A phylogenomic perspective on interspecific competition

Nicolas Louw[1,], Benjamin E. Wolfe[1], Lawrence H. Uricchio[1,†]

[1]Department of Biology, Tufts University, Medford, MA 02155

email: nicolas.louw@tufts.edu, benjamin.wolfe@tufts.edu, lawrence.uricchio@tufts.edu
correspondence: [†]Lawrence H. Uricchio, Department of Biology, 200 College Ave, Medford, MA 02155

Abstract: Evolutionary processes that shape trait variation may have substantial impacts on community assembly, but evidence for phylogenetic relatedness as a determinant of interspecific interaction strength remains mixed. Several possible explanations for this mixed support have been proposed, including differences in phylogenetic scale across studies and contrasting effects of phylogeny on niche and fitness differences. Here, we use simulations and experimental data to consider the possible impact of discordance between gene-trees and species-trees on inferences of phylogenetic signal. If only a small subset of genomic loci harbor trait-related genetic variation, then the phylogeny at these loci may be more informative of trait differences than the genome background. We develop a simple method to detect genomic loci at which the local gene tree branch lengths are significantly correlated with trait differences across species relative to the species phylogeny, and apply it to genomic data and an experimental measure of competition strength in eight *Penicillium* species. We find several loci with unusually strong correlations between competitive ability and the local gene tree. Our results show that phylogenetic signal that may be masked genome-wide could be detectable using phylogenomic techniques and may provide a window into the genetic basis for interspecific interactions.

Data & code accessibility: Data and code for this project are freely available at `github.com/uricchio/ILSComp` and will be permanently archived upon publication.

Statement of authorship: NL, BEW, and LHU designed the research; NL performed experiments; NL and LHU performed computational analyses; NL, BEW, and LHU wrote the manuscript.

Introduction

Interactions between species are major determinants of the structure of ecological communities, and evolutionary processes are likely to shape traits that determine interactions. One widely-used empirical approach for characterizing the evolutionary basis of ecological interactions between closely related species compares phylogenetic distances to trait differences (Cavender-Bares et al., 2009). For each species pair in a community of interest, the evolutionary distance between the species can be regressed against some measure of the interaction strength between the species, such as a competition coefficient or relative growth rate. If competition plays a substantial role in limiting community membership, then species that co-occur naturally in biological communities might be less phylogenetically related than is expected under a null model (Webb et al., 2002). In contrast, if biotic interactions are weak determinants of community structure, environmental filtering could cause clustering of related species in space because only species with similar traits may be able to persist in a specific set of environmental conditions (Losos, 1996). In either case, evolutionary history should be informative of community composition and/or the strength of interspecific interactions.

Many empirical studies have now performed such analyses, sometimes finding that relatedness is a strong predictor of interspecific competitive differences (*e.g.*, Jiang et al. 2010; Peay et al. 2012; Godoy et al. 2014) and sometimes not (*e.g.*, Cahill Jr et al. 2008; Narwani et al. 2013; Naughton et al. 2015). The inconsistency of phylogenetic signal across studies has been discussed and debated extensively, with potential explanations including contrasting effects of relatedness on niche and fitness differences (Mayfield and Levine, 2010) and differences in the phylogenetic or spatial scales considered (Parmentier et al., 2014; Graham et al., 2018; Jarzyna et al., 2021). Though this debate continues, ecologists seem to generally agree that phylogenetic structure would be more likely to emerge for traits that determine species interactions than for traits that are unrelated to community assembly processes or outcomes (Mayfield and Levine, 2010).

Phylogenetic inference could also play an important role in inferences of phylogenetic signal. When a single phylogeny is inferred based on concatenated genetic or genomic data for a particular set of focal species, regressing trait differences on phylogenetic differences could provide evidence for or against phylogenetic signal, but provides no additional insight into the specific genomic loci that determine trait differences because the inferred evolutionary history is shared across all loci. In general, "gene trees" (*i.e.*, the evolutionary tree that describes relatedness at a single locus in the genome) do not need to share the same topology as the species tree, a phenomenon which is known as discordance between gene trees and species trees (Degnan and Rosenberg, 2009). Discordance can arise through various processes, such as introgression, gene duplication, and incomplete lineage sorting. Supposing that only a small subset of genes in the genome may be relevant to determining co-occurrence patterns, phylogenetic signal may be more likely to arise at loci that are causal for trait differences than at the level of the genome background, especially when a gene harboring variation that is causal for the trait of interest has an evolutionary history that is discordant with the species tree. One recently proposed method known as `PhyloGWAS` leverages patterns of discordance to detect associations between traits and genomic loci (Pease et al., 2016; Wu et al., 2018), providing a window into genes that may be involved in causal differences in traits across species. However, `PhyloGWAS` has not (to our knowledge) been applied in context of biotic interaction strength, perhaps in part because it was designed primarily for categorical traits.

Here, we use simulations and competitive growth experiments to investigate the possible role of discordance between

gene trees and species trees as a determinant of phylogenetic signal for competition. We develop a new phylogenomic approach for detecting associations between genes and quantitative traits given a genome-wide phylogeny and species-level trait data and apply it to eight co-occurring *Penicillium species*. The *Penicillium* genus – with over 480 species described – has been the subject of extensive genomic (Van Den Berg et al., 2008), biomedical (Vanittanakom et al., 2006), and agricultural/food science studies (Frisvad and Samson, 2004; Giraud et al., 2010), but much is still unknown about interspecific interactions that determine *Penicillium* co-occurrence patterns in nature. While many species of *Penicillium* have been domesticated and are commonly used in both food production and biomedical applications, evidence of introgression between domesticated and wild strains has been presented by multiple studies (Cheeseman et al., 2014; Steenwyk et al., 2019). Such introgression events could allow for the transfer of genetic material that affects traits underlying community assembly, and may also result in signals of discordance between gene trees at the introgressed locus and the species trees. We show that our approach identifies several genomic loci that are associated with competition strength in *Penicillium*, restoring phylogenetic signal that would not be detected with the species tree alone. We discuss the interpretation of phylogenetic signal in studies of community structure in light of discordance patterns.

Materials & methods

Conceptual model

Our approach relies on the idea that causal loci for traits can have underlying gene tree topologies that differ from the species tree. If substitutions in a causal gene cause trait differences across species, then the topology and branch lengths of the gene tree at this locus may be more informative about phylogenetic signal than the genome background (Fig. 1). To test for statistical significance, our method relies on the generation of a null distribution of correlation coefficients between phylogenetic distances and pairwise interspecific trait differences under a Multi-Species Coalescent model (Edwards et al., 2016). The approach can be applied to any phylogenomic and pairwise trait dataset (such as the absolute value of trait differences between species pairs) to infer loci that are associated with trait variation. Here, we focus on the genomic basis of competition in eight species of *Penicillium* grown in a controlled environment on a single substrate at standardized cell densities. To motivate our results, we first perform simulations using a standard quantitative trait model, under which genes that are causal for a trait may harbor substitutions that affect trait values. Species that are more genetically differentiated at the causal locus are also more likely to be more differentiated for the trait. If we were to plot phylogenetic distance between species against the magnitude of the differences in their trait values, we would expect the two values to be correlated (Fig. 1, lower right).

An obvious challenge for performing locus-specific inference under this model is that correlations between gene trees and traits may be expected genome-wide, and not only at the causal locus, because the species-level phylogeny affects diversity across the whole genome and constrains the topologies of gene trees. However, due to incomplete lineage sorting, introgression, and gene duplications, gene tree topologies can differ from the species tree topology (Degnan and Rosenberg, 2009). In the limit of very high discordance between gene trees and species trees, we expect the topology and branch lengths of the gene tree for a causal gene to be more strongly informative of trait values than the species tree. We therefore take the approach of regressing gene tree branch lengths on trait differences, scanning for correlations between gene trees and traits that are larger than expected given the species tree (Fig. 1).

Our approach is conceptually related to `PhyloGWAS` (Pease et al., 2016), which scans the genome for genes that have larger than expected numbers of nonsynonymous substitutions that perfectly correlate with a categorical trait of interest. If the trait of interest is dispersed across the species tree (rather than being concentrated into a single monophyletic clade) then correlations between substitutions and traits could be due to 1) incomplete lineage sorting, in which a the species that share the trait are monophyletic in the gene tree but not the species tree, or 2) recurrent substitutions on distinct parts of the tree or 3) introgression. `PhyloGWAS` assesses the likelihood that the substitutions arose due to incomplete lineage sorting (which is taken as a neutral null) by simulating gene trees under the multi-species coalescent model. If a gene has a larger than expected number of correlated nonsynonymous substitutions such that the null hypothesis of incomplete lineage sorting can be rejected based on the simulated distribution of substitutions, then the gene is considered to be a candidate locus for the determination of the trait.

A substitution that is shared only by the subset of species that have a particular trait value induces a split (or bipartition) among the species, *i.e.* it separates all of the species in the phylogeny into two groups. Since `PhyloGWAS` scans the genomes for genes that have a larger-than-expected number of such substitutions, it is analogous to searching for gene trees that contain a longer-than-expected branch corresponding to this bipartition. By regressing the gene tree branch lengths on the traits (rather than counting substitutions along the branch of interest) we can naturally extend this idea to quantitative traits without determining an *a priori* split between the species. Effectively, we are searching for gene trees where the number of substitutions along the branch is more strongly predictive of differences in trait values than we would expect given the species-level phylogeny. This is a desirable property for quantitative traits because there is no natural grouping of the species, as there is with a categorical trait.

Simulations

We simulated quantitative traits under a simple evolutionary model of trait differentiation. Given the phylogeny of the species, we first simulate a gene tree for a causal gene under the multi-species coalescent model. We then simulate a quantitative trait evolving neutrally on the causal gene tree through the accumulation of substitutions. To perform these simulations, we used the `Python` package `DendroPy` (Sukumaran and Holder, 2010). We fixed the species tree topology in our simulations to the inferred topology for the eight *Penicillium* species in our study (see Phylogenomic analyses).

For our quantitative trait model, we suppose that the number of substitutions occurring on a branch of length T is Poisson distributed with parameter $T\theta$, where θ is the number of substitutions expected per unit of time T. We suppose that each substitution in a causal gene has an additive effect on the trait of interest, with effect sizes that are drawn from a standard normal distribution.

We compared three broad scenarios in our simulations, which correspond to high, moderate, and low discordance between gene trees and species trees. In these simulations, we suppose that each branch of the species tree has the same population size. We set the population size parameter to be very low relative to branch lengths such that most gene trees have the same topology as the species tree for low discordance scenarios, while population size is very large relative to branch lengths for high discordance scenarios. Parameters for these simulations are given in Tab. S2. Note that since the gene trees and species trees we use in the text have branch lengths scaled in units of substitutions per base-pair, the population sizes used in the simulations do not translate to numbers of individuals, but rather the rate of coalescence relative to the branch lengths.

In general, analyses of statistical power and Type-I error rates under our model are both dependent on the model parameters for the quantitative trait evolution and the topology of the species tree. Our trait model does not account for selection, interactions between loci, unknown distributions of mutational effects, and other complicating factors. As such, we take these simulations only as proof of concept and not a full exploration of the statistical performance of our approach under plausible evolutionary models. The analyses herein cannot be taken as general statements about power, but rather comparisons across a range of scenarios that encapsulate high and low discordance between gene trees and species trees. Software implementing these analyses is available at `github.com/uricchio/ILSComp`.

