010). Diversity enhances community recovery, but not resistance, after

drought. *Journal of Ecology*, 98, 81–86.
Venail, P.A., Narwani, A., Fritschie, K., Alexandrou, M.A., Oakley, T.H. & Cardinale, B.J. (2014). The influence of phylogenetic relatedness on species interactions among freshwater green algae in a mesocosm experiment. *Journal of Ecology*, 102, 1288–1299.
Venail, P., Gross, K., Oakley, T.H., Narwani, A., Allan, E., Flombaum, P., *et al.* (2015). Species richness, but not phylogenetic diversity, influences community biomass production and temporal stability in a re-examination of 16 grassland biodiversity studies. *Functional Ecology*, 29, 615–626.
Violle, C., Nemergut, D.R., Pu, Z. & Jiang, L. (2011). Phylogenetic limiting similarity and competitive exclusion. *Ecology letters*, 14, 782–787.
Webb, C.O. (2000). Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. *The American Naturalist*, 156, 145–155.
Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002). Phylogenies and community ecology. *Annual review of ecology and systematics*, 33, 475–505.
Weber, M.G., Wagner, C.E., Best, R.J., Harmon, L.J. & Matthews, B. (2017). Evolution in a community context: On integrating ecological interactions and macroevolution. *Trends in ecology & evolution*, 32, 291–304.
Wright, A.J., Wardle, D.A., Callaway, R. & Gaxiola, A. (2017). The overlooked role of facilitation in biodiversity experiments. *Trends in Ecology & Evolution*, 32, 383–390.
Xiao, Y., Angulo, M.T., Friedman, J., Waldor, M.K., Weiss, S.T. & Liu, Y.-Y. (2017). Mapping the ecological networks of microbial communities. *Nature communications*, 8, 2042.
Yguel, B., Jactel, H., Pearse, I.S., Moen, D., Winter, M., Hortal, J., *et al.* (2016). The evolutionary legacy of diversification predicts ecosystem function. *The American Naturalist*, 188, 398–410.