

<sup>1</sup> Increasing prevalence of plant-fungal symbiosis across two  
<sup>2</sup> centuries of environmental change

<sup>3</sup> Joshua C. Fowler<sup>1,2\*</sup>

Jacob Moutouama<sup>1</sup>

Tom E. X. Miller<sup>1</sup>

<sup>4</sup> 1. Rice University, Department of BioSciences, Houston, Texas 77006; 2. University of Miami,  
<sup>5</sup> Department of Biology, Miami, Florida;  
<sup>6</sup> \* Corresponding author; e-mail: jcf221@miami.edu.

<sup>7</sup> *Manuscript elements:* Figure 1 - Figure 6, Appendix A (including Figure A1 - Figure A15, Table  
<sup>8</sup> A1, and Supporting Methods).

<sup>9</sup> *Keywords:* climate change, plant-microbe symbiosis, herbarium, museum specimen, INLA,  
<sup>10</sup> spatially-varying coefficients model, *Epichloë*, *Poaceae*.<sup>1</sup>

<sup>11</sup> *Manuscript type:* Research Article.

<sup>12</sup> Prepared using the suggested L<sup>A</sup>T<sub>E</sub>X template for *Am. Nat.*

---

<sup>1</sup> *alphabetic order ?*

## Abstract

Species' distributions and abundances are shifting in response to ongoing global climate change. Mutualistic microbial symbionts can provide hosts with protection from environmental stress that may contribute towards resilience under environmental change, however this change may also disrupt species interactions and lead to declines in hosts and/or symbionts. Symbionts preserved within natural history specimens offer a unique opportunity to quantify changes in microbial symbiosis across broad temporal and spatial scales. We asked how the prevalence of seed-transmitted fungal symbionts of grasses (*Epichloë* endophytes) have changed over time in response to climate change, and how these changes vary across host species' distributions. Specifically, we examined 2,346 herbarium specimens of three grass host species (*Agrostis hyemalis*, *Agrostis perennans*, *Elymus virginicus*) collected over the past two centuries (1824 – 2019) for the presence or absence of *Epichloë* symbiosis. Analysis of an approximate Bayesian spatially-varying coefficients model implemented in INLA<sup>2</sup> revealed that endophytes increased in prevalence over the last two centuries from ca. 25% to ca. 75% prevalence, on average, across three host species. Changes in seasonal climate drivers were associated with increasing endophyte prevalence. Notably, increasing precipitation during the peak growing season for *Agrostis* species and decreasing precipitation for *E. virginicus* were associated with increasing endophyte prevalence. Changes in the variability of precipitation and temperature during off-peak seasons were also important predictors of increasing endophyte prevalence. Our analysis performed favorably in an out-of-sample predictive test with contemporary survey data, a rare extra step in collections-based research. However, we identified greater local-scale variability in endophyte prevalence in contemporary data compared to model predictions based on historic data, suggesting new directions that could improve predictive accuracy. Our results provide novel evidence for a cryp-

---

<sup>2</sup>This is the first time INLA is mentioned. It seems to be a suggestion from the reviewer. However, I think most readers may not be familiar with it. I appreciate the precision regarding the type of Bayesian statistics (approximate), but I wonder if that might be too technical for the abstract. I would understand if it were included in the methods section.

<sup>36</sup> tic biological response to climate change that may contribute to the resilience of host-microbe  
<sup>37</sup> symbiosis through **fitness benefits to symbiotic hosts.**

<sup>38</sup> Abstract : 300 words

## Introduction

40 Understanding how biotic interactions are altered by global change is a major goal of basic and  
41 applied ecological research (Blois et al., 2013; Gilman et al., 2010). Documented responses to  
42 environmental change, such as shifts in species' distributions (Aitken et al., 2008) and phenology  
43 (Piao et al., 2019), are typically blind to concurrent changes in associated biotic interactions.  
44 Empirically evaluating these biotic changes – whether interacting species shift in tandem with  
45 their partners or not (HilleRisLambers et al., 2013) – is crucial to predicting the reorganization  
46 of Earth's biodiversity under global change. Such evaluations have been limited because few  
47 datasets on species interactions extend over sufficiently long time scales of contemporary climate  
48 change (Poisot et al., 2021).

49 Natural history specimens, which were originally collected to study and preserve taxonomic  
50 diversity, present a unique opportunity to explore long-term changes in ecological interactions  
51 across broad spatial and temporal scales (Meineke et al., 2018). Natural history collections, built  
52 and maintained by the efforts of thousands of scientists, are invaluable time machines, primarily  
53 comprised of physical specimens of organisms along with information about the time and place  
54 of their collection. These specimens often preserve physical legacies of ecological processes and  
55 species' interactions from dynamically changing environments across time and space. For exam-  
56 ple, previous researchers have used plant collections (herbaria) to document shifts in phenology  
57 by examining reproductive structures (Berg et al., 2019; Park et al., 2019; Willis et al., 2017), polli-  
58 nation through examination of pollen removal (Duan et al., 2019; Pauw and Hawkins, 2011), and  
59 herbivory by documenting leaf damage (Meineke et al., 2019) related to anthropogenic climate  
60 change. However, few previous studies have leveraged biological collections to examine climate  
61 change-related shifts in a particularly common type of interaction: microbial symbiosis.

62 Microbial symbionts are common to all macroscopic organisms and can have important ef-  
63 fects on their hosts' survival, growth and reproduction (McFall-Ngai et al., 2013; Rodriguez et al.,  
64 2009). Many microbial symbionts act as mutualists, engaging in reciprocally beneficial interac-

tions with their hosts that can ameliorate environmental stress. For example, bacterial symbionts of insects, such as *Wolbachia*, can improve their hosts' thermal tolerance (Renoz et al., 2019; Truitt et al., 2019), and arbuscular mycorrhizal fungi, documented in 70-90% of families of land plants (Parniske, 2008), allow their hosts to persist through drought conditions by improving water and nutrient uptake (Cheng et al., 2021). On the other hand, changes in the mean and variance of environmental conditions may disrupt microbial mutualisms by changing the costs and benefits of the interaction for each partner, leading the interaction to deteriorate (Aslan et al., 2013; Fowler et al., 2024). Coral bleaching (the loss of symbiotic algae) due to temperature stress (Sully et al., 2019) is perhaps the best known example, but this phenomenon is not unique to corals. Lichens exposed to elevated temperatures experienced loss of photosynthetic function along with changes in the composition of their algal symbiont community (Meyer et al., 2022). How commonly and under what conditions microbial mutualisms deteriorate or strengthen under climate change remain unanswered questions (Frederickson, 2017). Previous work suggests that these alternative responses may depend on the intimacy and specialization of the interaction as well as the physiological tolerances of the mutualist partners (Rafferty et al., 2015; Toby Kiers et al., 2010; Warren and Bradford, 2014).

Understanding of how microbial symbioses are affected by climate change is additionally complicated by spatial heterogeneity in the direction and magnitude of environmental change (IPCC, 2021). Beneficial symbionts are likely able to shield their hosts from environmental stress in locations that experience a small degree of change, but symbionts in locations that experience changes of large magnitude may be pushed beyond their physiological limits (Webster et al., 2008). Additionally, symbionts are often unevenly distributed across their hosts' distribution. Facultative symbionts may be absent from portions of the host range (Afkhami et al., 2014), and hosts may engage with a diversity of partners (different symbiont species or locally-adapted strains) across their environments (Fowler et al., 2023; Fraude et al., 2008; Rolshausen et al., 2018). Identifying broader spatial trends in symbiont prevalence is therefore an important step in developing predictions for where to expect changes in the symbiosis in future climates.

92       *Epichloë* fungal endophytes are specialized symbionts of cool-season grasses, which have been  
93       documented in ~ 30% of cool-season grass species (Leuchtmann, 1992). They are transmitted  
94       vertically from maternal plants to offspring through seeds. Vertical transmission creates a feed-  
95       back between the fitness of host and symbiont (Douglas, 1998; Fine, 1975; Rudgers et al., 2009).  
96       Over time, endophytes that act as mutualists should rise in prevalence within a host population  
97       (Donald et al., 2021). *Epichloë* are known to improve their hosts' drought tolerance (Decunta  
98       et al., 2021) and protect their hosts against herbivores (Crawford et al., 2010) and pathogens (Xia  
99       et al., 2018) likely through the production of a diverse suite of alkaloids and other secondary  
100      metabolites. The fitness feedback induced by vertical transmission leads to the prediction that  
101      endophyte prevalence should be high in populations where these fitness benefits are most impor-  
102      tant. Previous survey studies of contemporary populations have documented large-scale spatial  
103      patterns in endophyte prevalence structured by environmental gradients (Afkhami, 2012; Bazely  
104      et al., 2007; Granath et al., 2007; Sneck et al., 2017). We predicted that prevalence should track  
105      temporal changes in environmental drivers (*i.e. drought*) that elicit strong fitness benefits.

106       Early research on *Epichloë* used herbarium specimens to describe the broad taxonomic di-  
107       versity of host species that harbor these symbionts (White and Cole, 1985), establishing that  
108       endophyte symbiosis could be identified in plant tissue from as early as 1851. However, no  
109       subsequent studies, to our knowledge, have used the vast resources of biological collections to  
110       quantitatively assess spatio-temporal trends in endophyte prevalence and their environmental  
111       correlates. Grasses are commonly collected and identified based on the presence of their re-  
112       productive structures, meaning that preserved specimens typically contain seeds, conveniently  
113       preserving the fungi along with their host plants on herbarium sheets. This creates the oppor-  
114       tunity to leverage the unique spatio-temporal sampling of herbarium collections to examine the  
115       response of the symbiosis to historical climate change. However, the predictive ability derived  
116       from historical analyses is rarely tested against contemporary data (Lee et al., 2024). Critically  
117       evaluating whether insights from historical reconstruction are predictive of variation across con-  
118       temporary populations is a crucial step for the field to move from reading signatures of the past

<sup>119</sup> to forecasting ecological dynamics into the future.

