

# Supplemental Analysis

ANONYMIZED

## Phylogenetic Specialism

Here we quantify the phylogenetic specialism of frugivores. To do this, we'll employ the `ses.mpd()` from the `picante` library. This function allows us to use a null model of our choice to find the standard effect size of pairwise distance in communities. In our case, we're interested in pairwise mean phylogenetic distance, and our "communities" are the total sets of plants that each frugivore interact with. We employ an independent swap null model, where the tip labels of the underlying phylogenetic tree are effectively swapped a given number of times (Gotelli, 2000).

```
set.seed(1)
load("../Data.nosync/DataSources/BIEN_subtree.Rda")
# Create cophenetic distance matrix
dmatrix <- stats::cophenetic(BIEN_subtree)
# Standardize by the maximum value
dmatrix <- dmatrix/max(dmatrix)

# String manipulation so tree tip names match
birds$plant_Tips <- gsub(pattern = " ", replace = "_", x = birds$Plant_Species)
# Only include plant species for which phylogenetic info is
# available
birds_phylo <- dplyr::filter(birds, plant_Tips %in% BIEN_subtree$tip.label)

comm <- as.data.frame.matrix(table(birds_phylo$Frugivore_Species,
    birds_phylo$plant_Tips)) #Binary, unweighted interaction matrix.
# The main function. Randomly swap tip labels 999 times and
# see whether the observed sets of plant interactors for
# each species is more phylogenetically clustered than this
# assumed distribution
test_mpd <- picante::ses.mpd(comm, dmatrix, null.model = "independentswap",
    abundance.weighted = FALSE, runs = 999, iterations = 1000)
# Remove singleton species (can't compute a phylogenetic
# distance)
phylospec <- dplyr::filter(test_mpd, ntaxa > 1)
# Remove singleton species (can't compute a phylogenetic
# distance)
sig <- nrow(dplyr::filter(phylospec, mpd.obs.p < 0.05))
print(sig)

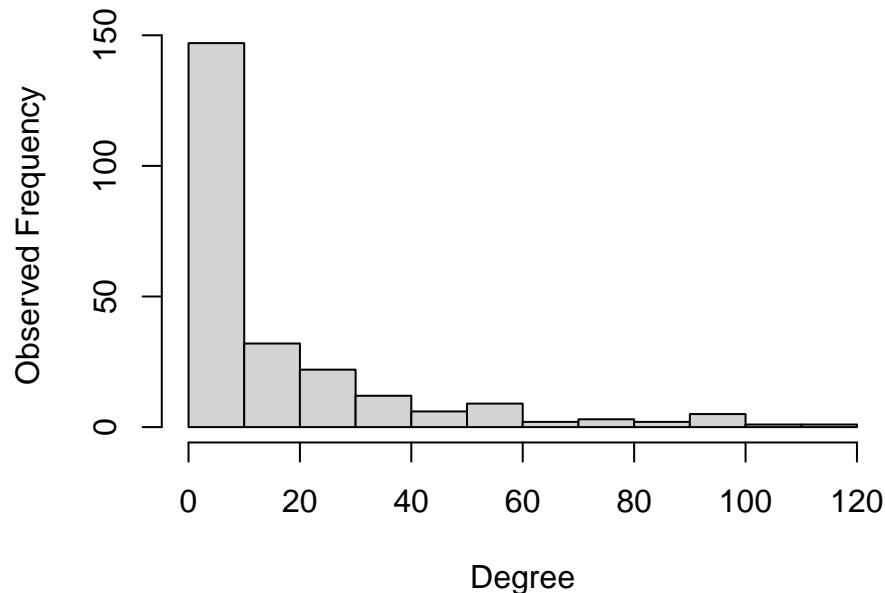
## [1] 9
```

We see that the number of frugivores whose plant partners are more significantly related than expected compared to a random draw is equal to 9 (The exact value may change slightly when the seed is changed due to the stochastic nature of the model, but the results should be qualitatively similar each time)