Statistical inference of phylogenetic signal

We compared three inference approaches in this study. First, we performed naive correlation analyses between gene tree branch lengths and squared differences in trait values, without correcting for the structure of the species tree. Second, we used Phylogenetic Least Squares (PGLS) as implemented in `ape` version 5.6-2 and `geiger` version 2.10.0 (Paradis and Schliep, 2019; Pennell et al., 2014) to regress squared trait differences on distances calculated with gene trees, taking the species tree as a co-variate. Third, we used coalescent simulations to generate a null distribution over the correlation between trait differences and gene tree branch lengths, given a species tree and a set of trait values at the leaves of the tree. We assumed a Brownian model of trait evolution in our PGLS analyses. We assessed the ability of the latter two approaches to correct for the expected correlation in trait values due to the shared phylogeny and identify significant outliers. We describe the third approach (MSC-based simulations) in more detail in the next section.

We calculated phylogenetic distances between species in both real and simulated gene trees using `DendroPy`. When we regress trait differences on phylogenetic distances in simulated data, we compare the squared difference between trait values to the phylogenetic distance because the expectation of the signed difference in trait values is zero under a neutral evolutionary model. When we consider real *Penicillium* gene trees and traits, we use an experimentally determined measure of competition strength as the trait values (RCC, see Competition Experiments below) for each species pair as the trait value. Because RCC for a pair of species i and j has the property that $\mathrm{RCC}_{i \leftarrow j} = {}^{1}/_{\mathrm{RCC}_{j \leftarrow i}}$, we use whichever of the RCC values is greater than one in our analyses. This is analagous to using the squared trait differences as we do in the simulations, since RCC = 1 indicates competitive equivalence and larger RCC values suggest larger overall differences in competitive ability. In other words, we are not concerned with the identity of the better competitor, but rather the magnitude of competitive differences between species as a function of phylogenetic distance. We used the function `linregress` in the `Python` package `scipy.stats` to calculate correlation coefficients and uncorrected p-values for the association between trait values and phylogenetic distances.

In all regression analyses performed in this study, we separately regressed distances between a single focal species and each of its seven competitors against the competition index (RCC) for each competitior with the focal species, rather than combining all eight species into a single analysis with all 28 pairwise measures included. We do not include self comparisons in any of our regressions, because they always have an evolutionary/trait distance of 0 by definition which could bias analyses. We reasoned that some genes might be of importance in determining only a subset of competitive outcomes, and that combining all 28 pairs could therefore decrease our power to detect some associations. This does come at some cost, in the sense that we could lose power to detect genes with weak shared effects across all (or most) of the pairs in our study. To address this limitation, we also performed tests of enrichment in which we

assess the evidence for a larger than expected number of genes with multiple associations across focal species using a hypergeometric model and coalescent simulations (see Tab. 1 and Tab. S3). Note that our approach could easily be used for applications in which it is desirable to analyze all species pairs simultaneously with only minor modifications to the software.

Determining a null distribution of ρ under the MSC

To calculate the null distribution of correlation coefficients between neutrally evolving gene trees and trait values under the coalescent model, we simulate neutral gene trees on the species tree given a fixed distribution of trait values. These fixed trait values correspond either to the simulated trait values for analyses of statistical performance, or to the experimentally determined RCC values for each pair of species. We then determine a null distribution of correlation coefficients between simulated phylogenetic distances and the observed trait values by regressing the simulated non-causal branch lengths of the neutral gene trees against the trait values.

While it is straightforward to perform these simulations when all relevant model parameters are known, we do not have empirically estimated population sizes for each species of *Penicillium* in our study. Since population sizes determine the rate of coalescence along a branch of the phylogenetic tree, we need this parameter value to perform our simulations. As the amount of discordance between gene trees and species trees generally increases as a function of population size, we can use an estimate of the discordance in a set of gene trees to determine an appropriate population size for our simulations. To determine an appropriate population size, we simulate sets of gene trees at a range of different population sizes and calculate the average of the symmetric Robinson-Foulds distance between the gene trees and the species tree. We fit a linear model to this data, and interpolate the population size that best fits our dataset from the linear model (Fig. S2).

We used our simulated null distributions of correlation coefficients to determine p-values of correlation coefficients. We use one-sided p-values in this study because we are only interested in the subset of loci where correlations between gene tree branches and traits are greater than expected. The p-values were determined by computing the fraction of simulated gene trees that had correlation coefficients in excess of the observed correlation coefficient for a given gene tree in our empirical dataset.

Species sampling and selection

Penicillium strains were isolated on PCAMS (plate count agar with milk and salt; Cosetta and Wolfe 2020) from a cheese aging facility in the midwest region of the United States. Rind samples were stored in 1X phosphate buffered saline (PBS) for downstream processing. To culture fungi from samples, we diluted the cheese/slurry to 6-fold in PBS and plated out the 10^{-3} and 10^{-4} dilutions which we plated on PCAMS and supplemented with 50 mg/L chloramphenicol to inhibit bacterial growth. After maturation ($\approx$7 days), we harvested fungal spores which we were stored at -80 Celcius for molecular processing. Air samples were collected by capturing spores in the facility with an open petri dish containing cheese curd agar (CCA; Cosetta and Wolfe 2020), which were incubated at room temperature and subsequently processed like the cheese rind samples. For *in vitro* competition experiments, we chose a subset of 8 species which are phylogenetically and morphologically diverse (Bodinaku et al., 2019). The eight species are *Penicillium cvjetkovicii*, *Penicillium breivcompactum*, *Penicillium* SP. *mb*, *Penicillium commune*, *Penicillium polonicum*, *Penicillium solitum*, *Penicillium verrucosum*, and *Penicillium chrysogenum*.

Competition experiments

To construct 28 unique interspecific pairwise microcosms representing all pairs of the eight focal species in our study, we standardised cell concentrations to 20 colony forming units (CFUs)/µL following methods as described earlier (Cosetta and Wolfe, 2020). Next, we spot-plated 10µL of standardized innoculum for competing strains 2.5 cm from opposing edges of petri dishes filled with CCA. For each strain, we included a pair of intraspecific competition as a control, with the competitor replaced by another replicate of the same strain at the same cell density. All microcosms were repeated for three biological replicates. Plates were incubated in low light levels at 14 Celcius — simulating conditions of a cheese-aging facility — for 21 days. After incubation, all microcosms were scanned using a Epson Perfection V37 document scanner to acquire standardized images of competitive outcomes.

To quantify competitive outcomes, we uploaded scanned images to `ImageJ` (Abràmoff et al., 2004) and used the "Polygon" function to measure the area occupancy of each strain per microcosm. Using the area occupancy data, we calculated a competitive index for each pairwise interaction. We used the Relative Crowding Coefficient (RCC) as a measure of competitive ability, which was originally developed for understanding competitive interactions between target crops and weeds (De Wit, 1960). RCC is calculated as $\mathrm{RCC}_{ij} = \frac{Y_{ij}/Y_{ii}}{Y_{ji}/Y_{jj}}$ of species i when grown with species j, where Y is the area of occupancy after growing together for a fixed period of time. Y_{ji} corresponds to the area occupied by species j when grown with competitor species i. All measurements in this study were taken after a period of 21 days. Since input cell densities between competitors were approximately equal, an RCC of greater than 1 indicates that species i is more competitive than j, whereas with an RCC score lower than 1, j is more competitive than i. We performed three replicate experiments for each pair of competitors, and took the average value of RCC across the replicates as our measure of competitive ability.

Genomic sequencing

DNA was extracted using the Powersoil Pro DNA extraction kit (MoBio, Carlsbad, CA, USA), following kit instructions. Library preparation was performed with NEBNext® Ultra™ II DNA Library Prep (New England Biolabs, Ipswich, MA, USA), which were sequenced at Tufts University Core Facility Genomics on an Illumina NextSeq550. Low quality reads were trimmed and removed for *de novo* assembly using `GenSAS` (Humann et al., 2019), while `Augustus` (Stanke et al., 2004) was use to annotate coding regions in these fungal genomes.

Phylogenomic analyses

To identify orthologous genes between fungal strains, we used `OrthoFinder` version 2.5.4 (Emms and Kelly, 2019), with default settings. We generated gene trees of orthologous genes using `FastTree` version 2.1.10 (Price et al., 2010), and performed a multiple sequence alignment on subsequent gene trees using `MAFFT` version 7.481 (Katoh and Standley, 2013). The alignments were concatenated together in a supermatrix alignment, which was used to perform a maximum likelihood phylogenetic reconstruction with `IQ-Tree` v1.6.12 (Nguyen et al., 2015), with the HIVb + FO model of substitution, which was fit with `ModelFinder` (Kalyaanamoorthy et al., 2017). One hundred bootstrap replicates were performed to infer branch support values. Only single-copy orthologous genes, for which we identified a single ortholog in all 8 fungal species, were included in our analyses, of which there were 4,544. Gene trees and sequences for all of these genes are available at `github.com/uricchio/ILSComp`.

Tests of functional enrichment

We performed several tests for pathway-level enrichment among sets of genes that were positively correlated with RCC using KEGG annotations. To perform these tests, we used the `FungiFun 2.2.8` (available as a webtool at `https://elbe.hki-jena.de/fungifun/`; Priebe et al. 2015). We downloaded 12,789 reference *P. chrysogenum* amino acid sequences from `UniProt`, taxon ID 500485 (`https://www.uniprot.org/taxonomy/500485`). We then used `OrthoFinder` to map orthologs between our sequenced *P. chrysogenum* isolate and this *P. chrysogenum* reference. We used the set of one-to-one orthologs between these datasets to create a "background" gene set, *i.e.* a set of all genes that could possibly be considered in our test for functional enrichment. Of the 4,544 single copy orthologs, 3,629 were mapped uniquely to genes in the *P. chrysogenum* reference genome to form this background set. We created several test sets of "foreground" genes that corresponded to different strategies for detecting associations (*i.e.* different p-value thresholds and number of associated species), as described in Tab. S4. We note that these tests of enrichment should be taken with caution for multiple reasons, perhaps most importantly because functionally related genes may be located proximally to each other on the genome, meaning that their gene trees may not be fully independent. This could violate the assumptions of the hypergeometric distribution, which was used to calculate the statistical significance of pathway enrichment. When possible, we used the aforementioned *P. chrysogenum* `UniProt` reference sequence and the `EnsemblFungi` genome browser (`http://fungi.ensembl.org/Penicillium_chrysogenum_gca_000710275/`) to determine the chromosomal location of each associated gene.

Results

Simulations of phylogenetic signal under the multi-species coalescent

We simulated a neutrally evolving trait, the value of which on each of the leaf nodes is determined by substitutions that occur along each branch of the gene tree for a causal gene. The gene trees were simulated according to the multi-species coalescent model, given a species tree for the 8 *Penicillium* species in our study (see Methods). We then simulated 1,000 additional non-causal genes for each set of simulated trait values under the MSC model, given the same species tree (Fig. 1A). We repeated this procedure 1,000 times, regressing the phylogenetic distances determined by gene tree branch lengths on squared differences in trait values for the species at the leaf nodes for both the true causal gene and the 1,000 non-causal genes. Results of these regressions (summarized by p-values) are reported in Fig. 2. When discordance between gene trees and species trees is very high, the true causal gene is often more correlated with trait values than a random gene sampled from the null distribution of gene trees (Fig. 2A), suggesting that under high discordance a naive calculation of correlation coefficients would have some power to identify causal loci under this simple model. However, as the amount of discordance decreases (*i.e.*, as population sizes get smaller relative to branch lengths of the species tree), the distribution of p-values under the null looks increasingly similar to the distribution of p-values for causal gene (Fig. 2B-C). This is expected, because branch lengths and topologies are increasingly similar between gene trees and species trees when discordance is low. Low discordance results in severely inflated p-values for correlations between gene tree distances and trait differences (Fig. 2D) and a ROC curve to suggests no ability to distinguish causal loci from the background (Fig. 2E). In contrast, high discordance results in inflation of p-values only under about $p = 1e - 2$ and provides some power to distinguish causal and non-causal genes (Fig. 2D-E). Note that while this qualitative pattern is expected with any species tree, the distribution of p-values for both null and true

positive genes depends on the topology of the tree and details of the evolutionary model that generates trait differences across species.