<sup>120</sup> In this study, we assessed the long-term responses of endophyte symbiosis to climate change  
<sup>121</sup> through the use of herbarium specimens of three North American host grass species (*Agrostis*  
<sup>122</sup> *hyemalis*, *Agrostis perennans*, and *Elymus virginicus*). We first addressed questions describing spa-  
<sup>123</sup> tial and temporal trends in endophyte prevalence: (i) How has endophyte prevalence changed  
<sup>124</sup> over the past two centuries? and (ii) How spatially variable are temporal trends in endophyte  
<sup>125</sup> prevalence across eastern North America? We then addressed how climate change may be driv-  
<sup>126</sup> ing trends in endophyte prevalence by asking: (iii) What is the relationship between temporal  
<sup>127</sup> trends in endophyte prevalence and associated changes in climate drivers? We predicted that  
<sup>128</sup> aggregate endophyte prevalence would increase over time in tandem with climate warming, and  
<sup>129</sup> that hotspots of endophyte change would correspond spatially to hotspots of climate change.  
<sup>130</sup> Finally, we evaluated (iv) how our model, built on data from historic specimens, performed in  
<sup>131</sup> an out-of-sample test - using data on endophyte prevalence from contemporary surveys of host  
<sup>132</sup> populations. To answer these questions we examined a total of 2,346 historic specimens collected  
<sup>133</sup> across eastern North America between 1824 and 2019, and evaluated model performance against  
<sup>134</sup> contemporary surveys comprising 1,442 individuals from 63 populations collected between 2013  
<sup>135</sup> and 2020.

## <sup>136</sup> Methods

### <sup>137</sup> Focal species

<sup>138</sup> Our surveys focused on three native North American grasses: *Agrostis hyemalis*, *Agrostis peren-*  
<sup>139</sup> *nans*, and *Elymus virginicus* that host *Epichloë* symbionts. These cool-season grass species are  
<sup>140</sup> commonly represented in natural history collections with broad distributions covering much the  
<sup>141</sup> eastern United States (Fig. 1). Cool-season grasses typically grow actively during the cooler tem-  
<sup>142</sup> peratures of spring and autumn due to their reliance on C<sub>3</sub> photosynthesis. *A. hyemalis* is a small  
<sup>143</sup> short-lived perennial species that germinates in spring and typically flowers between March and

<sup>144</sup> July (most common collection month: May). *A. perennans* is of similar stature but is longer lived  
<sup>145</sup> than *Agrostis hyemalis* and flowers in late summer and early autumn (most common collection  
<sup>146</sup> month: September). *A. perennans* is more sparsely distributed, tending to be found in shadier  
<sup>147</sup> and more moist habitats, while *A. hyemalis* is commonly found in open and recently disturbed  
<sup>148</sup> ground. Both *Agrostis* species are recorded from throughout the Eastern US, but *A. perennans* has  
<sup>149</sup> a slightly more northern distribution, whereas *A. hyemalis* is found rarely as far north as Canada  
<sup>150</sup> and is listed as a rare plant in Minnesota. *E. virginicus* is a larger and relatively longer-lived  
<sup>151</sup> species that is more broadly distributed than the *Agrostis* species. It begins flowering as early as  
<sup>152</sup> March or April but continues throughout the summer (most common collection month: July).

<sup>153</sup> Both *Agrostis* species host *Epichloë amarillans* (Craven et al., 2001; Leuchtmann et al., 2014),  
<sup>154</sup> while *Elymus virginicus* typically hosts *Epichloë elymi* (Clay and Schardl, 2002). The fungal sym-  
<sup>155</sup> bionts primarily reproduce asexually and are passed from mother to offspring by vertical trans-  
<sup>156</sup> mission through seeds. These traits contribute to highly specialized interactions between sym-  
<sup>157</sup> biont and host. Some host species have been shown to partner with multiple symbiont species  
<sup>158</sup> in these genus, and in some cases multiple symbiont species can co-exist within a host popu-  
<sup>159</sup> lation (Mc Cargo et al., 2014). However, suveys have typically found limited *Epichloë* genotypic  
<sup>160</sup> diversity within host populations(Treindl et al., 2023). Across host populations, concentrations  
<sup>161</sup> of biologically-active alkaloids and the genes associated with their production vary substantially  
<sup>162</sup> (Schardl et al., 2012). In this analysis, we focus on the presence/absence of *Epichloë* symbionts,  
<sup>163</sup> and we discuss potential implications of hidden genotypic diversity in the Discussion.

#### <sup>164</sup> *Herbarium surveys*

<sup>165</sup> We visited nine herbaria between 2019 and 2022 (see Table A1 for a summary of specimens in-  
<sup>166</sup> cluded from each collection). With permission from herbarium staff, we acquired seed samples  
<sup>167</sup> from 1135 *A. hyemalis* specimens collected between 1824 and 2019, 357 *A. perennans* specimens  
<sup>168</sup> collected between 1863 and 2017, and 854 *E. virginicus* specimens collected between 1839 and  
<sup>169</sup> 2019 (Fig. 1, Fig. 2A, Fig. A1). We chose our focal species in part because they are commonly

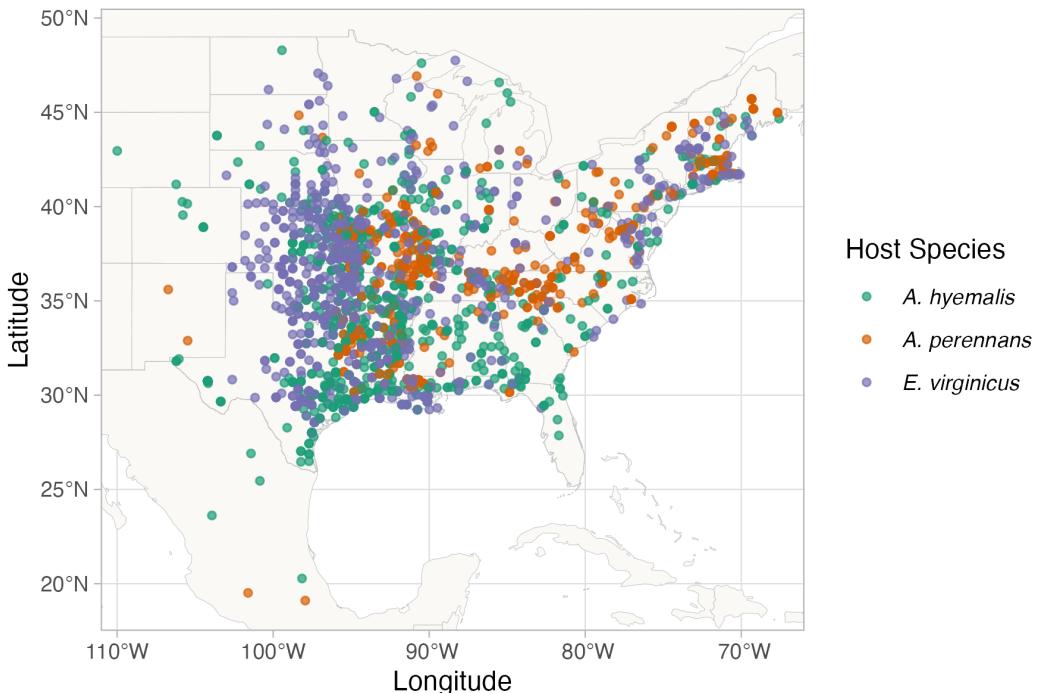
represented in herbarium collections, and produce high numbers of seeds, meaning that small samples would not diminish the value of the specimens for future studies. We collected up to 5-10 seeds per specimen after examining the herbarium sheet under a dissecting microscope to ensure that we collected mature seeds, not florets or unfilled seeds, fit for our purpose of identifying fungal endophytes with microscopy. We excluded specimens for which information about the collection location and date were unavailable. Each specimen was assigned geographic coordinates based on collection information recorded on the herbarium sheet using the geocoding functionality of the ggmap R package (Kahle et al., 2019). Many specimens had digitized collection information readily available, but for those that did not, we transcribed information printed on the herbarium sheet. Collections were geo-referenced to the nearest county centroid, or nearest municipality when that information was available. For fifteen of the oldest specimens, only information at the state level was available, and so we used the state centroid. The median pairwise distance between georeferenced coordinate points was 841 km. The median longitudinal width of the bounding boxes generated to geocode municipality, county, or state centroids was 44.7 km. Among those specimens geo-referenced at the state level, the largest bounding box, spanning the state of Texas, was 1233 km wide. The smallest bounding boxes were less than 1 km across for small municipalities (while this suggests high precision, we note that some specimens were collected in natural habitat nearby to small municipalities not encompassed by this bounding boxes).

Our visits focused on herbaria with historic strengths in *Poaceae* collections (e.g. Texas A&M, Missouri Botanic Garden) and other herbaria in the Southern Great Plains region of the United States. While these nine herbaria garnered specimens that span the focal species' ranges, our dataset unevenly samples across the study region. Texas, Oklahoma, Louisiana, and Missouri are the most represented states. Uneven sampling is most likely to be consequential for *A. perennans*, which has much of its range in the northeastern US. We explore the potential influence of spatial bias in sampling on our results through a simulation analysis (Appendix A - Supporting Methods).

After collecting seed samples, we quantified the presence or absence of *Epichloë* fungal hyphae in each specimen using microscopy. We first softened seeds with a 10% NaOH solution, then stained the seeds with aniline blue-lactic acid stain and squashed them under a microscope cover slip. We examined the squashed seeds for the presence of fungal hyphae at 200-400X magnification (Bacon and White, 2018). On average we scored 4.7 intact seeds per specimen of *A. hyemalis*, 4.2 seeds per specimen of *A. perennans*, and 3.8 seeds per specimen of *E. virginicus*; we scored 10,342 seeds in total. Due to imperfect vertical transmission, the production of symbiont-free offspring from symbiotic hosts (Afkhami and Rudgers, 2008), it is possible that symbiotic host-plants produce a mixture of symbiotic and non-symbiotic seeds. We therefore designated a specimen as endophyte-symbiotic if *Epichloë* hyphae were observed in one or more of its seeds, or non-symbiotic if *Epichloë* hyphae were observed in none of its seeds. To capture uncertainty in the endophyte identification process, we recorded both a "liberal" and a "conservative" endophyte score for each plant specimen. When we confidently identified endophytes within a specimen's seeds, we assigned matching liberal and conservative scores. When we identified potential endophytes with unusual morphology, low uptake of stain, or a small amount of fungal hyphae across the scored seeds, we recorded a positive identification for the liberal score and a negative identification for the conservative score. The liberal status assumed a potential endophyte identification was more likely to be endophyte-positive while the conservative status assumed that the potential endophyte identification was less likely to be endophyte-positive. 89% of scored plants had matching liberal and conservative scores, reflecting high confidence in endophyte status. The following analyses used the liberal status, however we repeated all analyses with the conservative status which yielded qualitatively similar results (Fig. A8).