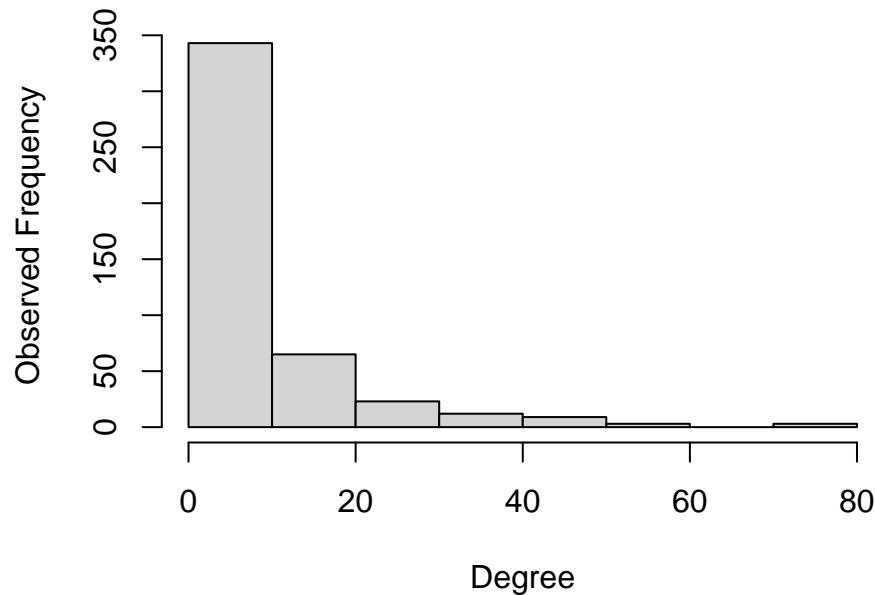
## Calculating the number of unique species in the dataset

In our dataset, we observe interactions between 242 unique frugivore species and 458 unique plant species.

## Degree Distributions of Mutualistic Partners



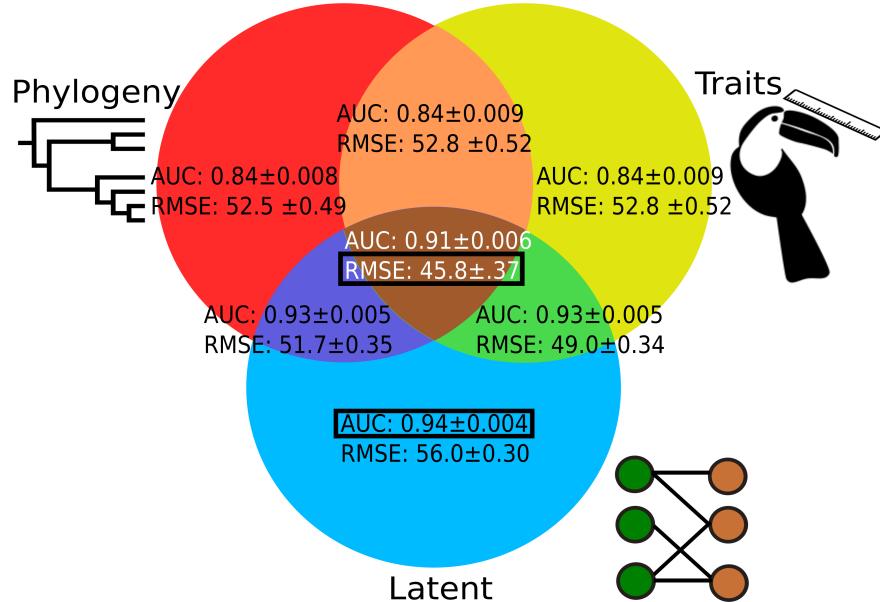
**Figure S1:** Observed degree distribution of avian frugivore species included in our analyses. X-axis represents node degree, or number of unique partners. Degree for frugivores ranged from 1 (53 species) to 120 unique plant interactions recorded for *Turdus rufiventris*; median frugivore degree was 5.