While each of these computational experiments resulted in some degree of inflation of p-values under the null, it is clear that lower discordance results in more inflation. We therefore repeated this computational experiment, but using PGLS as implemented in `ape` and `geiger` (Paradis and Schliep, 2019; Pennell et al., 2014) as an attempt to control for phylogenetic structure. As expected, this had little effect on the analysis when discordance is high – under high discordance, the topology of gene trees is only very weakly affected by the structure of the species tree, so including the species tree as a co-variate should not affect the analysis (Fig. S1). PGLS did moderately decrease inflation of p-values under $p = 1e-1$ for low and moderate discordance (Fig. S1D), but did not fully correct the p-value distribution under the null. ROC curves were similar between PGLS (Fig. S1E) and naive correlation analyses (Fig. 2E).

MSC-based correction for phylogenetic relatedness

We next used the multi-species coalescent model to determine a null distribution over correlation coefficients between trait values and gene tree branch lengths, given a species tree and a trait distribution at the tips of the tree. As in `PhyloGWAS` (Pease et al., 2016; Wu et al., 2018), we use a simulation-based approach because calculations under the MSC are not straightforward and depend on both tree topologies and trait distributions that are not known *a priori*. We simulate gene trees and calculate the null distribution over correlation coefficients, which we can then use to identify loci with unexpectedly large correlations. We set the population size using an empirically driven approach, which uses the observed amount of discordance in a set of real gene trees (summarized by the unweighted Robinson-Foulds distance) to infer an appropriate population size (Fig. S2).

In Fig. 3, we plot both the observed distribution of correlation coefficients and the distribution determined under the MSC model for all 8 species of *Penicillium* in our experiments, where the trait values correspond to RCC (a competitive index, see Methods). In general these simulation-based null distributions capture the overall shape and location of the observed distributions. For two species, (*P. brevicompactum* and *P. cvjetkovicii*, Fig. 3A and Fig. 3D, respectively) the observed distributions were peaked in similar locations as the MSC-based distributions, but were considerably wider. One peak expected under the MSC (Fig. 3D) is visibly absent in the observed data. Nonetheless, the breadth of the simulated null distributions is generally similar to the observed distributions, suggesting that the null distributions are useful for the calculation of simulation-based p-values.

We then compared the distribution of p-values obtained under our simulation-based approach to p-values obtained through naive correlations between gene tree distances and trait values (Fig. 4). We corrected p-values by using the distribution of correlation coefficients obtained under the MSC model as a null – the corrected p-value is the proportion of simulations in which the correlation coefficient exceeds the observed correlation coefficient. Naive calculations of correlation coefficients result in distributions of p-values that can be both strongly deflated and inflated, depending on how the species tree constrains gene trees for each individual species (Fig. 4A). Corrected p-values were approximately uniformly distributed, but may slightly over-correct for some species (*e.g.*, *P. verrucosum*) or under-correct for others (*e.g.*, *P. brevicompactum*).

Genomic loci associated with the strength of competition

Given our model-based null distributions of correlations between phylogenetic distances and trait differences, we scanned for outliers with correlations that were higher than expected under the MSC-based model for each focal

species. Since each species has its own null distribution for correlations that is determined by the species tree (Fig. 3), each species has its own cutoff for significance of correlation coefficients. We include a table of all orthogroups that were significant at the $p < 1e-4$ level, along with the top BLAST hit and putative gene functions (Tab. S1). Putative functions of these top hits were varied and included potentially relevant traits such as substrate specificity, growth, and hyphal extension.

Since our dataset includes only 8 focal species, we reasoned that we might have higher power and lower susceptibility to false positives if we pooled signals across species. While each focal species shares a single gene tree at each locus from which we calculate phylogenetic distances, competition indexes (RCC values) are measured independently for each pair of species, and hence only one measurement is shared between any pair of focal species across our estimates of correlation coefficients. Given this quasi-independence of the experimental data between focal species, we reasoned that genes that had a larger number of associations across focal species might be less likely to represent spurious associations. In our dataset, we observed three loci for which at least four of the eight focal species had corrected p-values under $1e-2$ (0.066% of of the 4,544 loci included in our experiment). In 100,000 independent coalescent simulations that used the same species tree and trait distribution, we observed only five such non-causal loci (or 0.005% of simulations) for which four or more focal species had $p < 1e-2$, suggesting that our observation of $\frac{3}{4{,}544}$ is not likely under the null (hypergeometric p-value $p = 2.05e-5$). We repeated this test for enrichment of loci harboring strong correlations across multiple species, using several different thresholds for significance and finding qualitatively similar results (Tab. S3). In Tab. 1 we list the three loci that were detected with the most stringent thresholds for significance.

Since the number of associated genes exceeds the expectation under our null model, we reasoned that the set of associated genes might also be enriched for interesting biological functions. We used several different filtering strategies to create sets of associated genes and tested for pathway level enrichment in these sets using `FungiFun v2` (Priebe et al., 2015). We note that this analysis is somewhat limited from a statistical power perspective, because only a subset of genes could be mapped to the *P. chrysogenum* reference and only 1,892 genes have pathway-level annotations in `FungiFun v2`. While several pathways had low un-adjusted p-values, none survived an FDR-adjustment to account for multiple testing. Tab. S4 lists all pathways for which the FDR-adjusted p-value was below 0.25 and at least two genes in the pathway were in the foreground set. Interestingly, both DNA replication and phagosome-related genes were consistently among the top categories across several different strategies for determining the cutoff for significant associations between gene trees and RCC values.

The locus with the largest number of focal species with $p < 1e-4$ is orthogroup 5719, with associations in both *Penicillium* SP. *mb* and *Penicillium polonicum*. It also had the most associations at the $\alpha = 5e-2$ level (in six out of eight of the species we studied). The top NCBI protein-protein BLAST hit for this protein was recoverin, which is a calcium signaling protein (Tamuli et al., 2011). In Fig. 5, we plot RCC values as a function of phylogenetic distance for each of the eight focal species, seven of which display a positive relationship between to phylogenetic distance and RCC. In contrast, the relationship between RCC and the species level phylogeny (based on naive linear models) was not significant for any single focal species at the $p = 1e-2$ level (though *Penicillium verrucosum* narrowly missed this threshold with $p = 0.0102$; Fig. S3). Though we treat all statistical associations in our study with caution, we consider the possible mechanistic role for recoverin in the determination of RCC values in our Discussion.

Discussion

Relatedness between species might play a substantial role in determining the magnitude of niche and fitness differences and hence could be a major driver of community assembly, but studies have found contrasting results when assessing the evidence for relatedness as a predictor of competition strength. Many possible explanations for these discrepant findings have been proposed, including differences in phylogenetic scale, contrasting effects of relatedness on niche and fitness differences, and the potential role of ecological filtering, to name a few. Here, we considered the possible role of discordance between gene trees and species trees as a mechanism for obscuring the role of phylogeny in competitive differences. Gene trees for genes that are causal for functional differences in traits related to competitive ability might be more strongly correlated with competition-related traits than the genome background. Our phylogenomic scan for such loci revealed several genes with strong correlations, in excess of what would be expected given a Multi-Species Coalescent-based null distribution. Three genes had branch lengths that were strongly correlated with competitive differences in four or more focal species, suggesting that they may be less likely to represent spurious signals. Our results show that phylogenetic signal that would not be detected on the basis of species trees can be recovered using gene trees, and that our genomic scan for correlated gene trees can be used to infer loci that are associated with trait differences.

Inferring the genetic basis of processes underlying community assembly could help us to develop better predictive ecological frameworks and interventions, with potential applications in medicine, agriculture, and food manufacturing (Sheth et al., 2016). For example, knowing the genetic basis for species interactions might facilitate the engineering of microbes that can out-compete pathological bacteria (Riglar and Silver, 2018), or allow for modulation of the rate of adaptation in response to therapeutics such as antibiotics or phages (Blazanin and Turner, 2021). As an important step towards this goal, there has been a recent surge in the development of trait-based approaches to study factors and processes underlying microbial community assembly (Malik et al., 2020; Bittleston et al., 2021), but our understanding of traits (and their genetic basis) underlying microbial ecological processes like competition, resource consumption, and abiotic stress tolerance is still developing. Some previous approaches have used presence/absence data to quantify the association between genes and species likelihood of colonization to screen for potentially fitness-related loci (Bradley et al., 2018), or have screened many traits against a single phylogeny to infer potentially relevant traits (Ruiz et al., 2022). Our approach, which screens distinct gene trees for correlations with a trait or group of traits, is complementary to these two approaches in that we leverage the distinct signature of gene tree, species tree discordance on tree:trait correlations.

Tractable microbial systems can be particularly useful for testing hypotheses about community assembly (Cadotte et al., 2005; Vega and Gore, 2018; Chappell and Fukami, 2018). In this study, we used eight species of *Penicillium* that were sampled from a cheese cave. Cheeserind microbial communities are relatively simple in species complexity compared to many animal and soil microbiomes, making them particularly amenable to experimentation, and members of the cheeserind species pool are culturable under *in vitro* conditions which accurately reflect their natural environment. While our primary interest in this study was to consider the possible role of discordance between gene trees and species trees on the inference of phylogenetic signal, we were also interested in the possible functional roles of our candidate loci underlying competitive interactions. The gene with the most associations in our study is recoverin, a calcium signaling protein. Calcium signaling is crucial for regulating fungal growth, development, and maintenance of

calcium homeostasis (Roy et al., 2021). Both substrate acidity (Wolfe et al., 2014) the availability of soluble calcium (Amenu and Deeth, 2007) can change substantially as microbes grow during cheese production, possibly affecting the relative growth rates of fungal species growing on cheese substrates.

While recoverin *per se* is not well-studied in fungi, the homolog Frq1 is well-characterized (Tong et al., 2021). In addition to a role in regulating the circadian rhythm in fungi and mammals (reviewed by Dunlap and Loros 2017), Frq1 has been implicated in functions which could influence the RCC values measured in this study, such as growth rates on high calcium or acidic substrates (Saitoh et al., 2003) and the ability to produce conidia (asexual spores; Dunlap and Loros 2017). We also used pathway-level enrichment analyses in an effort to detect any important functions that might be enriched amongst the set of correlated genes. Although no pathways were strongly enriched after performing an FDR-adjustment, two pathways (DNA replication and phagosome) were consistently among the set of strongest associations across several different filtering strategies. Upon closer inspection of the phagosome-related genes, we noted that several were subunits of the V-type proton ATPase (Pc13g08280, subunit C, chromosome I; Pc13g14670, V0, chromosome IV; Pc15g00440, subunit F, chromosome I; Pc21g12120, subunit G, location unknown; Pc22g24360, subunit E, chromosome I), and two others were subunits of Sec61 (Pc16g09660, subunit beta, chromosome I; Pc20g02810, subunit gamma, chromosome II). Defects in the function of V-type proton ATPases have been linked to growth defects in calcium-rich environments (Vasanthakumar and Rubinstein, 2020), while Sec61 has been implicated in calcium leakage in the endoplasmic reticulum (Lang et al., 2011). Although several of these genes fall on chromosome I, these patterns cannot be easily explained by genomic proximity because most of the gene pairs were megabases apart. Nonetheless, our experiment cannot confirm a role for recoverin or other associated genes in the determination of the competitive index that we measured in our experiments, but our results could be used to prioritize such outlier-loci for subsequent experiments to determine causality. For example, gene replacement knockout studies can be used to remove and exchange orthologs of recoverin between species that have large differences in competitive abilities. Repeating the pairwise competition assay on mutant strains could confirm whether differences in competitive abilities are explained by sequence dissimilarity of recoverin. In addition to mechanism-based experiments to confirm causality of outlier loci, scanning for correlations between gene tree topology differences against trait differences beyond competition would useful to test the efficacy of our approach.