### Modeling spatial and temporal changes in endophyte prevalence

We assessed spatial and temporal changes in endophyte prevalence across each host distribution, quantifying the "global" temporal trends averaged across space, and then examining spatial heterogeneity in the direction and magnitude of endophyte change (hotspots and coldspots) across



**Figure 1: Collection locations of herbarium specimens sampled for *Epichloë* endophytes.** Specimens span eastern North America from nine herbaria, and are colored by host species (*A. hyemalis*: green, *A. perennans*: orange, *E. virginicus*: purple). Map lines delineate study areas and do not necessarily depict accepted national boundaries.

the spatial extent of each host's distribution. To account for the spatial non-independence of georeferenced occurrences, we used an approximate Bayesian method, Integrated Nested Laplace Approximation (INLA), to construct spatio-temporal models of endophyte prevalence. INLA provides a computationally efficient method of ascertaining parameter posterior distributions for certain models that can be formulated as latent Gaussian Models (Rue et al., 2009). Many common statistical models, including structured and unstructured mixed-effects models, can be represented as latent Gaussian Models. We incorporated spatial heterogeneity into this analysis using spatially-structured intercept and slope parameters implemented as stochastic partial differential equations (SPDE) to approximate a continuous spatial Gaussian process. This SPDE approach is a flexible method of smoothing across space while explicitly accounting for spatial

dependence between data-points (Bakka et al., 2018; Lindgren et al., 2011). Fitting models with structured spatial effects is possible with MCMC sampling but can require long computation times, making INLA an effective alternative. This approach has been used to model spatial patterns in flowering phenology (Willems et al., 2022), the abundance of birds (Meehan et al., 2019) and butterflies (Crossley et al., 2022), the distribution of temperate trees (Engel et al., 2022) as well as the population dynamics of endangered amphibians (Knapp et al., 2016) and other ecological processes (Beguin et al., 2012).

We estimated global and spatially-varying trends in endophyte prevalence using a joint-likelihood model. For each host species  $h$ , endophyte presence/absence of the  $i^{th}$  specimen ( $P_{h,i}$ ) was modeled as a Bernoulli response variable with expected probability of endophyte occurrence  $\hat{P}_{h,i}$ . We modeled  $\hat{P}_{h,i}$  as a linear function of intercept  $A_h$  and slope  $T_h$  defining the global trend in endophyte prevalence specific to each host species as well as with spatially-varying intercepts  $\alpha_{h,l_i}$  and slopes  $\tau_{h,l_i}$  associated with location ( $l_i$ , the unique latitude-longitude combination of the  $i^{th}$  observation). The joint-model structure allowed us to “borrow information” across species in the estimation of shared variance terms for the spatially-dependent random effect  $\delta_{l_i}$ , intended to account for residual spatial variation, and  $\chi_{c_i}$  and  $\omega_{s_i}$  i.i.d.-random effects indexed for each collector identity ( $c_i$ ), and scorer identity ( $s_i$ ) of the  $i^{th}$  specimen.

$$\text{logit}(\hat{P}_{h,i}) = A_h + T_h * \text{year}_i + \alpha_{h,l_i} + \tau_{h,l_i} * \text{year}_i + \delta_{l_i} + \chi_{c_i} + \omega_{s_i} \quad (1)$$

By including random effects for collectors and scorers, we accounted for “nuisance” variance that may bias predictions for changes in endophyte prevalence. Previous work suggests that behavior of historical botanists may introduce biases into ecological inferences made from historic collections (Kozlov et al., 2020). Prolific collectors who contribute thousands of specimens may be more or less likely to collect certain species, or specimens with certain traits (Daru et al., 2018). Similarly, the process of scoring seeds for hyphae involved multiple researchers (or "scorers") who, even with standardized training, may vary in their likelihood of positively identifying *Epichloë*.

258 We performed model fitting using the inlabru R package (Bachl et al., 2019). Global intercept  
259 and slope parameters  $A$ , and  $T$ , were given vague priors. Collector and scorer random effects,  
260  $\chi$  and  $\omega$  respectively, were centered at 0 with precision parameters were assigned penalized  
261 complexity (PC) priors with parameter values  $U_{PC} = 1$  and  $a_{PC} = 0.01$  (Simpson et al., 2017). Each  
262 spatially-structured parameter depended on a covariance matrix according to the proximity of  
263 each pair of collection locations (Bakka et al., 2018; Lindgren et al., 2011). The covariance matrix  
264 was approximated using a Matérn covariance function, with each data point assigned a location  
265 according to the nodes of a mesh of non-overlapping triangles encompassing the study area (Fig.  
266 A2). We assessed model fit with visual posterior predictive checks (A3) and measurements of  
267 AUC (Figs. A4-A5) (Gelman and Hill, 2006). Priors for the Matérn covariance function, termed  
268 "range" and "variance", define how proximity effects decay with distance. Results presented in  
269 the main text reflect a prior range of 342 kilometers (i.e. a 50% probability of estimating a range  
270 less than 342 kilometers). We tested a range of values (from 68 kilometers to 1714 kilometers) and  
271 meshes (presented in the Supporting Methods), finding that while the magnitude and uncertainty  
272 of effects varied, model results were qualitatively similar, i.e. the same direction of effects across  
273 space. Through results and discussion that follow, we refer to the model described in this section  
274 as the "endophyte prevalence model".

### 275 *Modeling distributions of host species*

276 We modeled the geographic distribution of each host species with two goals: (1) generate realistic  
277 maps on which we could project the predictions of the INLA model, and (2) use the geographic  
278 distributions to test for relationships between climate change drivers and trends in endophyte  
279 prevalence. The herbarium records did not encompass the entirety of each host species' range.  
280 We followed the ODMAP (overview, data, model, assessment, prediction) protocol (Crossley  
281 et al., 2022) (see Supporting Methods). In short, we used presence-only observations of each host  
282 species from Global Biodiversity Information Facility (GBIF) between 1990 to 2020. We fit max-  
283 imum entropy (MaxEnt) models using the maxent function in the R package dismo (Hijmans

284 et al., 2017) using the same set of seasonal climate predictors considered below calculated for  
285 the 1990-2020 climate normals: mean and standard deviation of spring, summer, and autumn  
286 temperature, and mean and standard deviation of spring, summer, and autumn cumulative pre-  
287 cipitation. We generated 10,000 pseudo-absences as background points, and split the occurrence  
288 data into 75% for model training and 25% for model testing. The performance of models was  
289 evaluated with AUC (Jiménez-Valverde, 2012). We found AUC values of 0.862, 0.838, 0.821 re-  
290 spectively for *Agrostis hyemalis*, *Agrostis perennans*, and *Elymus virginicus* indicating good model  
291 fit to data. To convert the continuous predicted probabilities into binary presence - absence maps  
292 on which we projected INLA predictions, we used the training sensitivity (true positive rate) and  
293 specificity threshold (true negative rate) (Liu et al., 2005).

#### 294 *Assessing the role of climate drivers*

295 We assessed how the magnitude of climate change may have driven changes in endophyte preva-  
296 lence by assessing correlations between changes in climate and changes in endophyte prevalence  
297 predicted from our spatial model at evenly spaced pixels across the study area. We first down-  
298 loaded monthly temperature and precipitation rasters from the PRISM climate group (Daly and  
299 Bryant, 2013) covering the time period between 1895 and 2020 using the 'prism' R package (Hart  
300 and Bell, 2015). Prism provides reconstructions of historic climate variables across the United  
301 States by spatially-interpolating weather station data (Di Luzio et al., 2008). We calculated 30-  
302 year climate normals for seasonal mean temperature and cumulative precipitation for the recent  
303 (1990 to 2020) and historic (1895 to 1925) periods. We used three four-month seasons within  
304 the year (Spring: January, February, March, April; Summer: May, June, July, August; Autumn:  
305 September, October, November, December). This division of seasons allowed us to quantify dif-  
306 ferences in climate associated with the two "cool" seasons, when we expected our focal species  
307 to be most biologically active (*A. hyemalis* flowering phenology: spring; *E. virginicus*: spring and  
308 summer; *A. perennans*: autumn). In addition to mean climate conditions, environmental vari-  
309 ability itself can influence population dynamics (Tuljapurkar, 1982) and changes in variability

are a key prediction of climate change models (IPCC, 2021; Stocker et al., 2013). Therefore, we calculated the standard deviation for each annual and seasonal climate driver across each 30-year period. We then took the difference between recent and historic periods for the mean and standard deviation for each climate driver (Figs. A13-A15). All together, we assessed twelve potential climate drivers: the mean and standard deviation of spring, summer, and autumn temperature, as well as the mean and standard deviation of spring, summer, and autumn cumulative precipitation.

We then evaluated whether areas that have experienced the greatest changes in endophyte prevalence (hotspots of endophyte change) are associated with high degrees of change in climate (hotspots of climate change). To do so, we modeled the fitted, spatially-varying slopes of endophyte change through time ( $\tau_{[h]l}$ ) as a linear function of environmental covariates, with a Gaussian error distribution for a set of pixels across each host distribution. The continuous SPDE approach taken for our endophyte prevalence model allows us to generate predictions of temporal trends in prevalence at arbitrarily many pixels across each host distribution. Balancing computation time with resolution, we generated predicted trends for 546, 645, and 753 pixels across each host distribution for *A. perennans*, *A. hyemalis*, and *E. virginicus* respectively (pixel dimensions: *A. perennans* = 65 km x 36 km; *A. hyemalis* = 61km x 45 km; *E. virginicus* = 62 km x 40 km). Fitting regressions to many pixels across the study region risks artificially inflating confidence in our results due to large sample sizes, and so we performed this analysis using only a random subsample of 250 pixels across the study region; other sizes of subsample yielded similar results. Data from each host species were analyzed separately. Through the results and discussion that follow, we refer to this analysis as the “*post hoc* climate regression analysis”.