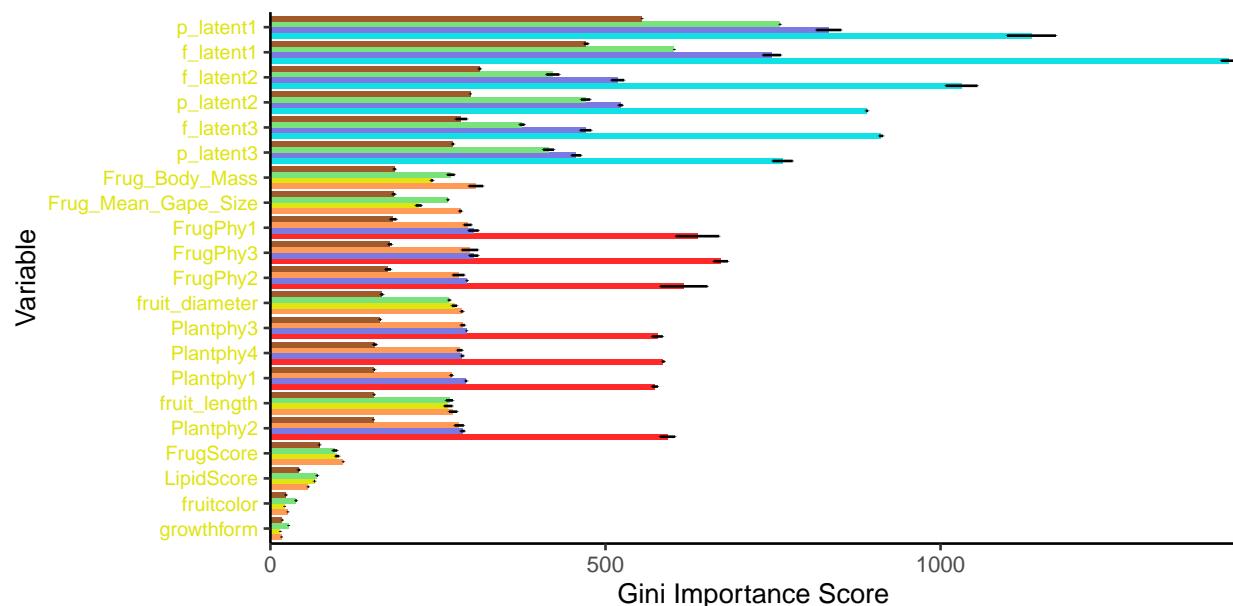
**Figure S2:** Observed degree distribution of avian frugivore species included in our analyses. X-axis represents node degree, or number of unique partners. Degree on average tended to be lower for plants than frugivores; plant node degree ranged from 1 (128 species) to 80 unique frugivorous interactions recorded for *Myrsine coriacea*; median plant degree was 5.