Our phylogenomic scan is conceptually related to a recent approach known as `PhyloGWAS` (Pease et al., 2016), which scans the genome for genes that have a larger than expected number of nonsynonymouys substitutions that perfectly correlate with trait differences. `PhyloGWAS` detects associations between a set of nonsynonymous substitutions and a trait of interest, taking into account the background phylogeny by employing a multi-species coalescent (MSC) model. For each gene of interest, the number of nonsynonymous alleles that perfectly correlate with the trait is calculated, and a likelihood can be assigned to this substitution pattern by performing coalescent simulations given the species tree. This process naturally controls number of false-positive loci due to hemiplasy (*i.e.*, patterns of discordance that can be well-explained by the null model of incomplete lineage sorting), and has been shown to be a powerful approach for detecting gene/trait associations. A minor limitation of `PhyloGWAS` led us to develop a new method for this study. As `PhyloGWAS` counts only perfectly correlated substitutions, it could be prone to decreased power in large sample sizes, where even one non-conforming lineage would discount the inclusion of a site of potential interest. Likewise, the inclusion of only perfectly correlated sites also limits the generalizability of method to quantitative phenotypes, where

there is no similar notion of a perfect correlation between trait values and substitution patterns. While these issues could be addressed by placing a lower threshold on the correlation-coefficient, there is not a clear *a priori* way to set such a threshold in general. Our approach of regressing gene tree branch lengths on trait values while controlling for the species tree under an MSC-based model extends this idea naturally to quantitative traits.

Like all correlative methods, our approach has several limitations. Although we took substantial care to infer an appropriate null distribution of correlation coefficients across gene trees, minor discrepancies between the null model and the true processes underlying variation in gene tree topology could cause mis-calibration and result in inflated test statistics that increase our false-positive rate. We note that our null is likely to be conservative, because we used the empirical distribution of the Robinson-Foulds metric to set the parameters of the null. Any gene trees that have discrepant phylogenies due to processes that are not well explained by the multi-species coalescent, such as introgression events or mis-inference of the true topology, will tend to increase the observed levels of discordance and hence make it more difficult to detect genes with aberrant topologies. In other words, our choice of null model should often result in a conservative reduction in power rather than an inflation in test statistics because the null is constructed from a gene set that includes both positive and negative genes, as well as introgressed genes. Like all genomic correlative methods, our approach comes with a substantial multiple testing burden. When we applied a strict Bonferroni cutoff, we retained only five genes with exceptionally strong correlations (Tab. S1), and only a handful of genes with strong correlations across multiple species (Tab. S3). However, because the gene tree topologies in our dataset and others are unlikely to be completely independent (due to evolutionarily maintained syntenic blocks and linkage disequillibrium), a less stringent threshold could be applied by assessing the level of independence and the effective number of statistical tests. We also focused only on genes for which we could reliably infer a single-copy ortholog. Some genes of interest for competitive differences may have expanded in copy number and could have been excluded from our analyses.

Characterization and prediction of the mechanistic processes that determine community structure remains a substantial goal of community ecology, with much interest in how evolutionary processes may affect species interactions. Studies comparing trait differences to phylogenetic distances have provided evidence both for and against the hypothesis that competitive interactions are stronger between closely related species, with many potential resolutions proposed (*e.g.*, Losos 1996; Godoy et al. 2014; Parmentier et al. 2014). How relatedness influences competition and coexistence between species likely depends on specific traits under selection, rather than aggregate trait differences. Genes encoding for specific traits influencing competition might not necessarily follow the same patterns of relatedness as the species tree. Our study thereby suggests a complementary role for discordance between gene trees and species trees in the interpretation of phylogenetic signal. We suggest that continued development of eco-evolutionary approaches that leverage the distinct evolutionary signatures in genes that are causal for traits, including traits that determine species interactions, will substantially enrich our understanding of the joint role of evolution and ecology in determining community assembly.

Acknowledgements

This work was partially funded by a grant from the United States National Science Foundation (CAREER IOS/BIO 1942063) to B.E.W. The authors are grateful to Dr. Megan Biango-Daniels for collecting the original strains used in this experiment. We would also like to express our gratitude to Philip Carl Bentz for helpful discussions.

References

M. D. Abràmoff, P. J. Magalhães, and S. J. Ram. Image processing with `ImageJ`. *Biophotonics International*, 11(7): 36–42, 2004.

M. Adnan, W. Zheng, W. Islam, M. Arif, Y. S. Abubakar, Z. Wang, and G. Lu. Carbon catabolite repression in filamentous fungi. *International Journal of Molecular Sciences*, 19(1):48, 2017.

B. Amenu and H. Deeth. The impact of milk composition on cheddar cheese manufacture. *Australian Journal of Dairy Technology*, 62(3):171, 2007.

L. S. Bittleston, Z. B. Freedman, J. R. Bernardin, J. J. Grothjan, E. B. Young, S. Record, B. Baiser, and S. M. Gray. Exploring microbiome functional dynamics through space and time with trait-based theory. *mSystems*, 6(4): e00530–21, 2021.

M. Blazanin and P. E. Turner. Community context matters for bacteria-phage ecology and evolution. *The ISME Journal*, 15(11):3119–3128, 2021.

I. Bodinaku, J. Shaffer, A. B. Connors, J. L. Steenwyk, M. N. Biango-Daniels, E. K. Kastman, A. Rokas, A. Robbat, and B. E. Wolfe. Rapid phenotypic and metabolomic domestication of wild penicillium molds on cheese. *mBio*, 10 (5):e02445–19, 2019.

P. H. Bradley, S. Nayfach, and K. S. Pollard. Phylogeny-corrected identification of microbial gene families relevant to human gut colonization. *PLoS Computational Biology*, 14(8):e1006242, 2018.

M. W. Cadotte, J. A. Drake, and T. Fukami. Constructing nature: laboratory models as necessary tools for investigating complex ecological communities. *Advances in Ecological Research*, 37:333–353, 2005.

J. F. Cahill Jr, S. W. Kembel, E. G. Lamb, and P. A. Keddy. Does phylogenetic relatedness influence the strength of competition among vascular plants? *Perspectives in Plant Ecology, Evolution and Systematics*, 10(1):41–50, 2008.

J. Cavender-Bares, K. H. Kozak, P. V. Fine, and S. W. Kembel. The merging of community ecology and phylogenetic biology. *Ecology Letters*, 12(7):693–715, 2009.

C. R. Chappell and T. Fukami. Nectar yeasts: a natural microcosm for ecology. *Yeast*, 35(6):417–423, 2018.

K. Cheeseman, J. Ropars, P. Renault, J. Dupont, J. Gouzy, A. Branca, A.-L. Abraham, M. Ceppi, E. Conseiller, R. Debuchy, et al. Multiple recent horizontal transfers of a large genomic region in cheese making fungi. *Nature Communications*, 5(1):2876, 2014.

C. M. Cosetta and B. E. Wolfe. Deconstructing and reconstructing cheese rind microbiomes for experiments in microbial ecology and evolution. *Current Protocols in Microbiology*, 56(1):e95, 2020.

C. T. De Wit. *On competition*. Pudoc, Wageninoen, The Netherlands, 1960.

J. H. Degnan and N. A. Rosenberg. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24(6):332–340, 2009.

J. Douglas. `UglyTrees`: a browser-based multispecies coalescent tree visualizer. *Bioinformatics*, 37(2):268–269, 2021.

J. C. Dunlap and J. J. Loros. Making time: conservation of biological clocks from fungi to animals. *Microbiology Spectrum*, 5(3):5–3, 2017.

S. Dupont, G. Lemetais, T. Ferreira, P. Cayot, P. Gervais, and L. Beney. Ergosterol biosynthesis: a fungal pathway for life on land? *Evolution*, 66(9):2961–2968, 2012.

S. V. Edwards, Z. Xi, A. Janke, B. C. Faircloth, J. E. McCormack, T. C. Glenn, B. Zhong, S. Wu, E. M. Lemmon, A. R. Lemmon, et al. Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. *Molecular Phylogenetics and Evolution*, 94:447–462, 2016.

D. M. Emms and S. Kelly. `OrthoFinder`: phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20:1–14, 2019.

J. C. Frisvad and R. A. Samson. Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*. a guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins. *Studies in Mycology*, 49(1):1–174, 2004.

F. Giraud, T. Giraud, G. Aguileta, E. Fournier, R. Samson, C. Cruaud, S. Lacoste, J. Ropars, A. Tellier, and J. Dupont. Microsatellite loci to recognize species for the cheese starter and contaminating strains associated with cheese manufacturing. *International Journal of Food Microbiology*, 137(2-3):204–213, 2010.

O. Godoy, N. J. Kraft, and J. M. Levine. Phylogenetic relatedness and the determinants of competitive outcomes. *Ecology letters*, 17(7):836–844, 2014.

G. H. Goldman, S. L. McGuire, and S. D. Harris. The DNA damage response in filamentous fungi. *Fungal Genetics and Biology*, 35(3):183–195, 2002.

C. H. Graham, D. Storch, and A. Machac. Phylogenetic scale in ecology and evolution. *Global Ecology and Biogeography*, 27(2):175–187, 2018.

J. L. Humann, T. Lee, S. Ficklin, and D. Main. Gene Prediction: Methods and Protocols. chapter "Structural and functional annotation of eukaryotic genomes with GenSAS", pages 29–51. Springer, 2019.

Y. Itoh, A. Naschberger, N. Mortezaei, J. M. Herrmann, and A. Amunts. Analysis of translating mitoribosome reveals functional characteristics of translation in mitochondria of fungi. *Nature Communications*, 11(1):5187, 2020.

M. A. Jarzyna, I. Quintero, and W. Jetz. Global functional and phylogenetic structure of avian assemblages across elevation and latitude. *Ecology Letters*, 24(2):196–207, 2021.

L. Jiang, J. Tan, and Z. Pu. An experimental test of Darwin's naturalization hypothesis. *The American Naturalist*, 175(4):415–423, 2010.

S. Kalyaanamoorthy, B. Q. Minh, T. K. Wong, A. Von Haeseler, and L. S. Jermiin. `ModelFinder`: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6):587–589, 2017.

K. Katoh and D. M. Standley. `MAFFT` multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4):772–780, 2013.

S. Lang, F. Erdmann, M. Jung, R. Wagner, A. Cavalie, and R. Zimmermann. Sec61 complexes form ubiquitous ER Ca2+ leak channels. *Channels*, 5(3):228–235, 2011.