### Validating model performance with in-sample and out-of-sample tests

We evaluated the predictive ability of the model using both in-sample training data from the herbarium surveys, and with out-of-sample test data, an important but rarely used strategy in ecological studies (Lee et al., 2024; Tredennick et al., 2021). We generated out-of-sample test

336 data from contemporary surveys of endophyte prevalence in natural populations of *A. hyemalis*  
337 and *E. virginicus* in Texas and the southern US. Surveys of *E. virginicus* were conducted in 2013  
338 as described in Sneck et al. (2017), and surveys of *A. hyemalis* took place between 2015 and  
339 2020. Population surveys of *A. hyemalis* were initially designed to cover longitudinal variation  
340 in endophyte prevalence towards its range edge, while surveys of *E. virginicus* were designed to  
341 cover latitudinal variation. In total, we visited 43 populations of *A. hyemalis* and 20 populations  
342 of *E. virginicus* across the south-central US, with emphasis on Texas and neighboring states (Fig  
343 A12). During surveys, we collected seeds from up to 30 individuals per population (average  
344 number of plants sampled per population: 22.9); note that this sampling design provided greater  
345 local depth of information than the herbarium records, where only one plant was sampled at  
346 each locality. We quantified the endophyte status of each individual with staining microscopy  
347 as described for the herbarium surveys (with 5-10 seeds scored per individual), and calculated  
348 the prevalence of endophytes within the population (proportion of plants that were endophyte-  
349 symbiotic). For each population, we compared the observed fraction of endophyte-symbiotic  
350 hosts to the predicted probability of endophyte occurrence  $\hat{P}$  derived from the model for that  
351 location and year. The contemporary survey period (2013-2020) is at the most recent edge of the  
352 time period encompassed by the historical observations used for model fitting.

353

## Results

354 *How has endophyte prevalence changed over time?*

355 Across more than 2300 herbarium specimens dating back to 1824, we found that prevalence of  
356 *Epichloë* endophytes increased over the last two centuries for all three grass host species (Fig.  
357 2). On average, endophytes of *A. perennans* and *E. virginicus* increased from ~ 40 % to 70%  
358 prevalence across the study region, and *A. hyemalis* increased from ~ 25% to over 50% prevalence.  
359 Our model indicates a high **confidence** that overall temporal trends are positive across species  
360 (99% probability of a positive overall year slope in *A. hyemalis*, 92% probability of a positive

<sup>361</sup> overall year slope in *A. perennans*, and 91% probability of a positive overall year slope in *E.*  
<sup>362</sup> *virginicus*) (Fig. A6).

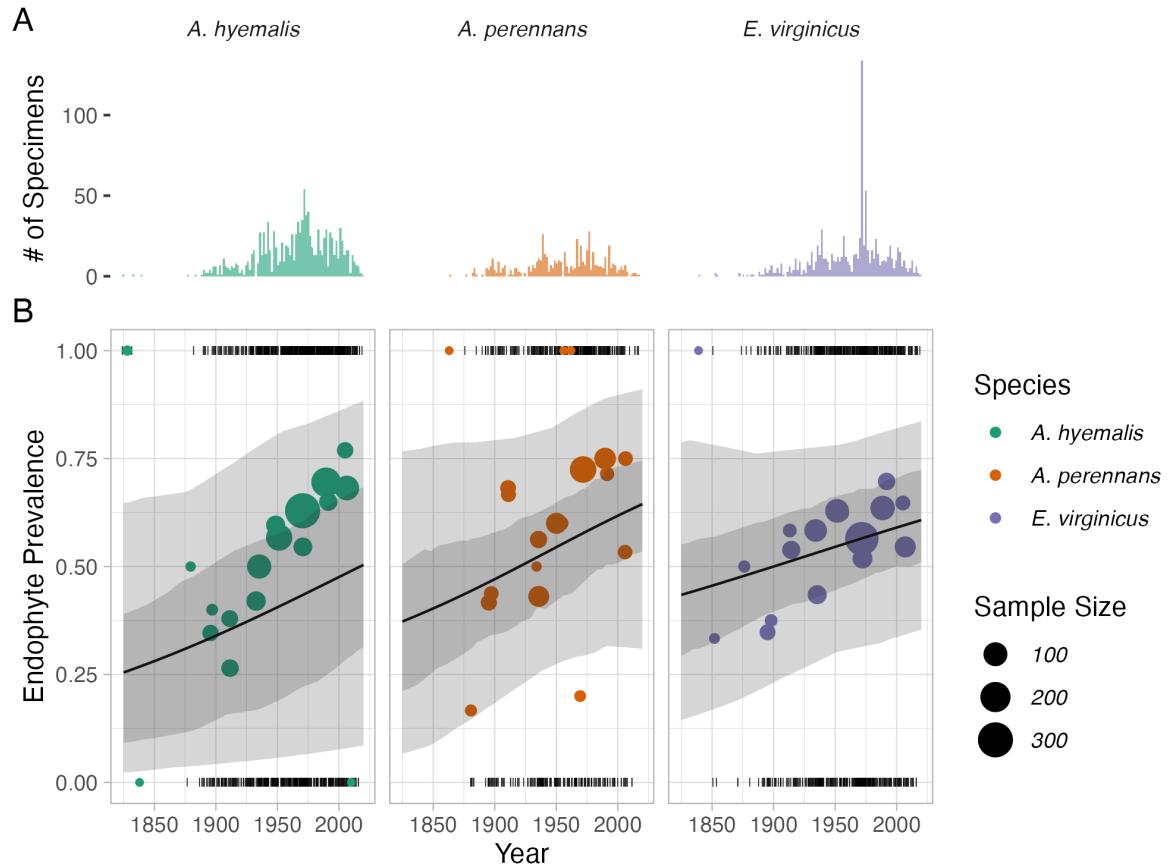


Figure 2: **Temporal trends in endophyte prevalence.** (A) Histograms show the frequency of scored specimens through time for each host species. (B) Lines show predicted mean endophyte prevalence over the study period along with the 50% and 95% CI bands incorporating uncertainty associated with collector and scorer random effects. Trends are estimated from the endophyte prevalence model. Colored points are binned means of the observed endophyte presence/absence data (black dashes). Colors represent each host species (*A. hyemalis*: green, *A. perennans*: orange, *E. virginicus*: purple) and point size represents the number of specimens.

<sup>363</sup> The model appears to under-predict the observed increase in endophyte prevalence relative  
<sup>364</sup> to the data, particularly for *A. hyemalis* (Fig. 2B), but the model is accounting for random effects

365 and spatial non-independence that are not readily seen in the figure. We found no evidence  
366 that collector biases influenced our results. Collector random effects were consistently small (Fig.  
367 A9), and models fit with and without this random effect provide qualitatively similar results. The  
368 identity of individual scorers, **the researchers who identified endophyte status microscopically**,  
369 did contribute to observed patterns in endophyte prevalence. For example, 3 of the 25 scorers  
370 were more consistently likely than average to assign positive endophyte status, as indicated by  
371 95% credible intervals greater than zero (Fig. A10).

### 372 *How spatially variable are temporal trends in endophyte prevalence?*

373 While there was an overall increase in endophyte prevalence, our model revealed hotspots and  
374 coldspots of change across the host species' ranges, which are mapped in Fig. 3 across geo-  
375 graphic ranges predicted by MaxEnt species distribution models. In some regions, posterior  
376 mean estimates of spatially varying temporal trends indicate that *A. hyemalis* and *A. perennans*  
377 experienced increases in prevalence by as much as 2% per year over the study period, while  
378 *E. virginicus* experienced increases up to around 1% per year. Both *Agrostis* species show areas  
379 of strong increase and areas of declining prevalence, while *E. virginicus* had an overall weaker  
380 and geographically more homogeneous increase in endophyte prevalence. Notably, endophytes  
381 increased most strongly towards the western range edge of *A. hyemalis* (Fig. 3A) and across the  
382 northeastern US for *A. perennans* (Fig. 3B). **Broad increases in prevalence on average, along with**  
383 **increases towards range edges that had low historic prevalence result in range expansions of the**  
384 **symbiosis for both *Agrostis* species (Fig. 4)**. Increases in prevalence were strongest in regions  
385 **with low historic prevalence for the *Agrostis* species (Fig. A11 A-B)**, but for *E. virginicus* trends  
386 **did not differ according to historic prevalence (A11 C)**. Posterior estimates of uncertainty in spa-  
387 **tially varying slopes indicate that these hotspots of change may have experienced increases of up**  
388 **to 5% per year while declines in prevalence may be as great as 4% per year for *A. hyemalis* and *A.***  
389 ***perennans*. For *E. virginicus*, uncertainty ranges between 3.5% increases and 2.5% decreases (Fig.**  
390 **A7)**.

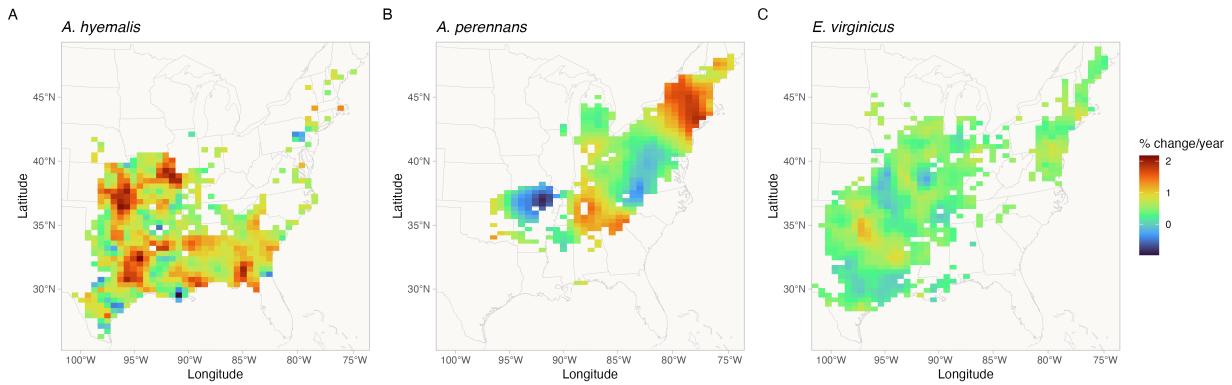


Figure 3: Predicted posterior mean of spatially-varying slopes representing change in endophyte prevalence for each host species. **Spatially-varying trends are estimated from the endophyte prevalence model.** Color indicates the relative change in predicted endophyte prevalence. Map lines delineate study areas and do not necessarily depict accepted national boundaries.