## Impact of Phylogenetic Imputation

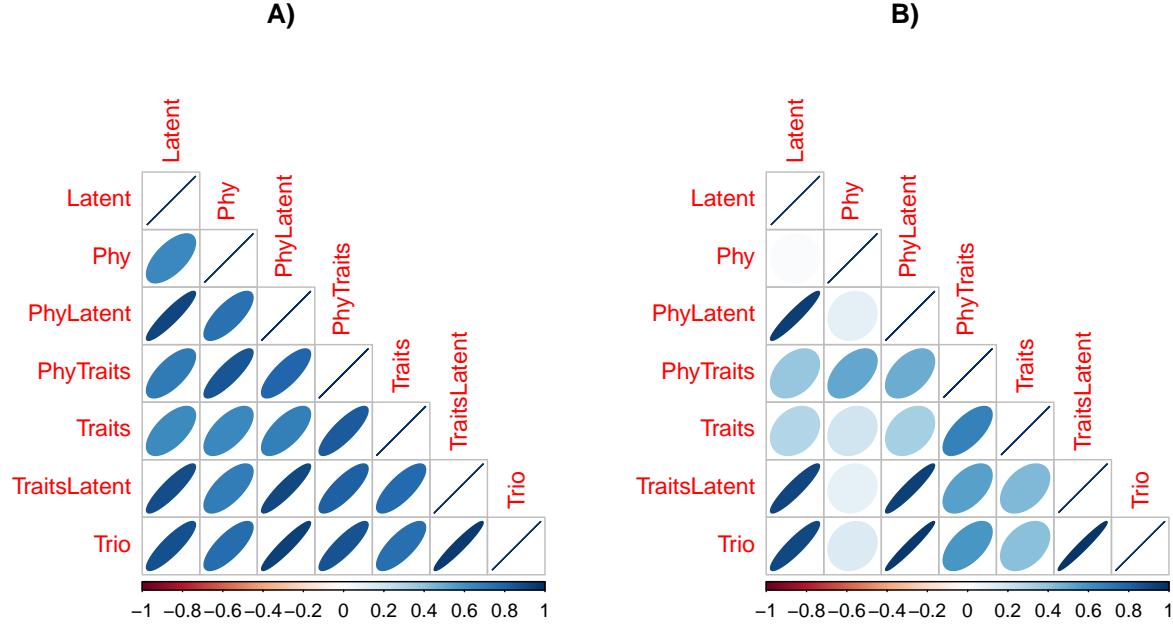
In our main text results, for the 46 plant species and 2 bird species which did not appear in the appropriate phylogenies but did have congeners, we randomly assigned those species to be a polytomy within their parent genera. When these species are excluded from the analysis, our results are still qualitatively indistinguishable. Below are versions of the three main-text figures with those 48 species removed from the network.



**Figure S3:** Summary performance metrics of all 7 models *when species without full phylogenetic information are excluded*, as measured by area under the receiver operating characteristic curve (AUC) and root mean squared error (RMSE); highest performing models for each metric are outlined in black. Mean metric values are presented from 100 replicates of each model structure alongside standard deviation. As in the main text, model discriminatory power between links and non-links is maximized by including latent structural features, with the inclusion of trait, phylogenetic information, or both actually slightly decreasing discriminatory power. However, inclusion of trait and phylogenetic information, while not improving AUC, does increase overall model accuracy as measured by mean root squared error.