J. B. Losos. Phylogenetic perspectives on community ecology. *Ecology*, 77(5):1344–1354, 1996.

A. A. Malik, J. B. Martiny, E. L. Brodie, A. C. Martiny, K. K. Treseder, and S. D. Allison. Defining trait-based microbial strategies with consequences for soil carbon cycling under climate change. *The ISME Journal*, 14(1):1–9, 2020.

M. M. Mayfield and J. M. Levine. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters*, 13(9):1085–1093, 2010.

M. L. L. Motta, R. R. de Melo, L. M. Zanphorlin, C. A. dos Santos, A. P. de Souza, et al. A novel fungal metal-dependent α-l-arabinofuranosidase of family 54 glycoside hydrolase shows expanded substrate specificity. *Scientific Reports*, 11(1):1–10, 2021.

A. Muszewska, S. Piłsyk, U. Perlińska-Lenart, and J. S. Kruszewska. Diversity of cell wall related proteins in human pathogenic fungi. *Journal of Fungi*, 4(1):6, 2017.

V. K. Nagarajan, C. I. Jones, S. F. Newbury, and P. J. Green. XRN 5'→ 3' exoribonucleases: structure, mechanisms and functions. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1829(6-7):590–603, 2013.

A. Narwani, M. A. Alexandrou, T. H. Oakley, I. T. Carroll, and B. J. Cardinale. Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. *Ecology Letters*, 16(11):1373–1381, 2013.

H. Naughton, M. Alexandrou, T. Oakley, and B. Cardinale. Phylogenetic distance does not predict competition in green algal communities. *Ecosphere*, 6(7):1–19, 2015.

L.-T. Nguyen, H. A. Schmidt, A. Von Haeseler, and B. Q. Minh. `IQ-TREE`: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1):268–274, 2015.

E. Paradis and K. Schliep. `ape 5.0`: an environment for modern phylogenetics and evolutionary analyses in `R`. *Bioinformatics*, 35(3):526–528, 2019.

H.-S. Park, T.-Y. Nam, K.-H. Han, S. C. Kim, and J.-H. Yu. Velc positively controls sexual development in aspergillus nidulans. *PloS One*, 9(2):e89883, 2014.

I. Parmentier, M. Réjou-Méchain, J. Chave, J. Vleminckx, D. W. Thomas, D. Kenfack, G. B. Chuyong, and O. J. Hardy. Prevalence of phylogenetic clustering at multiple scales in an african rain forest tree community. *Journal of Ecology*, 102(4):1008–1016, 2014.

J. B. Pease, D. C. Haak, M. W. Hahn, and L. C. Moyle. Phylogenomics reveals three sources of adaptive variation during a rapid radiation. *PLoS Biology*, 14(2):e1002379, 2016.

K. G. Peay, M. Belisle, and T. Fukami. Phylogenetic relatedness predicts priority effects in nectar yeast communities. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729):749–758, 2012.

M. W. Pennell, J. M. Eastman, G. J. Slater, J. W. Brown, J. C. Uyeda, R. G. FitzJohn, M. E. Alfaro, and L. J. Harmon. `geiger v2.0`: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, 30(15):2216–2218, 2014.

J. Pöhlmann, C. Risse, C. Seidel, T. Pohlmann, V. Jakopec, E. Walla, P. Ramrath, N. Takeshita, S. Baumann, M. Feldbrügge, et al. The Vip1 inositol polyphosphate kinase family regulates polarized growth and modulates the microtubule cytoskeleton in fungi. *PLoS Genetics*, 10(9):e1004586, 2014.

M. N. Price, P. S. Dehal, and A. P. Arkin. `FastTree 2`–approximately maximum-likelihood trees for large alignments. *PloS One*, 5(3):e9490, 2010.

S. Priebe, C. Kreisel, F. Horn, R. Guthke, and J. Linde. `FungiFun2`: a comprehensive online resource for systematic analysis of gene lists from fungal species. *Bioinformatics*, 31(3):445–446, 2015.

R. Ramanujam and N. I. Naqvi. PdeH, a high-affinity cAMP phosphodiesterase, is a key regulator of asexual and pathogenic differentiation in *Magnaporthe oryzae*. *PLoS Pathogens*, 6(5):e1000897, 2010.

D. T. Riglar and P. A. Silver. Engineering bacteria for diagnostic and therapeutic applications. *Nature Reviews Microbiology*, 16(4):214–225, 2018.

A. Roy, A. Kumar, D. Baruah, and R. Tamuli. Calcium signaling is involved in diverse cellular processes in fungi. *Mycology*, 12(1):10–24, 2021.

J. Ruiz, M. de Celis, J. Diaz-Colunga, J. C. C. Vila, B. Benitez-Dominguez, J. Vicente, A. Santos, A. Sanchez, and I. Belda. Predictability of the community-function landscape in wine yeast ecosystems. *bioRxiv*, 2022.

K.-i. Saitoh, T. Arie, T. Teraoka, I. Yamaguchi, and T. Kamakura. Targeted gene disruption of the neuronal calcium sensor 1 homologue in rice blast fungus, *Magnaporthe grisea*. *Bioscience, Biotechnology, and Biochemistry*, 67(3): 651–653, 2003.

W. Seufert, J. McGrath, and S. Jentsch. Ubc1 encodes a novel member of an essential subfamily of yeast ubiquitin-conjugating enzymes involved in protein degradation. *The EMBO Journal*, 9(13):4535–4541, 1990.

R. U. Sheth, V. Cabral, S. P. Chen, and H. H. Wang. Manipulating bacterial communities by in situ microbiome engineering. *Trends in Genetics*, 32(4):189–200, 2016.

M. Stanke, R. Steinkamp, S. Waack, and B. Morgenstern. `AUGUSTUS`: a web server for gene finding in eukaryotes. *Nucleic Acids Research*, 32(suppl_2):W309–W312, 2004.

J. L. Steenwyk, X.-X. Shen, A. L. Lind, G. H. Goldman, and A. Rokas. A robust phylogenomic time tree for biotechnologically and medically important fungi in the genera aspergillus and penicillium. *mBio*, 10(4):e00925–19, 2019.

S. W. Stevens and J. Abelson. Purification of the yeast U4/U6 U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proceedings of the National Academy of Sciences*, 96(13):7226–7231, 1999.

T. Stuwe, C. J. Bley, K. Thierbach, S. Petrovic, S. Schilbach, D. J. Mayo, T. Perriches, E. J. Rundlet, Y. E. Jeon, L. N. Collins, et al. Architecture of the fungal nuclear pore inner ring complex. *Science*, 350(6256):56–64, 2015.

J. Sukumaran and M. T. Holder. `DendroPy`: a python library for phylogenetic computing. *Bioinformatics*, 26(12): 1569–1571, 2010.

R. Tamuli, R. Kumar, and R. Deka. Cellular roles of neuronal calcium sensor-1 and calcium/calmodulin-dependent kinases in fungi. *Journal of Basic Microbiology*, 51(2):120–128, 2011.

S.-M. Tong, B.-J. Gao, H. Peng, and M.-G. Feng. Essential roles of two frq proteins (frq1 and frq2) in beauveria bassiana's virulence, infection cycle, and calcofluor-specific signaling. *Applied and Environmental Microbiology*, 87 (6):e02545–20, 2021.

M. A. Van Den Berg, R. Albang, K. Albermann, J. H. Badger, J.-M. Daran, A. J. M Driessen, C. Garcia-Estrada, N. D. Fedorova, D. M. Harris, W. H. Heijne, et al. Genome sequencing and analysis of the filamentous fungus penicillium chrysogenum. *Nature Biotechnology*, 26(10):1161–1168, 2008.

N. Vanittanakom, C. R. Cooper Jr, M. C. Fisher, and T. Sirisanthana. *Penicillium marneffei* infection and recent advances in the epidemiology and molecular biology aspects. *Clinical Microbiology Reviews*, 19(1):95–110, 2006.

T. Vasanthakumar and J. L. Rubinstein. Structure and roles of v-type atpases. *Trends in biochemical sciences*, 45(4): 295–307, 2020.

N. M. Vega and J. Gore. Simple organizing principles in microbial communities. *Current opinion in microbiology*, 45: 195–202, 2018.

A.-L. Veuthey and G. Bittar. Phylogenetic relationships of fungi, plantae, and animalia inferred from homologous comparison of ribosomal proteins. *Journal of Molecular Evolution*, 47:81–92, 1998.

C. O. Webb, D. D. Ackerly, M. A. McPeek, and M. J. Donoghue. Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, 33(1):475–505, 2002.

B. E. Wolfe, J. E. Button, M. Santarelli, and R. J. Dutton. Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. *Cell*, 158(2):422–433, 2014.

K. H. Wong, M. J. Hynes, R. B. Todd, and M. A. Davis. Transcriptional control of nmrA by the bZIP transcription factor MeaB reveals a new level of nitrogen regulation in *Aspergillus nidulans*. *Molecular Microbiology*, 66(2):534–551, 2007.

M. Wu, J. L. Kostyun, M. W. Hahn, and L. C. Moyle. Dissecting the basis of novel trait evolution in a radiation with widespread phylogenetic discordance. *Molecular Ecology*, 27(16):3301–3316, 2018.