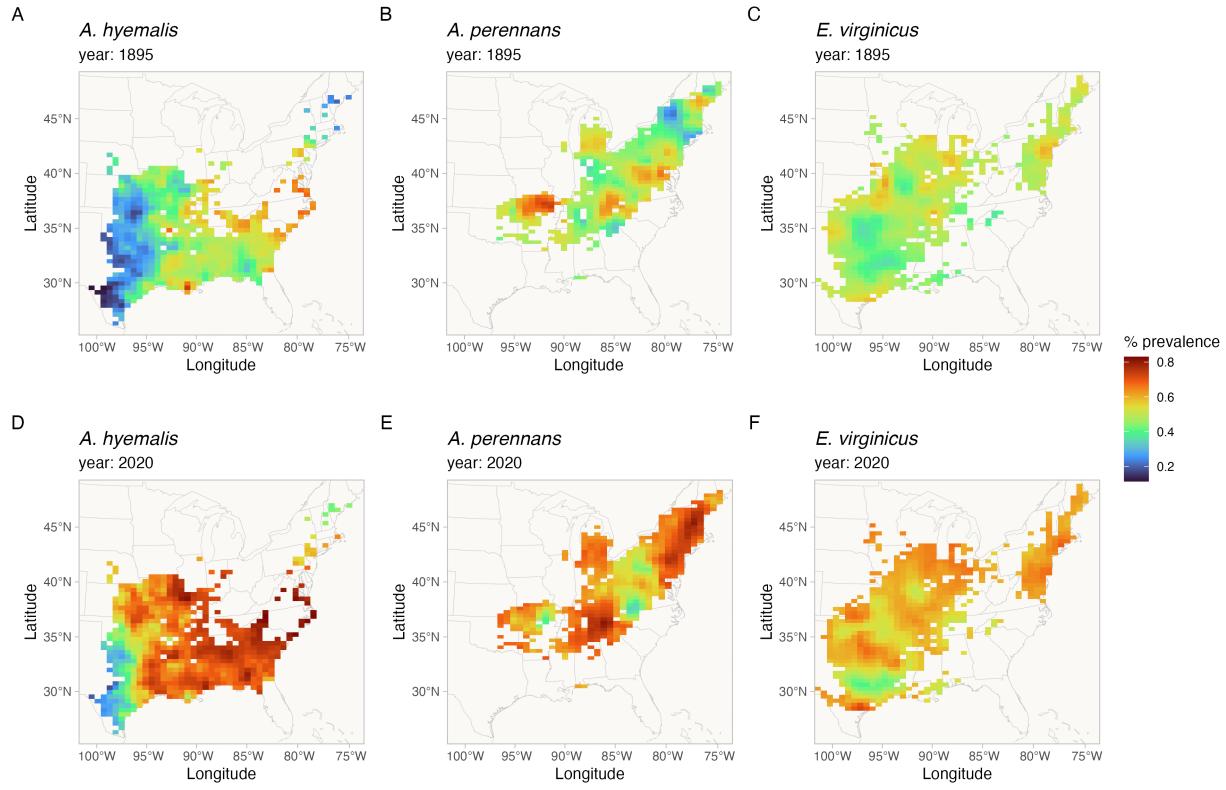


Figure 4: Predicted endophyte prevalence for each host species in 1895 and 2020. Predictions of prevalence come from the endophyte prevalence model. Color indicates the posterior mean endophyte prevalence for *A. hyemalis* in (A) 1895 and (D) 2020, for *A. perennans* in (B) 1895 and (E) 2020, and for *E. virginicus* in (C) 1895 and (F) 2020. Map lines delineate study areas and do not necessarily depict accepted national boundaries.

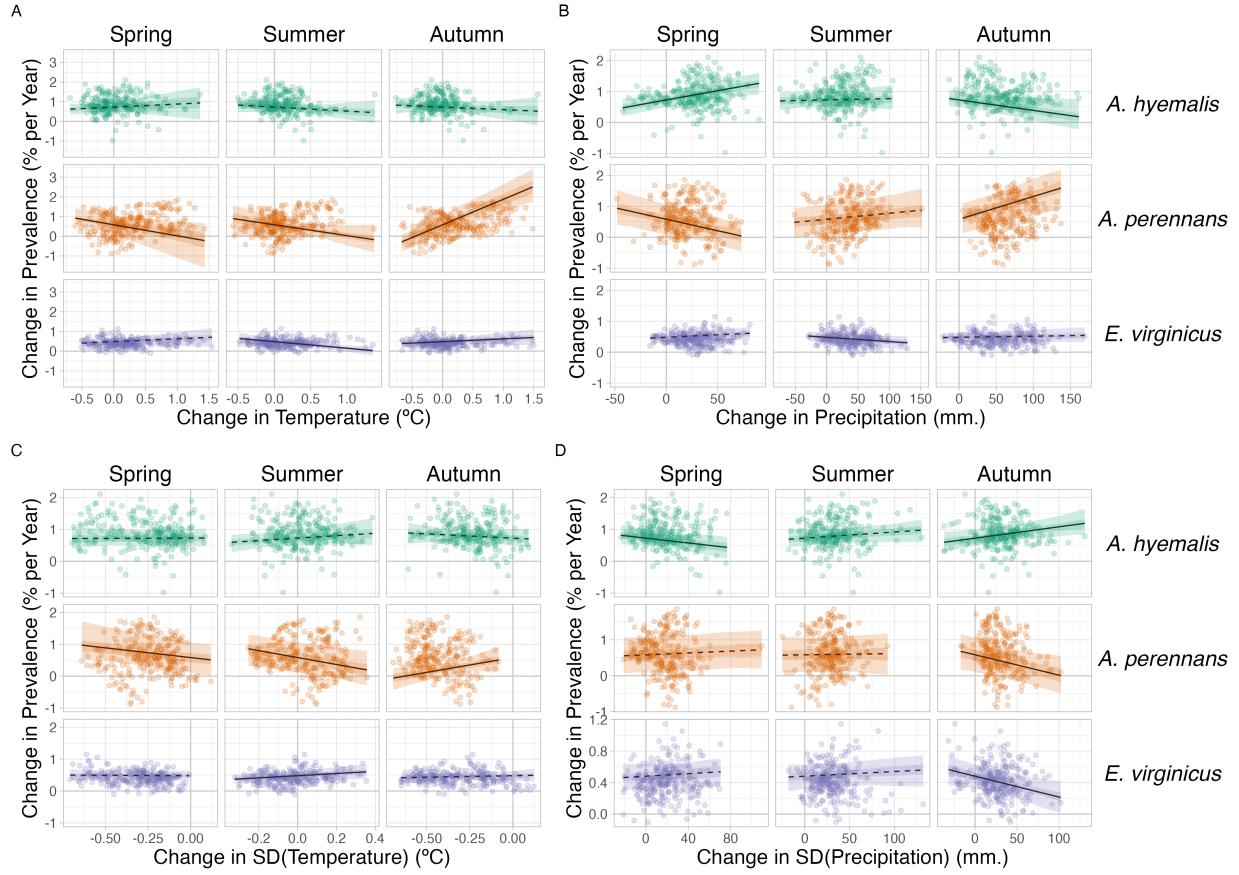
391     What is the relationship between variation in temporal trends in endophyte  
 392        prevalence and changes in climate drivers?

393     We found that trends in endophyte prevalence were strongly associated with seasonal climate  
 394        change drivers (Fig. 5). For the majority of the study region, the climate has become wetter (an  
 395        average increase in annual precipitation of 60 mm.) with relatively little temperature warming  
 396        (an average increase in annual temperature of 0.02 °C) over the last century (Fig. A13-A15), a con-

397 sequence of regional variation in global climate change (IPCC, 2021). Within the region, climate  
398 changes were spatially variable; certain locations experienced increases in annual precipitation  
399 as large as 375 mm. or decreases up to 54 mm. across the last century, while annual temper-  
400 ature changes ranged from warming as great as 1.4 °C to cooling by 0.46 °C. Spatially variable  
401 climate trends were predictive of trends in endophyte prevalence. For example, strong increases  
402 in endophyte prevalence for *A. perennans* were most strongly associated with increasing autumn  
403 precipitation and with increasing mean and variability in autumn temperature (greater than 97%  
404 posterior probabilities of positive slopes). For this species, a 1 °C increase in autumn temper-  
405 ature was associated with a 1.07 % increase per year in endophyte prevalence (Fig. 5A) and a  
406 100 mm. increase in precipitation was associated with a 0.8% increase per year in endophyte  
407 prevalence (Fig. 5B). This result aligns with the species' autumn active growing season, however  
408 other seasonal climate drivers were also associated with increasing endophyte prevalence. In  
409 particular, we found cooler and drier springs and cooler summers to be associated with increas-  
410 ing endophyte prevalence (greater than 99% posterior probabilities of negative slopes) however  
411 these slopes were generally of smaller magnitude than those for autumn climate drivers.

412 Changes in endophyte prevalence across the ranges of *A. hyemalis* and *E. virginicus* were less  
413 strongly driven by changes in climate. Like *A. perennans*, climate during peak growing season  
414 (spring for *A. perennans* and summer for *E. virginicus*) emerged most commonly as drivers of  
415 changes in endophyte prevalence. Increases in mean spring precipitation were the strongest pre-  
416 dictor of increasing trends in endophyte prevalence for *A. hyemalis* (Fig. 5B) (greater than 99%  
417 posterior probability of a positive slope). For this species, an increase of 100 mm. in spring pre-  
418 cipitation led to an increase of 0.6% per year in endophyte prevalence. The next greatest slopes  
419 were those associated with variability in spring precipitation (greater than 96% posterior proba-  
420 bility of a negative slope), as well as in the mean and variability of autumn climate (greater than  
421 98% probability of negative and positive slopes, respectively). Changes in endophyte prevalence  
422 in *E. virginicus* were not strongly associated with changes in most climate drivers, but regions  
423 with reduced variability in autumn precipitation (Fig. 5B) and with cooler and more variable

<sup>424</sup> summer temperatures (Fig. 5A,C) experienced the largest increases in endophyte prevalence.  
<sup>425</sup> Our analysis indicated relatively high confidence that these climate drivers influence endophyte  
<sup>426</sup> prevalence shifts in *E. virginicus*(greater than 99% posterior probability of either negative or posi-  
<sup>427</sup> tive slopes respectively), however they translate to less than 0.2% change in endophyte prevalence  
<sup>428</sup> per year for a change of 100 mm. change in precipitation over the century. Repeating this analysis  
<sup>429</sup> using all pixels across each species' distribution were qualitatively similar to these results.