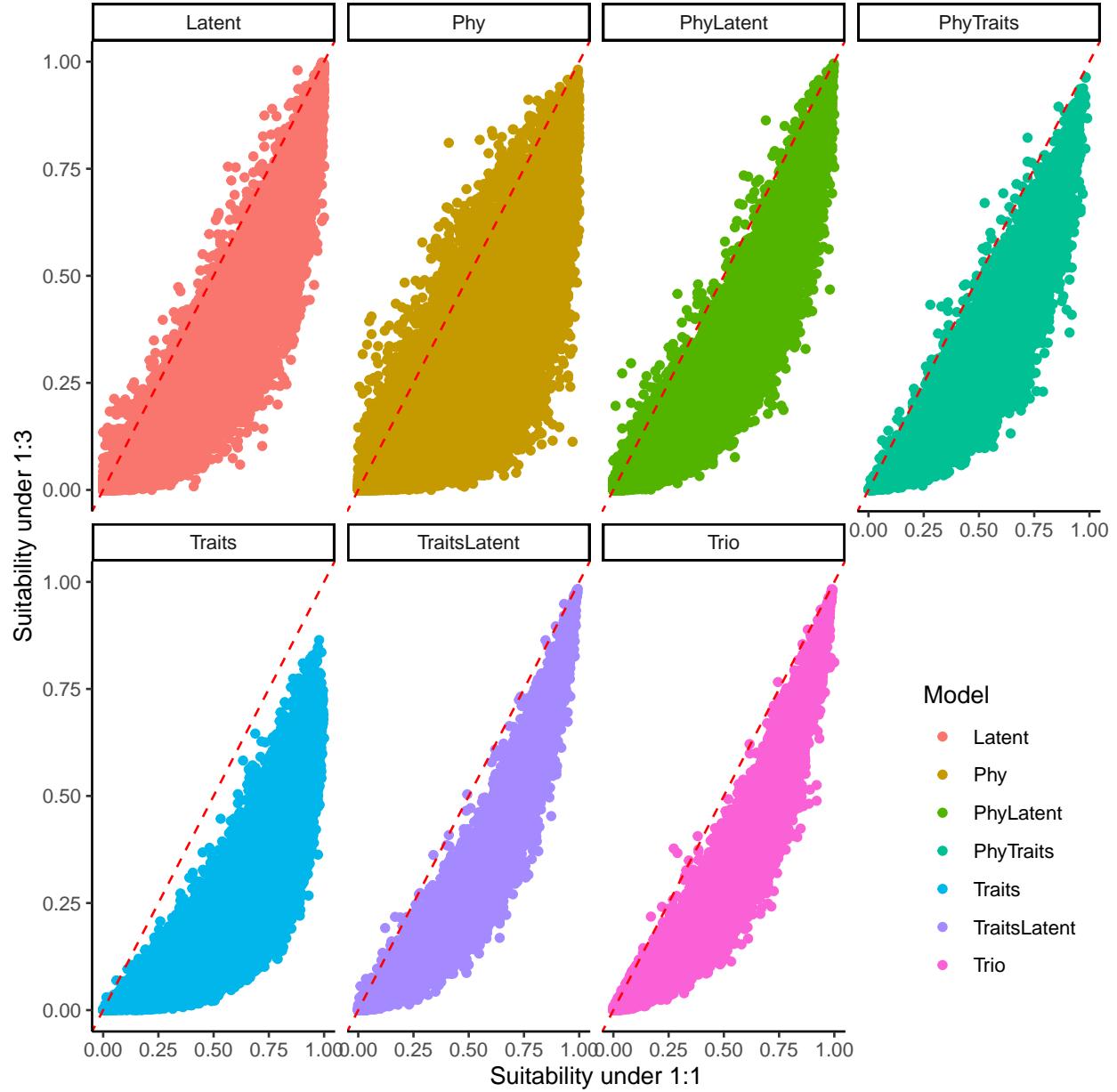
**Figure S4:** Variable importance across all models as measure by Gini importance score *when species without full phylogenetic information are excluded*; color scheme is consistent with figure S3. As in the maintext, latent traits were consistently the most important variables for prediction. These were followed by continuous frugivore traits (body mass, gape size), and frugivore phylogenetic axes. Plant phylogenies and continuous trait information were generally less important for prediction than frugivore traits. Categorical plant traits (Lipid content, fruit color, growth form) were the least important variables for prediction.