Figures & Tables

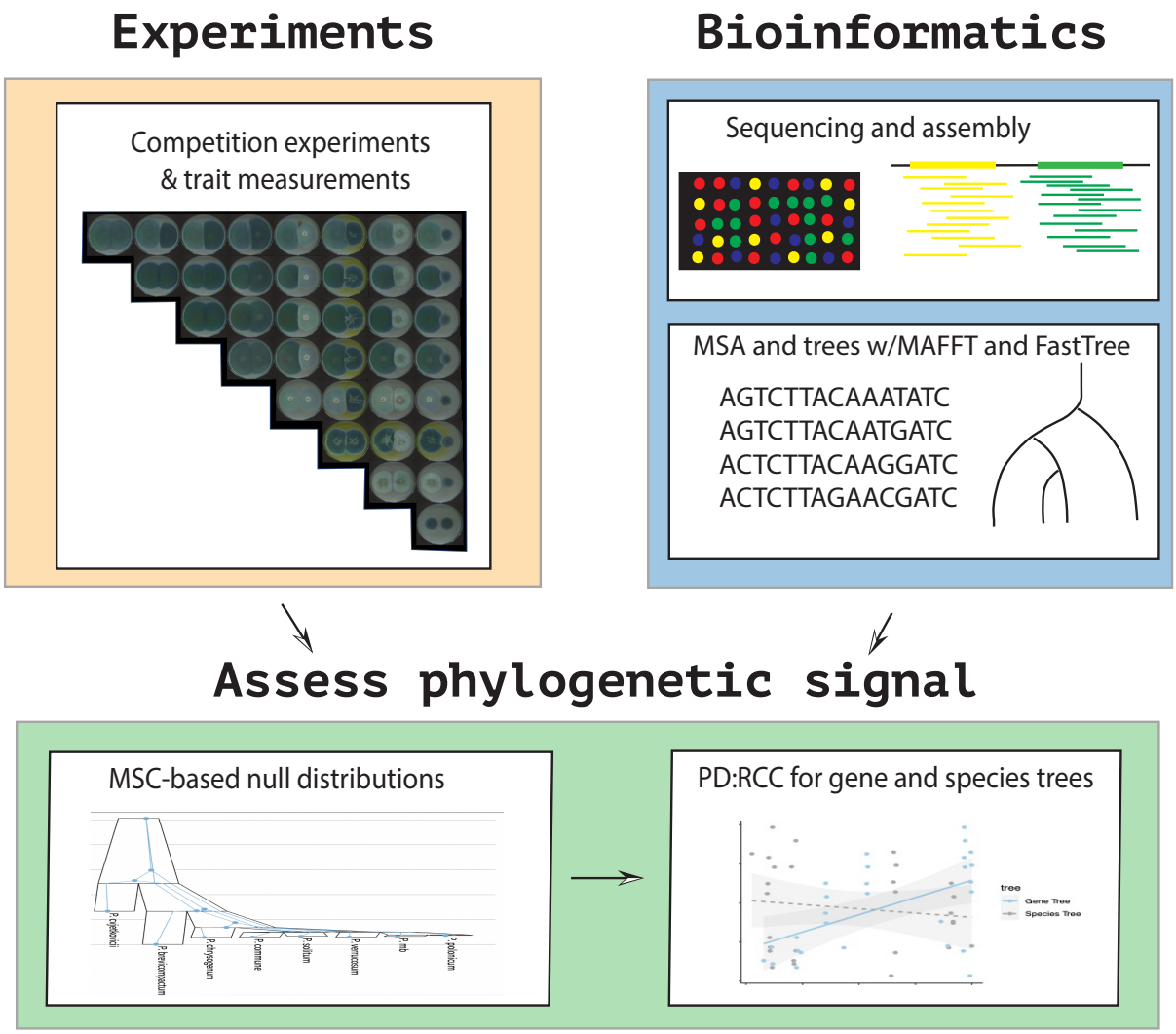

Figure 1: Flow chart of our computational and experimental analyses. We performed pairwise competition assays for eight *Penicillium* species. For each species, we built multiple sequence alignments of single copy ortholog genes (using `OrthoFinder` and `MAFFT`) and built gene trees with `FastTree`. We then assessed phylogenetic signal by comparing phylogenteic distances (PD) to trait values (a competitive index, RCC) for each gene tree. To assess significance, we used a null distribution of correlation coefficients based on the Multi-Species Coalescent (MSC) model. On the bottom left, the wide tree with black lines represents a hypothetical species tree and the blue lines represent a gene tree simulated on this species tree. We used `UglyTrees` (Douglas, 2021) to make this image. Trait values may be more correlated with gene tree branch lengths than the species tree, either due to randomness under the MSC model or because the gene tree harbors substitutions that are causal for trait differences.

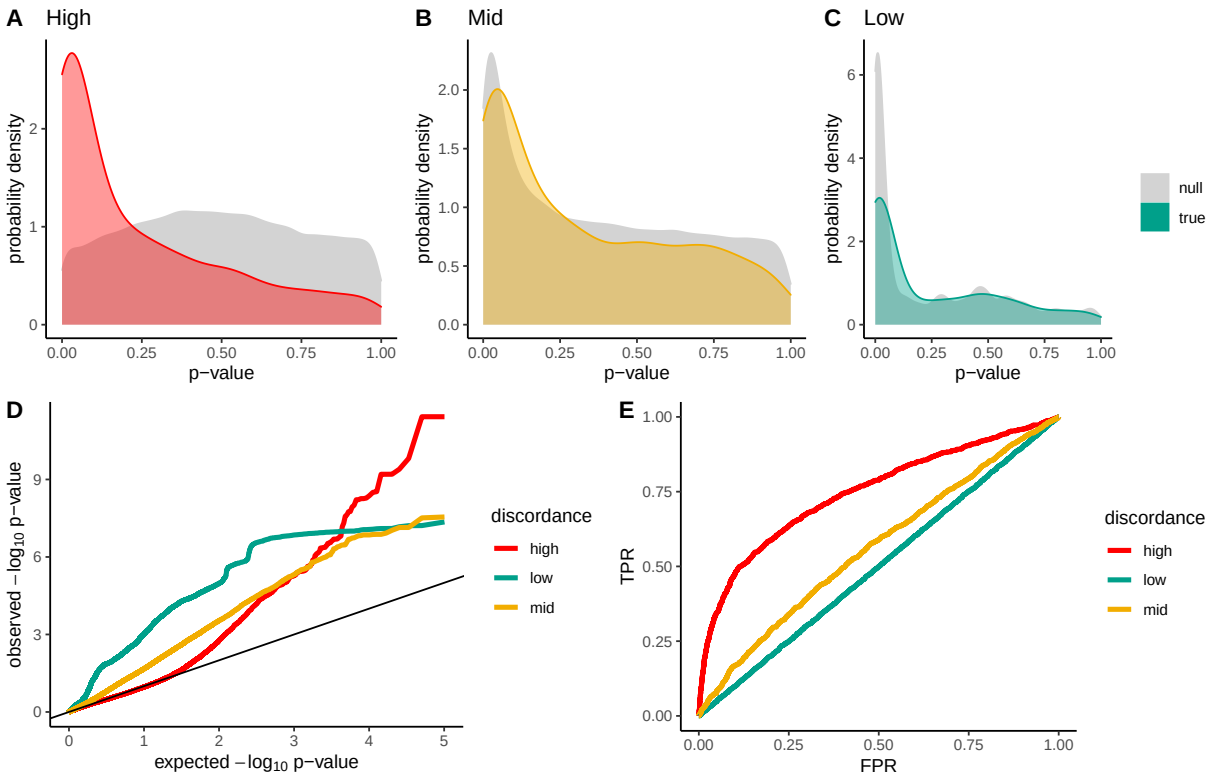

Figure 2: A-C: Probability density of p-values obtained for the correlation between phylogenetic distances and simulated trait differences in trait values under the "null" (*i.e.*, random gene trees) and for the "true" causal gene tree for each of three levels of discordance. Each panel in this figure corresponds only to differences between *P. commune* and each of its competitor species because each species has it's own null dsitribution. D: QQ-plot for p-values obtained under the null for high, low, and moderate gene/species discordance. E: ROC curve for identification of causal genes based on p-values of the correlation between gene tree branch lengths and pairwise trait differences.

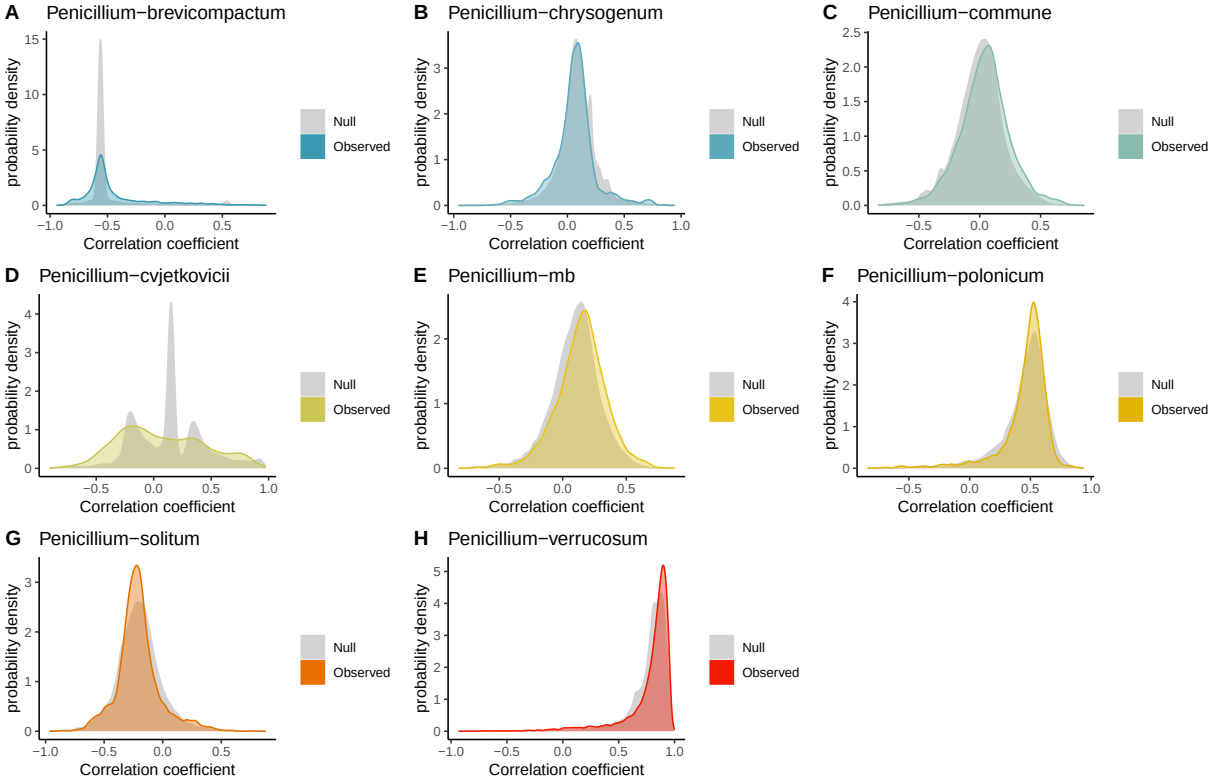

Figure 3: A-H: Null distribution of correlation coefficients between RCC and phylogenetic distances (gray) as compared to the observed distribution obtained from the collection of inferred gene trees (colors), for each focal species of *Penicillium* in our study.

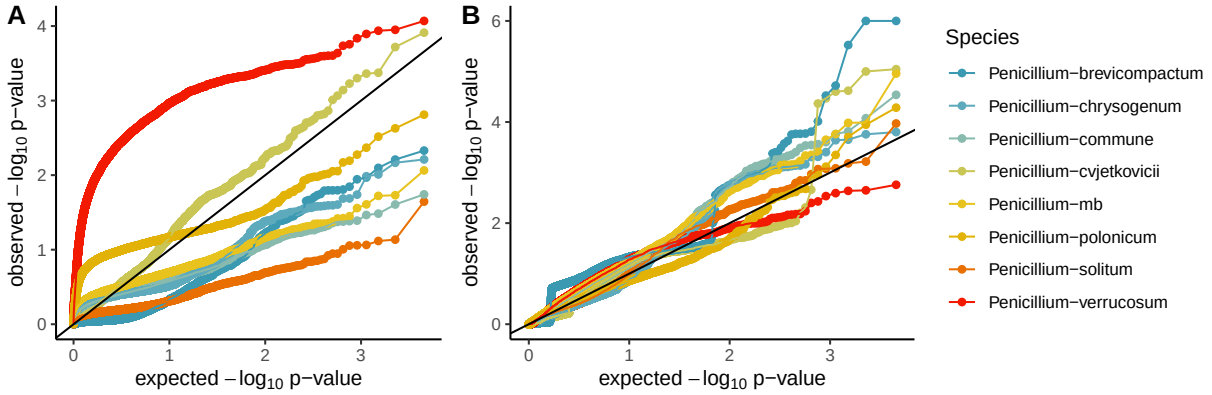

Figure 4: A: QQ-plot for p-values obtained by naive correlation analysis between RCC and phylogenetic distances, for each focal species of *Penicillium*. B: QQ-plot for p-values obtained under our simulation-based null model between RCC and phylogenetic distances, for each focal species of *Penicillium*

.