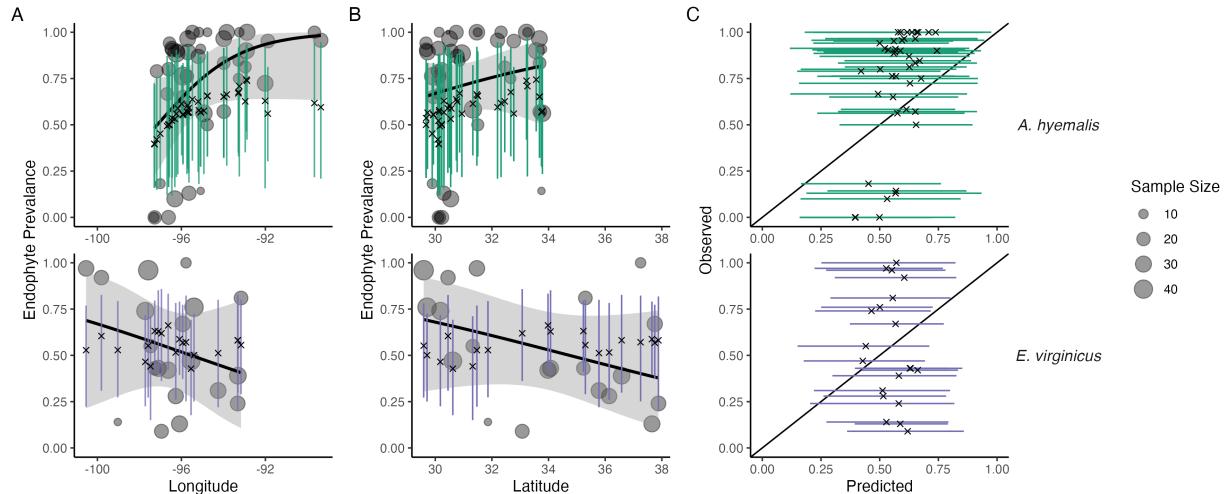


**Figure 5: Relationships between predicted trends in endophyte prevalence and changes in seasonal climate drivers.** Lines show marginal predicted relationship between spatially-varying trends in endophyte prevalence and changes in mean and variability of climate ((A): mean temperature, (B): cumulative precipitation, (C): standard deviation in temperature, (D): standard deviation in precipitation) **estimated from the *post hoc* climate regression analysis**. Confidence bands represent the 50 and 95% CI, colored by host species (*A. hyemalis*: green, *A. perennans*: orange, *E. virginicus*: purple). Slopes with greater than 95% probability of being either positive or negative are represented as solid lines while those that have less than 95% probability are dashed. Points are the **values of pre-computed SVC trends and climate drivers at 250 randomly sampled pixels across each host's distribution used in model fitting for the *post hoc* climate regression analysis**.

430

## *Evaluation of model performance on an out-of-sample test*

431 Tests of models' predictive performance as quantified by AUC and by visual posterior predic-  
432 tive checks, indicated good predictive ability. Model performance was similar between historic  
433 herbarium specimens used as training data and the out-of-sample test data from contemporary  
434 surveys (AUC = 0.79 and 0.77 respectively; Fig. A5-A4). The model successfully captured broad  
435 regional trends in endophyte prevalence seen in the contemporary survey data, such as decline  
436 endophyte prevalence in *A. hyemalis* towards western longitudes (Fig. 6A) and northern latitudes  
437 (Fig. 6B). **It is notable that** model predictions for endophyte prevalence exhibited relatively little  
438 local geographic variation, whereas the out-of-sample survey data were maximally variable with  
439 populations spanning 0% to 100% endophyte-symbiotic plants (Fig. 6C).



**Figure 6: Predictive performance for contemporary test data.** (A) The endophyte prevalence model, trained on historic herbarium collection data, performed modestly at predicting prevalence in contemporary population surveys. The model captured regional trends across (A) longitude and (B) latitude. Crosses indicate predicted mean prevalence along with the 95% CI (colored lines: *A. hyemalis*: green, orange, *E. virginicus*: purple) from the herbarium model. Contemporary prevalence is represented by grey points (point size reflects sample size) along with trend lines from generalized linear models (black line and shaded 95% confidence interval). (C) Comparison of observed vs. predicted endophyte prevalence shows that contemporary test data had more variance between populations than contemporary predictions.

440

## Discussion

441 Our examination of historic plant specimens revealed previously hidden shifts in microbial sym-  
 442 biosis over the last two centuries. For the three host species we examined, there have been strong  
 443 increases in prevalence of *Epichloë* endophyte symbiosis. We interpret increases in prevalence of  
 444 *Epichloë*, which are vertically transmitted, as adaptive changes that improve the fitness of their  
 445 hosts under increasing environmental stress. This interpretation is in line with theory predicting  
 446 that the positive fitness feedback caused by vertical transmission leads beneficial symbionts to

447 rise in prevalence within a population (Donald et al., 2021; Fine, 1975). We further found that  
448 trends in endophyte prevalence varied across the distribution of each species in association with  
449 changes in climate drivers, suggesting that the increases in endophyte prevalence are driven by  
450 context-dependent benefits to hosts that confer resilience under environmental change. Taken  
451 together, this suggests an overall strengthening of host-symbiont mutualism over the last two  
452 centuries.

453 *Responses of host-microbe symbioses to climate change*

454 Differences across host species underscore that while all of these  $C_3$  grasses share similar broad-  
455 scale distributions, each engages in unique biotic interactions and has unique responses to en-  
456 vironmental drivers. We identified hotspots of change for *A. perennans*, which was the species  
457 that experienced the strongest absolute changes in endophyte prevalence (Fig. 3). Declines of  
458 0.9% per year in the southern portion of its range and increases of up to 2% per year in the  
459 north suggest a potential poleward range shift of endophyte-symbiotic plants (whether the over-  
460 all host distribution is shifting in parallel is an exciting next question.) Based on previous work  
461 demonstrating that endophytes can shield their hosts from drought stress (reviewed in Decunta  
462 et al. (2021)), we generally predicted that drought conditions would be a driver of increasing en-  
463 dophyte prevalence. In contrast to this expectation, increasing prevalence for *A. perennans* were  
464 associated with increasing autumn temperature and precipitation (Fig. 5). To our knowledge,  
465 the response of the symbiosis in *A. perennans* to drought has not been examined experimentally,  
466 but in a greenhouse experiment, endophytes had a positive effect on host reproduction under  
467 shaded, low-light conditions (Davitt et al., 2010). Our results also hint that it may be useful to  
468 investigate whether lagged climate effects are important predictors of host fitness in this system  
469 (Evers et al., 2021). Endophyte prevalence of the autumn-flowering *A. perennans* was strongly  
470 linked with decreasing spring precipitation, and that of the spring-flowering *A. hyemalis* was  
471 associated with decreasing autumn precipitation (Fig. 5B). For *A. hyemalis*, endophytes could  
472 be playing a role helping hosts weather autumn-season droughts, which is likely also an im-

473 portant time for the species' germination. Previous work has demonstrated drought benefits in  
474 a greenhouse manipulation with this species (Davitt et al., 2011), and early life stages may be  
475 particularly vulnerable to prolonged droughts. For *E. virginicus*, which experienced the most  
476 modest changes in endophyte prevalence overall (ranging between 1.1% increases and 0.2% de-  
477 creases), we only found modest associations with changes in climate drivers. Surveys by Sneck  
478 et al. (2017), used as part of the test data in this study, identified a drought index (SPEI) that  
479 integrates precipitation with estimated evapotranspiration as an important predictor of **contem-**  
480 **porary** endophyte prevalence **in this species**. The diverse relationships we detect between trends  
481 in endophyte prevalence and climate drivers suggest a more complicated picture than the simple  
482 explanation that drought alone, as measured through changes in annual precipitation, causes  
483 increasing endophyte prevalence through context-dependent fitness benefits.

484 While we show consistent increasing trends in prevalence between the three species, the  
485 mechanisms that explain these changes may be diverse and idiosyncratic. First, climate change  
486 responses may depend on genotype-specific responses that are ignored in our current analy-  
487 sis. While *Epichloë* symbioses are highly specialized, surveys have demonstrated genotypic and  
488 chemotypic diversity of the symbionts across populations and even within populations (Treindl  
489 et al., 2023; von Cräutlein et al., 2021). Genotypic variation in *Epichloë* endophytes, particularly  
490 in genes responsible for alkaloid production, produces "chemotypes" with differing benefits for  
491 hosts against insect or mammalian herbivores mediated by environmental conditions (Saikkonen  
492 et al., 2013; Schardl et al., 2012). Genotypic variation of the hosts themselves can also influ-  
493 ence interaction outcomes (Gundel et al., 2011; Parker et al., 2017). Whether hotspots of change  
494 in endophyte prevalence reflect selection for genotype-pairings with particularly strong fitness  
495 benefits is an unanswered question. Additionally, *Epichloë* endophytes have been connected to  
496 a suite of non-drought related fitness benefits including herbivory **defense** (Brem and Leucht-  
497 mann, 2001), salinity resistance (Wang et al., 2020), and mediation of pathogens (Vikuk et al.,  
498 2019) and the soil microbiome (Roberts and Ferraro, 2015). **Broad changes in the distribution**  
499 **and abundance of natural enemies, along with stresses from anthropogenic changes in landcover**

500 and pollution (Sage, 2020) likely influence the benefits of symbiosis (Rudgers et al., 2020). The  
501 historic trends that we observed result from the combination of these fitness benefits playing out  
502 across the heterogenous map of a changing climate.

503 Our results indicate that *Epichloë* symbiosis has likely improved host fitness in stressful en-  
504 vironments leading to increasing prevalence. What is less clear is how this will influence future  
505 range shifts. Based on our analysis, it is likely that the symbiosis will facilitate range shifts for  
506 hosts by improving population growth at range edges. Previous population surveys attributed  
507 environment-dependent gradients in endophyte prevalence (Rudgers and Swafford, 2009; Sem-  
508 martin et al., 2015; Sneck et al., 2017) to symbiont-derived fitness benefits allowing their hosts  
509 to persist in environments where they otherwise could not (Afkhami et al., 2014; Kazenel et al.,  
510 2015). However, symbiont-facilitated range shifts require that endophytes be present in the pop-  
511 ulations to be able to support population growth. The arid western range edge of *A. hyemalis*  
512 has had historically low endophyte prevalence (Fig. 4), and while prevalence has increased most  
513 quickly in the regions with historically low endophyte prevalence (Fig. A11), the complete ab-  
514 sence of endophytes range edges would make it impossible for prevalence to increase without  
515 dispersal of symbiotic seeds (Fowler et al., 2023). These factors potentially contribute to the abil-  
516 ity of the host species to track its environmental niche. Another interesting question is the degree  
517 to which symbiotic and non-symbiotic hosts, which occupy overlapping but distinct niches, are  
518 likely to experience distribution shifts in tandem or at different spread rates in future work. More  
519 extreme climate stresses, which are expected more frequently in the future (Seneviratne et al.,  
520 2021), have the potential to alter the costs and benefits of the interaction. The past indicates a  
521 resilient interaction, but understanding the future climate conditions that may tip this interaction  
522 to deteriorate will be crucial.