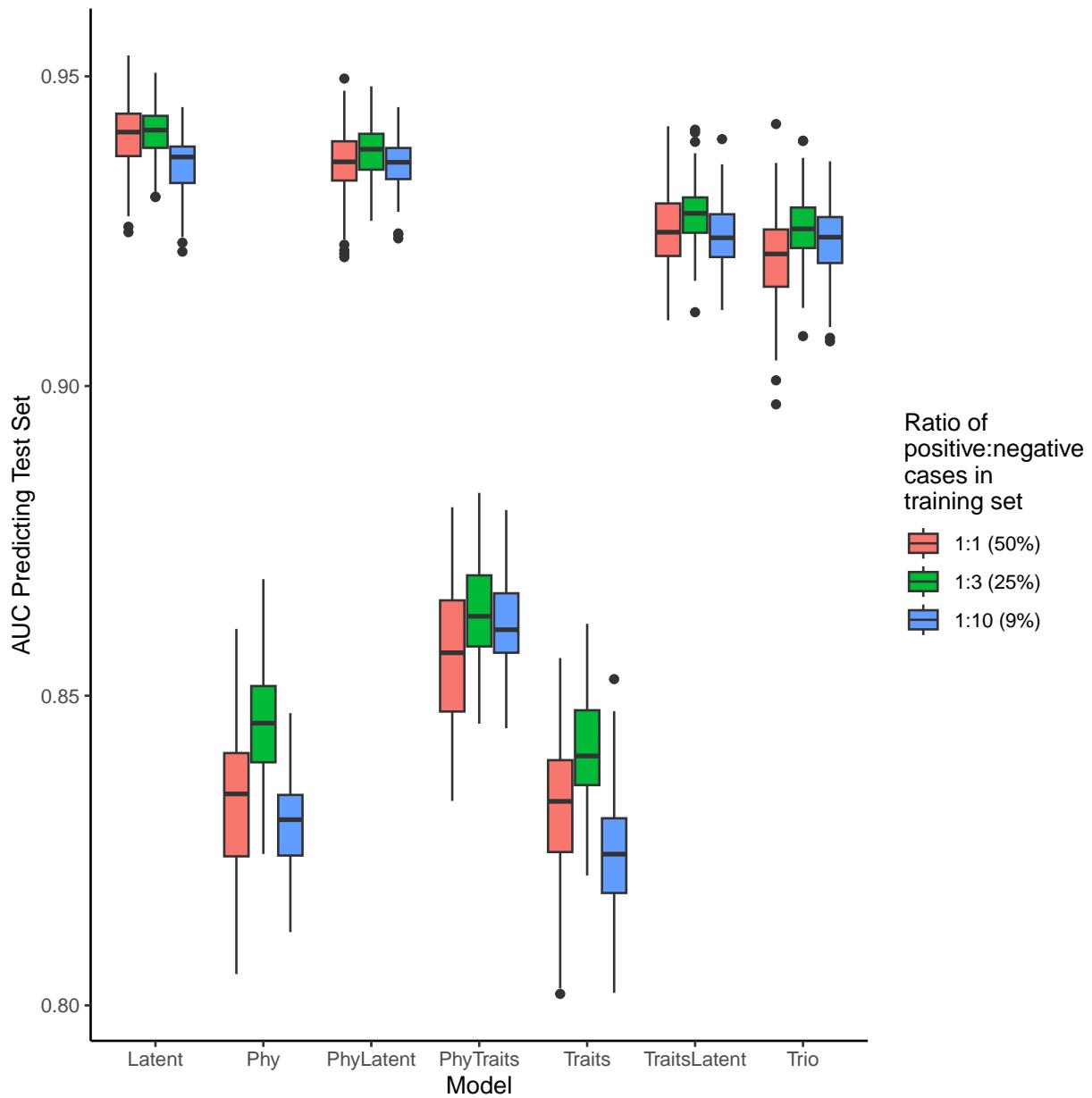
**Figure S4:** Pairwise Spearman's rank correlations of link suitabilities across models for all potential interaction (A), as well as only unobserved interactions (B), *after species without full phylogenetic information are excluded*. The latter set of links represents both true forbidden links, as well as other potential interactions not observed in our data-set.