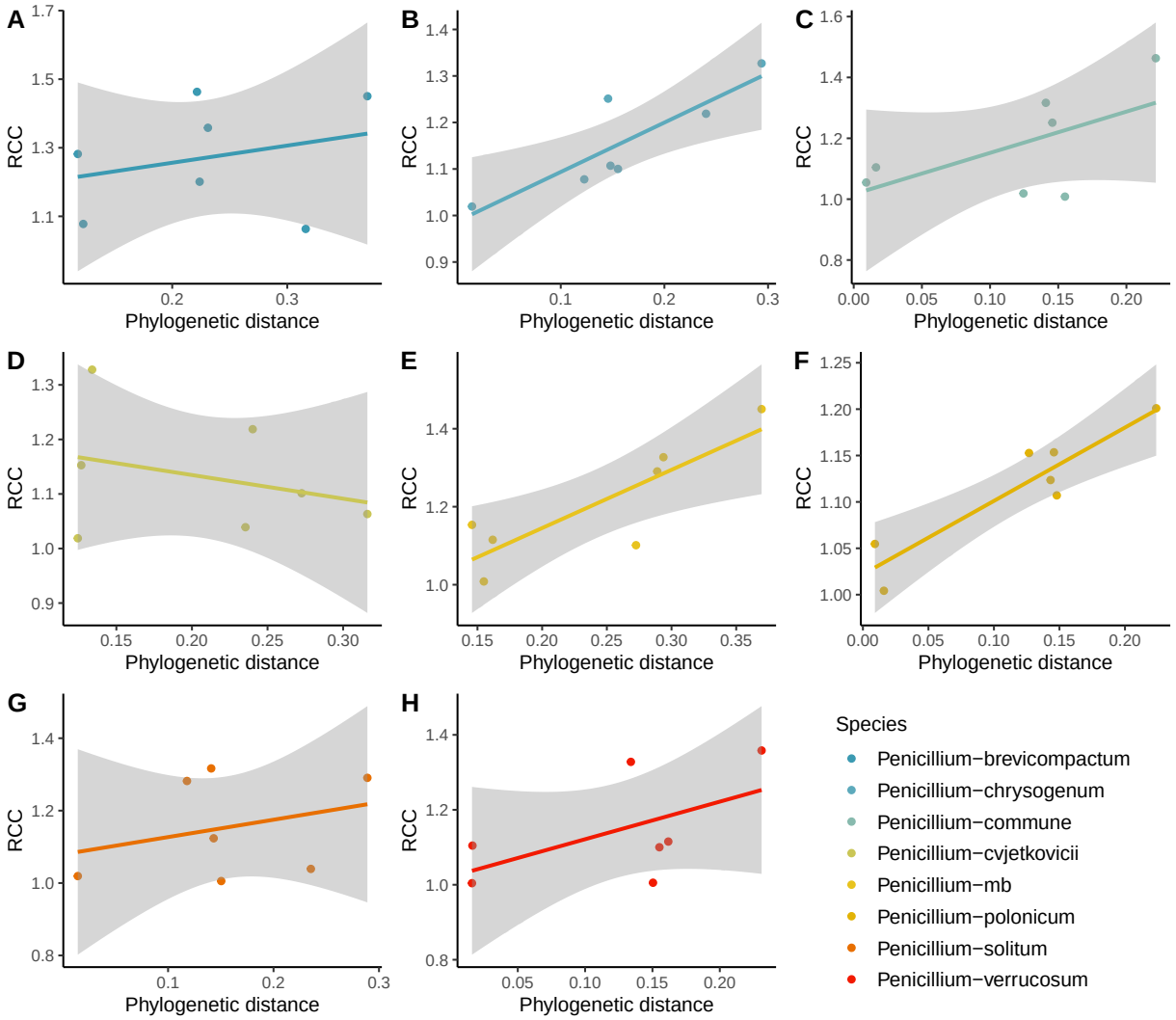

Figure 5: Relationship between phylogenetic distance and RCC for each of the 8 species for the top candidate locus in our analysis (orthogroup 5719, for which the top BLAST hit was recoverin). Linear models were fit with `geom_smooth` in the `R` package `ggplot2`, using the method `lm`.

Orthogroups with ≥ 4 species with $p < 1e-2$		
Orthogroup	BLAST hit	Function
5719	Recoverin	Calcium signaling (Tamuli et al., 2011)
4031	Ubiquitin-conjugating enzyme	Protein degredation, DNA repair, and sporulation (Seufert et al., 1990)
2568	Helicase-C-terminal	DNA replication, repair, and translation (Goldman et al., 2002)

Table 1: Three genes had four or more focal species with p-values under 0.01. All three genes (and putative functions based on BLAST hits) are listed in the above table. The orthogroup numbers can be cross-referenced with our dataset, available at `github.com/uricchio/ILSComp`.

Supplemental Figures & Tables

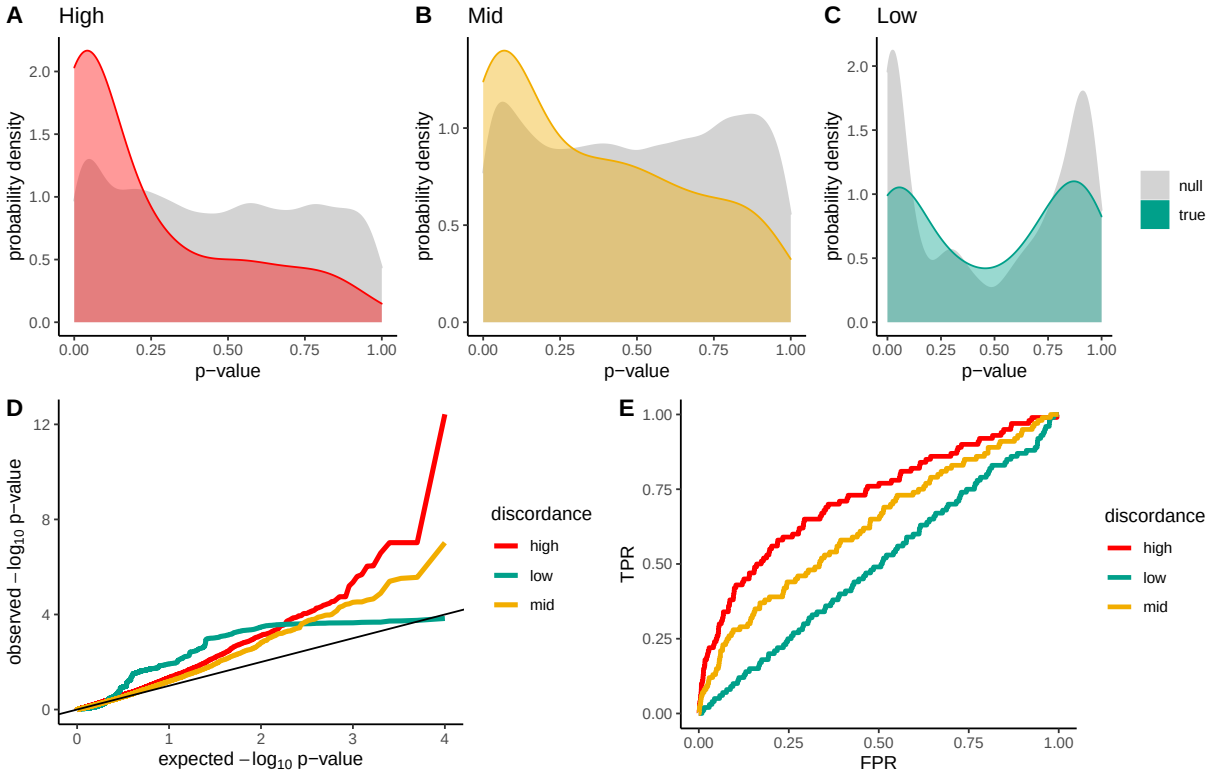

Figure S1: A-C: Probability density of p-values obtained for the correlation between phylogenetic distances and simulated trait differences in trait values under the "null" (*i.e.*, random gene trees) and for the "true" causal gene tree for each of three levels of discordance, when PGLS is used to perform the regression. Each panel in this figure corresponds only to differences between *P. commune* and each of its competitor species because each species has it's own null dsitribution. D: QQ-plot for p-values obtained under the null for high, low, and moderate gene/species discordance. E: ROC curve for identification of causal genes based on p-values of the correlation between gene tree branch lengths and pairwise trait differences

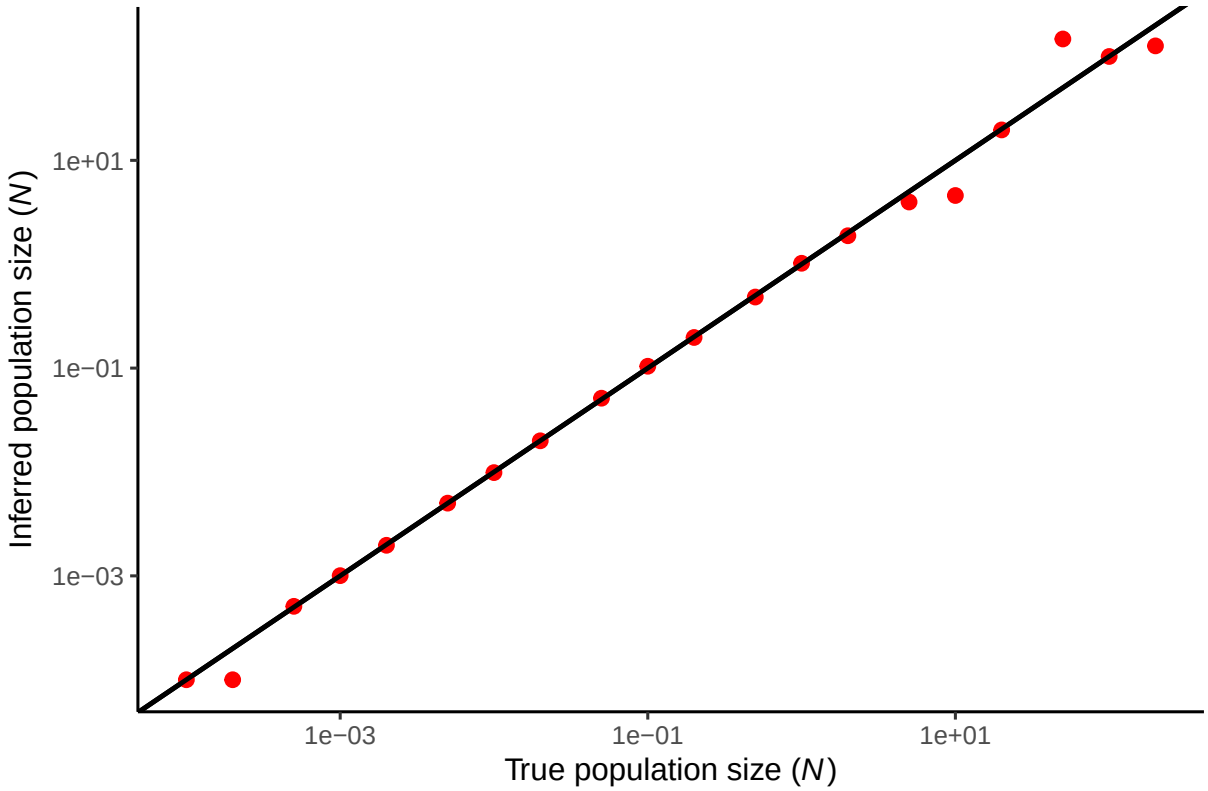

Figure S2: Inferred population sizes as a function of true population sizes using our population size inference approach. Each point represents the predicted population size for a set of 10,000 simulated gene trees, based on the observed mean of the Robinson-Foulds distance between the gene trees and the (known) species tree. Inferred sizes are less reliable when discordance is very low (small N) or very high (large N) because the Robinson-Foulds metric saturates to its minimum or maximum value, respectively, and contains less information about population size.