523 *Steps towards forecasts of host-microbe symbioses*

524 The combination of a spatially-explicit model and historic herbarium specimens allowed us to  
525 identify regions of both increasing and decreasing endophyte prevalence, however we see sev-

526 eral next steps towards the goal of predicting host and symbiont niche-shifts in response to  
527 future climate change. While the model recreated the large-scale spatial trends observed in  
528 contemporary population surveys, test data contained more population-to-population variabil-  
529 ity in prevalence. We interpret this to mean that the model captures coarse-scale spatial and  
530 temporal trends reasonably well, but is not equipped to capture local-scale nuances that gener-  
531 ate population-to-population differences. Validating our model predictions with this test, a rare  
532 extra step in collections-based studies, allows us to evaluate places to improve the model's out-  
533 of-sample predictive ability. Lack of information on local variability in symbiont prevalence may  
534 simply be a feature of data derived from herbarium specimens. Natural history collectors sample  
535 one or a few specimens from local populations, which our analyses aggregates across to derive  
536 broad-scale model estimates. This suggests that increasing local replication should be a factor  
537 considered in future collection efforts of natural history specimens, balanced with the required  
538 time and effort and with limitations on storage space within collections. Poor predictive ability  
539 at local scales in this grass-endophyte system is not surprising, as previous studies have found  
540 that local variation, even to the scale of hundreds of meters can structure endophyte-host niches  
541 (Kazenel et al., 2015). An important step would be integrating data from local and regional scales  
542 through modeling to constrain estimates of local and regional variation.

543 Predicting future niche-shifts of hosts and symbionts will require considering the coupled  
544 dynamics of host-symbiont dispersal in addition to fitness benefits. For example, transplanting  
545 symbiotic and non-symbiotic plants beyond the range edge of *A. hyemalis* could tell us whether  
546 low endophyte prevalence in that area (Fig. 4A) is a result of environmental conditions that lead  
547 the symbiosis to negative fitness consequences, or is a result of some historical contingency or  
548 dispersal limitation that has thus far limited the presence of symbiotic hosts from a region where  
549 they would otherwise flourish and provide resilience. Incorporating available climatic and soil  
550 layers as covariates is another obvious step that could improve predictions. These steps will  
551 bridge gaps that often exist between large but broad bioclimatic and biodiversity data and small  
552 but local data on biotic interactions, and move towards the goal of predicting the dynamics of

553 microbial symbioses under climate change (Isaac et al., 2020; Miller et al., 2019).

554 *Herbaria for global change research*

555 Our analysis advances the use of herbarium specimens in global change biology in two ways.

556 First and foremost, this is one of a growing number of studies to examine microbial symbiosis us-  
557 ing specimens from natural history collections, and the first, to our knowledge, to link long-term  
558 changes in the symbioses to changes in climate. The responses of microbial symbioses are a rich  
559 target for future studies within historic specimens, particularly those that take advantage of ad-  
560 vances in sequencing technology. While we used relatively coarse presence/absence data based  
561 on fungal morphology, other studies have examined historic plant microbiomes using molecu-  
562 lar sequencing and sophisticated bioinformatics techniques, but these studies have so far been  
563 limited to relatively few specimens at limited spatial extents (Bieker et al., 2020; Bradshaw et al.,  
564 2021; Gross et al., 2021; Heberling and Burke, 2019; Yoshida et al., 2015). Continued advances  
565 in capturing historic DNA and in filtering out potential contamination during specimen storage  
566 (Bakker et al., 2020; Daru et al., 2019; Raxworthy and Smith, 2021) will be imperative in the effort  
567 to scale up these efforts. This scaling up will be essential to be able to quantify changes not just  
568 in the prevalence of symbionts, but also in symbionts' intraspecific variation and evolutionary  
569 responses to climate change, as well as in changes in the wider host microbiome. With improved  
570 molecular insights from historic specimens, we could ask whether the broad increases in en-  
571 dophytes that we have identified reflect selection for particular genetic strains or chemotypes  
572 and how this selection varies across space. Answering these questions as well as the unknown  
573 questions that future researchers may ask also reiterates the value in capturing meta-information  
574 during ongoing digitization efforts at herbaria around the world and during the accession of  
575 newly collected specimens (Edwards et al.; Lendemer et al., 2020).

576 The second major advance in this analysis is in accounting for several potential biases in  
577 the data observation process that may be common to many collections-based research questions  
578 by using a spatially-explicit random effects model. Potential biases introduced by the sampling

habits of collectors (Daru et al., 2018), and variation between contemporary researchers during the collection of trait data, if not corrected for could lead to over-confident inference about the strength and direction of historic change (Fig. 2). Previous studies that have quantified the effects of collector biases typically find them to be small (Davis et al., 2015; Meineke et al., 2019), and we similarly did not find that collector has a strong effect on the results of our analysis, but that scorer identity did impact results. It is difficult to distinguish whether the impact of scorers was driven by true differences in scorers' biases during the seed scoring process or by unintended spatial or temporal clustering of the specimens examined by each scorer (Clayton et al., 1993; Urdangarin et al., 2023). By under-weighting endophyte-positive samples that are clustered spatially or by collector or observer, the endophyte prevalence model is appropriately accounting for nuisance variables and providing a conservative inference of endophyte change relative to the raw data. Spatial autocorrelation is another phenomenon<sup>3</sup> likely common in data derived from herbarium specimens (Willems et al., 2022), which our spatially-explicit analysis models among samples. Beyond spatial autocorrelation of outcomes, systematic differences in sampling across space can result in spatial bias. One strength of herbaria as vehicles for global change research is the relative ease with which specimens from many distinct geographic locations can be examined. We visited just nine institutions in the central southern United States, and we were able to sample seeds from specimens across an area spanning over 300,000 sq. km, including specimens from Mexico and Canada. The specimens we examined are concentrated in the South-Central United States, with fewer specimens in the rapidly warming northeastern United States. We provide a simulation analysis exploring the potential impact of spatially and temporally biased sampling (Appendix A - Supporting Methods). We found that the spatially-varying coefficient model had a strong ability to re-capitulate temporal trends across space in simulated data, and that this result was robust to relatively high levels of spatial bias (80% of data missing from region). Simulation analyses that extend this work to consider the myriad ways herbarium data may be biased (i.e. testing different spatial arrangements and scales of spatial bias, or testing

---

<sup>3</sup>phenomenon ?

605 different sample sizes) would be extremely valuable (Daru et al., 2018; Erickson and Smith, 2021;  
606 Gaul et al., 2020; Meineke and Daru, 2021).

607 *Conclusion*

608 Ultimately, a central goal of global change biology is to generate predictive insights into the  
609 future of natural systems on a rapidly changing planet. Beyond host-microbe symbioses, de-  
610 tecting ecological responses to anthropogenic global change and attributing their causes would  
611 inform public policy decision-makers and adaptive management strategies. This survey of his-  
612 toric endophyte prevalence is necessarily correlative, yet it serves as a foundation to develop  
613 better predictive models of the response of microbial symbioses to climate change. By compar-  
614 ing detected ecological responses with alternative mechanistic simulations of the past, we could  
615 attribute their cause, in a manner similar to methods from climate science and economics (Car-  
616 leton and Hsiang, 2016; Stott et al., 2010; Trenberth et al., 2015). Combining the insights from  
617 this type of regional-scale survey with field experiments and physiological performance data  
618 could be invaluable to identify mechanisms driving shifts in host-symbiont dynamics. Evidence  
619 is strong that certain dimensions of climate change correlated with endophytes' temporal re-  
620 sponds, however we do not know why trends in prevalence were weak in some areas or how  
621 endophytes would respond to more extreme changes in climate. The "time machine" of natu-  
622 ral history collections revealed evidence of mutualism resilience for grass-endophyte symbioses  
623 in the face of environmental change, but more extreme changes could potentially push one or  
624 both partners beyond their physiological limits, leading to the collapse of the mutualism; more  
625 research is needed to understand what those limits might be.

626 **Acknowledgments**

627 We thank Dr. Jessica Budke for help in drafting our initial destructive sampling plan, and to the  
628 many staff members of herbaria who facilitated our research visits, as well as to the hundreds  
629 of collectors who contributed to the natural history collections. Several high schooler and un-  
630 dergraduate researchers contributed to data collection, including A. Appio-Riley, P. Bilderback,

631 E. Chong, K. Dickens, L. Dufresne, B. Gutierrez, A. Johnson, S. Linder, E. Scales, B. Scherick,  
632 K. Schrader, E. Segal , G. Singla, and M. Tucker. This research was supported by funding from  
633 National Science Foundation (grants 1754468 and 2208857) and by funding from the Texas Ecolab  
634 Program.

635 **Statement of Authorship**

636 J.C.F. contributed to research conception, data collection, data analysis, and led manuscript draft-  
637 ing. J.M. contributed to data analysis and manuscript revisions. T.E.X.M. contributed to research  
638 conception, data collection, data analysis, and manuscript revisions.

639 **Data and Code Availability**

640 Data from this publication can be found through a publicly available repository  
641 (<https://doi.org/10.5061/dryad.rn8pk0pn0>). Code for analyses can be found through a pub-  
642 licly available repository (<https://github.com/joshuacfowler/EndoHerbarium>) that will be per-  
643 manently archived upon publication.

644

## Literature Cited

- 645 Michelle E Afkhami. Fungal endophyte–grass symbioses are rare in the California floristic  
646 province and other regions with mediterranean-influenced climates. *Fungal ecology*, 5(3):345–  
647 352, 2012.
- 648 Michelle E Afkhami and Jennifer A Rudgers. Symbiosis lost: imperfect vertical transmission of  
649 fungal endophytes in grasses. *The American Naturalist*, 172(3):405–416, 2008.
- 650 Michelle E Afkhami, Patrick J McIntyre, and Sharon Y Strauss. Mutualist-mediated effects on  
651 species' range limits across large geographic scales. *Ecology letters*, 17(10):1265–1273, 2014.
- 652 Sally N Aitken, Sam Yeaman, Jason A Holliday, Tongli Wang, and Sierra Curtis-McLane. Adap-  
653 tation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary*  
654 *applications*, 1(1):95–111, 2008.
- 655 Clare E Aslan, Erika S Zavaleta, Bernie Tershy, and Donald Croll. Mutualism disruption threatens  
656 global plant biodiversity: a systematic review. *PLoS one*, 8(6):e66993, 2013.
- 657 Fabian E Bachl, Finn Lindgren, David L Borchers, and Janine B Illian. inlabru: an R package for  
658 bayesian spatial modelling from ecological survey data. *Methods in Ecology and Evolution*, 10(6):  
659 760–766, 2019.
- 660 Charles W Bacon and James F White. Stains, media, and procedures for analyzing endophytes.  
661 In *Biotechnology of endophytic fungi of grasses*, pages 47–56. CRC Press, 2018.
- 662 Haakon Bakka, Håvard Rue, Geir-Arne Fuglstad, Andrea Riebler, David Bolin, Janine Illian, Elias  
663 Krainski, Daniel Simpson, and Finn Lindgren. Spatial modeling with r-inla: A review. *Wiley*  
664 *Interdisciplinary Reviews: Computational Statistics*, 10(6):e1443, 2018.
- 665 Freek T Bakker, Vanessa C Bieker, and Michael D Martin. Herbarium collection-based plant  
666 evolutionary genetics and genomics, 2020.