#### Sensitivity to Class Imbalances

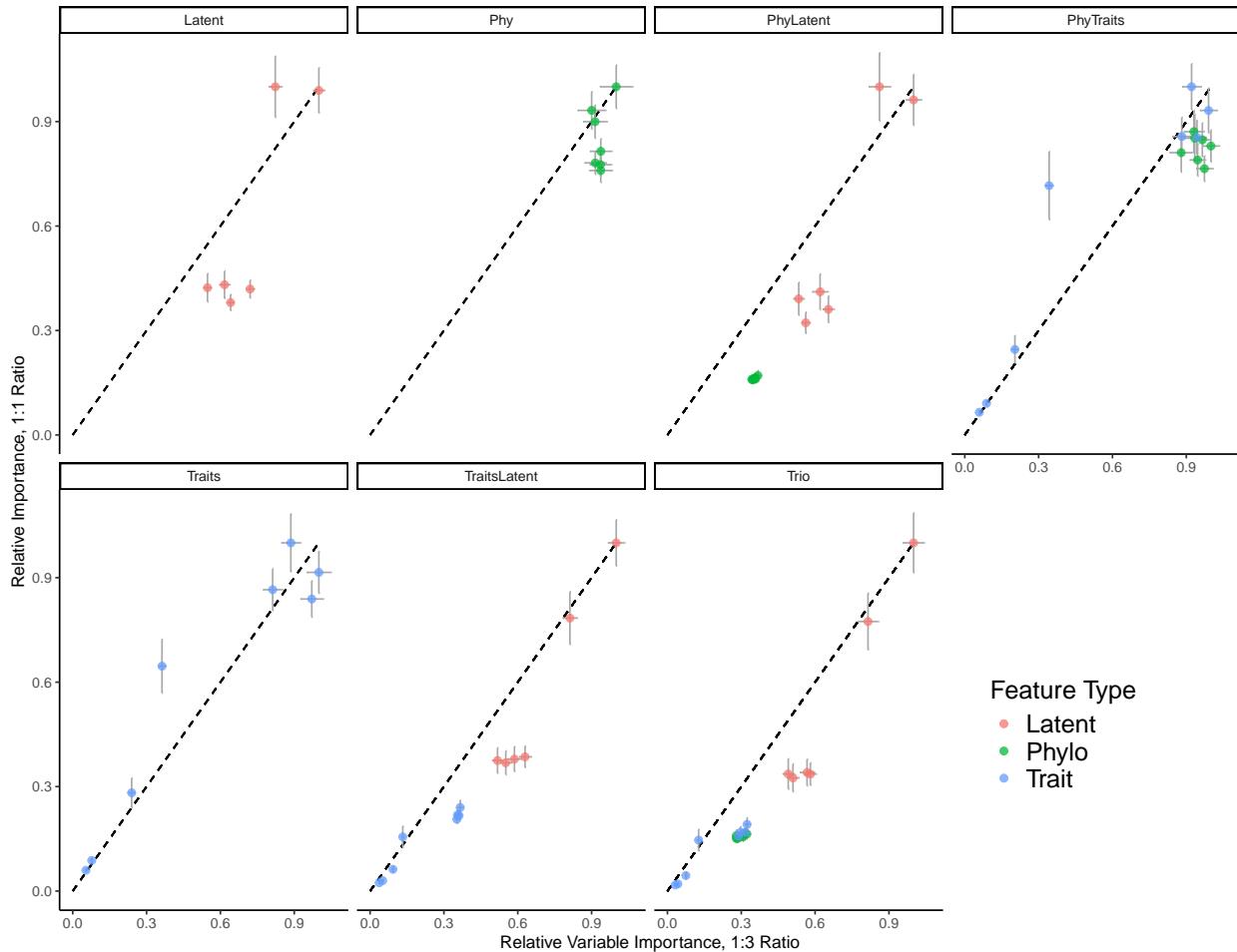
In order to deal with the high degree of sparceness in our network, we trim the training set to enforce class balancing. In the maintext, we present the results of a 3:1 ratio unobserved:observed links. Here, show the methods is relatively insensitive to this exact proportion by repeating our analysis with alternative training ratios of 1:1 and 1:10.



**Figure S5:** Scatterplot of relative suitability values of the same models trained under either a 1:3 ratio of present to absent interactions (y-axis) or a 1:1 ratio (x-axis); red dashed lines represent a 1:1 line. While the absolute value of suitability values tend to be reduced when the prevalence of true positives is reduced (most points fall below the 1:1 line), the relative rank of suitability values tends is very closely aligned (Spearman's  $\rho = 0.964$  across all predictions).



**Figure S6:** Box plots of AUC values for model predictions on a 20% test set after being conditioned on training data with either 1:1, 1:3, or 1:10 ratios of presence to absence values. We see that while there is some variation in model performance according to training prevalence, the relative performance of each model is still qualitatively the same across the range of prevalence values.



**Figure S7:** Variable importance between the same models trained on either a 1:1 ratio of present to absent interactions (y-axis) or a 1:3 ratio (x-axis). Dotted black lines represent a 1:1 line; points that fall on the line indicate that relative importance of that variable was unaffected by training prevalence. We see that the majority of features fall very close to the 1:1 line, including the most important variables (top right of each inset plot). This indicates are models are relatively insensitive to training prevalence.

## References

Gotelli N.J. (2000) Null model analysis of species co-occurrence patterns. *Ecology*, 81, 2606-2621