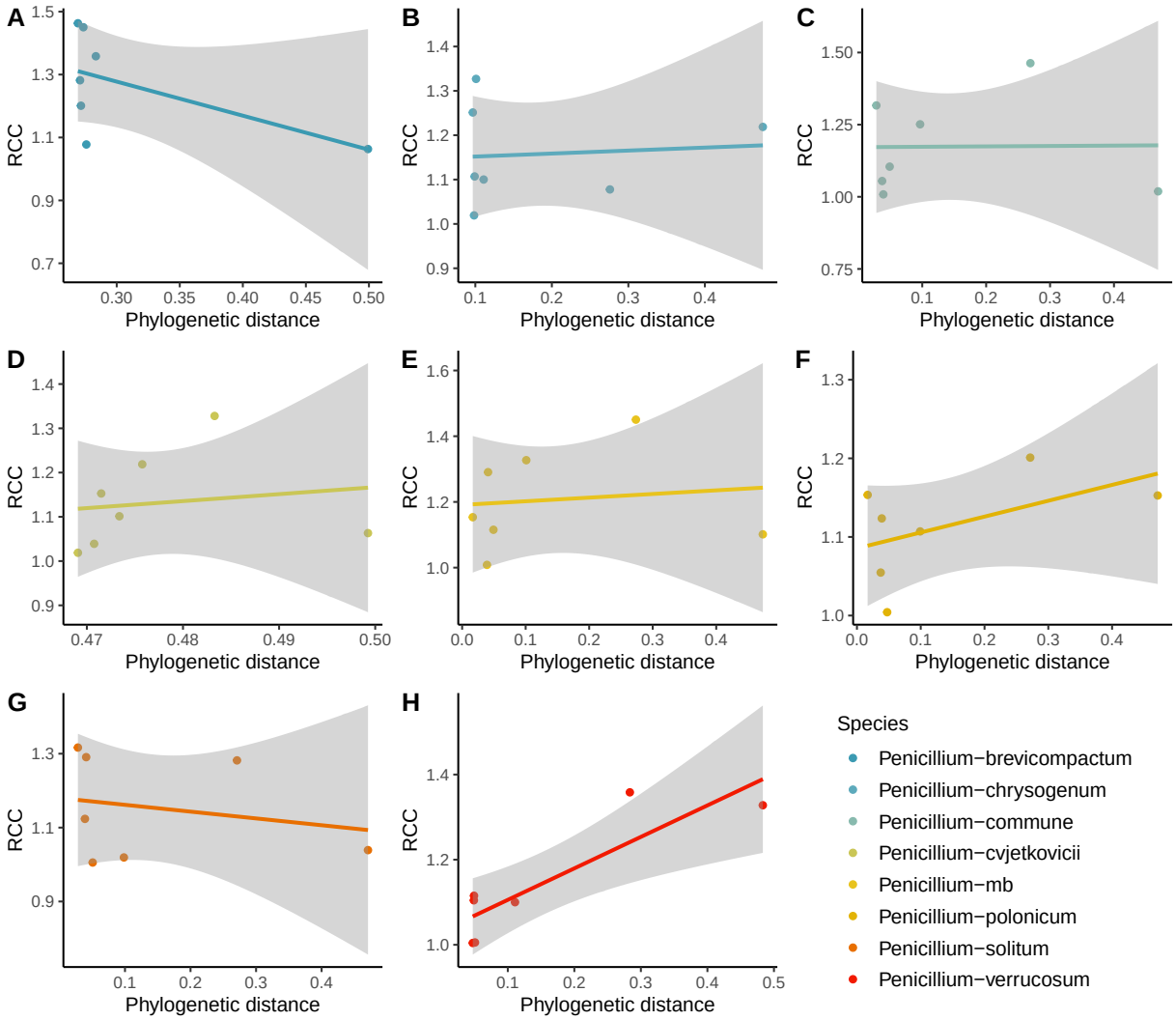

Figure S3: Relationship between phylogenetic distance and RCC for each of the 8 species for the species tree. Linear models were fit with `geom_smooth` in the `R` package `ggplot2`, using the method `lm`.

Orthogroups with $p < 1e-4$						
Focal species	Gene	ρ	p	naive p	BLAST hit	Function
Penicillium cvjetkovicii	5315	0.94745	0.00004	0.00132	Metal-dependent protein hydrolase	Substrate specificity and cellular maintenance (Motta et al., 2021)
Penicillium cvjetkovicii	2604	0.95086	0.00003	0.00110	Nuclear pore protein	Transport channel for molecular exchange between the nucleus and cytoplasm (Stuwe et al., 2015)
Penicillium cvjetkovicii	6062	0.95359	0.00003	0.00088	Type I phosphodiesterase	Regulator of cAMP signaling, influences hyphal growth and asexual development (Ramanujam and Naqvi, 2010)
Penicillium cvjetkovicii	5401	0.95409	0.00002	0.00066	Ceramide glucosyltransferase	Cell wall maintenance (Muszewska et al., 2017)
Penicillium cvjetkovicii	4273	0.96668	0.00001*	0.00044	BTB/POZ fold	Growth, development, and sporulation (Park et al., 2014)
Penicillium cvjetkovicii	2418	0.97213	0.00001*	0.00022	5'-3' exoribonuclease	RNA degradation and quality control (Nagarajan et al., 2013)
Penicillium brevicompactum	2077	0.80608	0.0001	0.00132	Ribosomal protein S13	Protein synthesis (Veuthey and Bittar, 1998)
Penicillium brevicompactum	5693	0.82370	0.00003	0.00110	Inositol transporter 1	Polarized hyphal growth (Pöhlmann et al., 2014)
Penicillium brevicompactum	4457	0.82775	0.00002	0.00088	RmlC-like cupin domain-containing protein	Unknown
Penicillium brevicompactum	6133	0.84777	<0.00001*	0.00066	Hexokinase	Carbon source sensing and growth (Adnan et al., 2017)
Penicillium brevicompactum	2994	0.86302	<0.00001*	0.00044	small nuclear ribonucleoprotein G	Growth and splicing of pre-RNA molecules (Stevens and Abelson, 1999)
Penicillium brevicompactum	5192	0.87784	<0.00001*	0.00022	Ergosterol biosynthesis protein Erg28	Maintenance of membrane stability, permeability and synthesis (Dupont et al., 2012)
Penicillium commune	5645	0.75546	0.00008	0.00044	bZIP transcription factor	Nitrogen metabolism, growth and secondary metabolism (Wong et al., 2007)
Penicillium commune	5628	0.78570	0.00003	0.00022	Protein translocation complex subunit Sss1	Translocation and quality control of protein (Itoh et al., 2020)
Penicillium SP. *mb*	5719	0.84270	0.00001	0.00022	Recoverin	Calcium signaling (Tamuli et al., 2011)
Penicillium polonicum	5719	0.92241	0.00005	0.00022	see above	see above

Table S1: Othogroups with $p < 1e-4$. Genes that were significant with a strict Bonferroni correction assuming 4,544 independent tests (one for each gene) are indicated with an asterisk. Naive p refers to the p-value given by a naive linear model, while p corresponds to our multi-species coalescent based estimate of the probability of the observed correlation under the null model. The Gene column indicates the number of the orthogroup in our dataset, which can be accessed at `github.com/uricchio/ILSComp`.

Simulation parameters		
Model	$\theta = 2Nu$	Population size (N)
Low discordance	10^4	10^{-3}
Mid discordance	10^2	10^{-1}
High discordance	10^{-4}	10^5

Table S2: Parameters relevant to our simulations of gene-tree, species-tree discordance. Lower population sizes (relative to branch lengths) result in increased levels of discordance. $\theta = 2NuL$ corresponds to the mutation rate over a gene of length L, such that the expected number of substitutions along a simulated branch of length T is $T\theta$. Note that the branch lengths are scaled in units of substitutions per site, so the population sizes here do not correspond to numbers of individuals. We allowed the mutation rate to decrease with population size, because larger population sizes result in longer branches for gene trees, and hence would have more substitutions at a given value of θ.

Tests for enrichment in the number of associated genes				
Sig. threshold	No. of focal species	Null	Observed	p-value
0.01	4	5	3	2.05e-5
0.01	3	140	23	2.15e-8
0.05	4	1,213	73	0.007
0.05	3	4,283	357	$< 2e - 16$

Table S3: Results of enrichment tests with various parameters. "Sig. threshold" corresponds to the threshold p-value at which single focal species was considered to be a positive association, and "No. of focal species" is the number of species that needed to have p-values below the threshold at a single locus in order for the locus to be considered a significant hit. The Null column indicates the number of simulations (out of 100,000) for which we observed a positive association given the parameters, and Observed indicates the number of associations we observed (out of 4,544) in our real dataset. The p column indicates the probability of the observed data given the simulations under a hypergeometric model.

Tests of functional enrichment					
Sig. threshold	Focal species	Pathway	Genes	Un-adjusted p-value	FDR-adjusted p-value
0.01	1	Phagosome	Pc13g08280, Pc13g14670, Pc15g00440, Pc20g02810	0.004733925	0.1420177
0.05	2	DNA replication	Pc12g04440, Pc13g07420, Pc16g10450, Pc20g14580, Pc21g08060, Pc21g10710, Pc21g11440, Pc21g14450, Pc21g15650, Pc21g21700	0.001998831	0.1339216
0.05	2	Phagosome	Pc13g08280, Pc13g14670, Pc15g00440, Pc16g09660, Pc18g01990, Pc20g02810, Pc21g12120, Pc22g24360	0.007726087	0.242686
0.05	2	Spliceosome	Pc12g04780, Pc13g03680, Pc13g04070, Pc13g06570, Pc13g10450, Pc13g11030, Pc16g04930, Pc16g04940, Pc20g01340, Pc20g09280, Pc20g10480, Pc20g15290, Pc21g00700, Pc21g04460, Pc22g04980	0.01086654	0.242686
0.05	3	DNA replication	Pc12g04440, Pc13g07420, Pc16g10450, Pc20g14580, Pc21g10710, Pc21g11440, Pc21g15650, Pc21g21700	0.0006394325	0.03580822
0.05	3	Phagosome	Pc13g08280, Pc13g14670, Pc16g09660, Pc18g01990, Pc20g02810, Pc21g12120	0.005832991	0.1633238
0.05	3	Mismatch repair	Pc12g04440, Pc13g07420, Pc20g14580, Pc21g11440, Pc21g21700	0.0130423	0.2434563
0.05	4	Mismatch repair	Pc20g14580, Pc21g21700	0.01857454	0.1672918
0.05	4	Spliceosome	Pc16g04930, Pc20g09280, Pc22g04980	0.02521325	0.1672918
0.05	4	DNA replication	Pc20g14580, Pc21g21700	0.03584823	0.1672918
0.05	4	Nucleotide excision repair	Pc20g14580, Pc21g21700	0.05744045	0.2010416

Table S4: Results of tests for KEGG pathway enrichment with various parameters that define the "foreground" set of genes. "Sig. threshold" corresponds to the threshold p-value for positive associations, and "Focal species" is the number of species that needed to have p-values below the threshold at a single locus in order for the locus to be considered a significant hit. "Pathway" is the name of the putatively enriched pathway, and "Genes" is the set of genes that were among the foreground set. The un-adjusted p-value column indicates the probability of the observed pathway enrichment under a hypergeometric model, while the FDR adjustment represents a correction for multiple testing. These analyses were performed with `FungiFun v.2` (Priebe et al., 2015), and the gene identifiers correspond to the *P. chrysogenum* reference identifiers in the `FungiFun v.2` database.