- 667 Dawn R Bazely, John P Ball, Mark Vicari, Andrew J Tanentzap, Myrtille Bérenger, Tomo Rakoc  
668 evic, and Saewan Koh. Broad-scale geographic patterns in the distribution of vertically-  
669 transmitted, asexual endophytes in four naturally-occurring grasses in sweden. *Ecography*,  
670 30(3):367–374, 2007.
- 671 Julien Beguin, Sara Martino, Håvard Rue, and Steven G Cumming. Hierarchical analysis of  
672 spatially autocorrelated ecological data using integrated nested laplace approximation. *Methods*  
673 in *Ecology and Evolution*, 3(5):921–929, 2012.
- 674 Colette S Berg, Jason L Brown, and Jennifer J Weber. An examination of climate-driven flowering-  
675 time shifts at large spatial scales over 153 years in a common weedy annual. *American Journal*  
676 of *Botany*, 106(11):1435–1443, 2019.
- 677 Vanessa C Bieker, Fátima Sánchez Barreiro, Jacob A Rasmussen, Marie Brunier, Nathan Wales,  
678 and Michael D Martin. Metagenomic analysis of historical herbarium specimens reveals a  
679 postmortem microbial community. *Molecular ecology resources*, 20(5):1206–1219, 2020.
- 680 Jessica L Blois, Phoebe L Zarnetske, Matthew C Fitzpatrick, and Seth Finnegan. Climate change  
681 and the past, present, and future of biotic interactions. *Science*, 341(6145):499–504, 2013.
- 682 Michael Bradshaw, Uwe Braun, Marianne Elliott, Julia Kruse, Shu-Yan Liu, Guanxiu Guan, and  
683 Patrick Tobin. A global genetic analysis of herbarium specimens reveals the invasion dynamics  
684 of an introduced plant pathogen. *Fungal Biology*, 125(8):585–595, 2021.
- 685 D Brem and A Leuchtmann. Epichloë grass endophytes increase herbivore resistance in the  
686 woodland grass *brachypodium sylvaticum*. *Oecologia*, 126(4):522–530, 2001.
- 687 Tamara A Carleton and Solomon M Hsiang. Social and economic impacts of climate. *Science*, 353  
688 (6304):aad9837, 2016.
- 689 Shen Cheng, Ying-Ning Zou, Kamil Kuča, Abeer Hashem, Elsayed Fathi Abd\_Allah, and Qiang-

- 690 Sheng Wu. Elucidating the mechanisms underlying enhanced drought tolerance in plants  
691 mediated by arbuscular mycorrhizal fungi. *Frontiers in Microbiology*, 12:4029, 2021.
- 692 Keith Clay and Christopher Schardl. Evolutionary origins and ecological consequences of endo-  
693 phyte symbiosis with grasses. *the american naturalist*, 160(S4):S99–S127, 2002.
- 694 David G Clayton, Luisa Bernardinelli, and Cristina Montomoli. Spatial correlation in ecological  
695 analysis. *International journal of epidemiology*, 22(6):1193–1202, 1993.
- 696 KD Craven, PTW Hsiau, A Leuchtmann, W Hollin, and CL Schardl. Multigene phylogeny of  
697 epichloë species, fungal symbionts of grasses. *Annals of the Missouri Botanical Garden*, pages  
698 14–34, 2001.
- 699 Kerri M Crawford, John M Land, and Jennifer A Rudgers. Fungal endophytes of native grasses  
700 decrease insect herbivore preference and performance. *Oecologia*, 164:431–444, 2010.
- 701 Michael S Crossley, Timothy D Meehan, Matthew D Moran, Jeffrey Glassberg, William E Snyder,  
702 and Andrew K Davis. Opposing global change drivers counterbalance trends in breeding north  
703 american monarch butterflies. *Global change biology*, 28(15):4726–4735, 2022.
- 704 Christopher Daly and Kirk Bryant. The prism climate and weather system—an introduction.  
705 *Corvallis, OR: PRISM climate group*, 2, 2013.
- 706 Barnabas H Daru, Daniel S Park, Richard B Primack, Charles G Willis, David S Barrington,  
707 Timothy JS Whitfeld, Tristram G Seidler, Patrick W Sweeney, David R Foster, Aaron M Ellison,  
708 et al. Widespread sampling biases in herbaria revealed from large-scale digitization. *New  
709 Phytologist*, 217(2):939–955, 2018.
- 710 Barnabas H Daru, Elizabeth A Bowman, Donald H Pfister, and A Elizabeth Arnold. A novel proof  
711 of concept for capturing the diversity of endophytic fungi preserved in herbarium specimens.  
712 *Philosophical Transactions of the Royal Society B*, 374(1763):20170395, 2019.

- 713 Charles C Davis, Charles G Willis, Bryan Connolly, Courtland Kelly, and Aaron M Ellison.  
714 Herbarium records are reliable sources of phenological change driven by climate and pro-  
715 vide novel insights into species' phenological cueing mechanisms. *American journal of botany*,  
716 102(10):1599–1609, 2015.
- 717 Andrew J Davitt, Marcus Stansberry, and Jennifer A Rudgers. Do the costs and benefits of fungal  
718 endophyte symbiosis vary with light availability? *New Phytologist*, 188(3):824–834, 2010.
- 719 Andrew J Davitt, Chris Chen, and Jennifer A Rudgers. Understanding context-dependency in  
720 plant–microbe symbiosis: the influence of abiotic and biotic contexts on host fitness and the  
721 rate of symbiont transmission. *Environmental and Experimental Botany*, 71(2):137–145, 2011.
- 722 Facundo A Decunta, Luis I Pérez, Dariusz P Malinowski, Marco A Molina-Montenegro, and  
723 Pedro E Gundel. A systematic review on the effects of epichloë fungal endophytes on drought  
724 tolerance in cool-season grasses. *Frontiers in plant science*, 12:644731, 2021.
- 725 Mauro Di Luzio, Gregory L Johnson, Christopher Daly, Jon K Eischeid, and Jeffrey G Arnold.  
726 Constructing retrospective gridded daily precipitation and temperature datasets for the con-  
727 terminous united states. *Journal of Applied Meteorology and Climatology*, 47(2):475–497, 2008.
- 728 Marion L Donald, Teresa F Bohner, Kory M Kolis, R Alan Shadow, Jennifer A Rudgers, and  
729 Tom EX Miller. Context-dependent variability in the population prevalence and individual  
730 fitness effects of plant–fungal symbiosis. *Journal of Ecology*, 109(2):847–859, 2021.
- 731 AE Douglas. Host benefit and the evolution of specialization in symbiosis. *Heredity*, 81(6):599–  
732 603, 1998.
- 733 Yuan-Wen Duan, Haibao Ren, Tao Li, Lin-Lin Wang, Zhi-Qiang Zhang, Yan-Li Tu, and Yong-Ping  
734 Yang. A century of pollination success revealed by herbarium specimens of seed pods. *New  
735 Phytologist*, 224(4):1512–1517, 2019.

- 736 Erika J Edwards, Brent D Mishler, and Charles D Davis. University herbaria are uniquely impor-  
737 tant. *Trends in Plant Science*.
- 738 Markus Engel, Tobias Mette, and Wolfgang Falk. Spatial species distribution models: Using  
739 bayes inference with inla and spde to improve the tree species choice for important european  
740 tree species. *Forest Ecology and Management*, 507:119983, 2022.
- 741 Kelley D Erickson and Adam B Smith. Accounting for imperfect detection in data from muse-  
742 ums and herbaria when modeling species distributions: combining and contrasting data-level  
743 versus model-level bias correction. *Ecography*, 44(9):1341–1352, 2021.
- 744 Sanne M Evers, Tiffany M Knight, David W Inouye, Tom EX Miller, Roberto Salguero-Gómez,  
745 Amy M Iler, and Aldo Compagnoni. Lagged and dormant season climate better predict plant  
746 vital rates than climate during the growing season. *Global Change Biology*, 27(9):1927–1941,  
747 2021.
- 748 Paul EM Fine. Vectors and vertical transmission: an epidemiologic perspective. *Annals of the New  
749 York Academy of Sciences*, 266(1):173–194, 1975.
- 750 Joshua C Fowler, Marion L Donald, Judith L Bronstein, and Tom EX Miller. The geographic  
751 footprint of mutualism: How mutualists influence species' range limits. *Ecological Monographs*,  
752 93(1):e1558, 2023.
- 753 Joshua C Fowler, Shaun Ziegler, Kenneth D Whitney, Jennifer A Rudgers, and Tom EX Miller.  
754 Microbial symbionts buffer hosts from the demographic costs of environmental stochasticity.  
755 *Ecology Letters*, 27(5):e14438, 2024.
- 756 PR Frade, F De Jongh, F Vermeulen, J Van Bleijswijk, and RPM Bak. Variation in symbiont  
757 distribution between closely related coral species over large depth ranges. *Molecular Ecology*,  
758 17(2):691–703, 2008.

- 759 Megan E Frederickson. Mutualisms are not on the verge of breakdown. *Trends in ecology &*  
760 *evolution*, 32(10):727–734, 2017.
- 761 Willson Gaul, Dinara Sadykova, Hannah J White, Lupe Leon-Sanchez, Paul Caplat, Mark C  
762 Emmerson, and Jon M Yearsley. Data quantity is more important than its spatial bias for  
763 predictive species distribution modelling. *PeerJ*, 8:e10411, 2020.
- 764 Andrew Gelman and Jennifer Hill. *Data analysis using regression and multilevel/hierarchical models*.  
765 Cambridge university press, 2006.
- 766 Sarah E Gilman, Mark C Urban, Joshua Tewksbury, George W Gilchrist, and Robert D Holt. A  
767 framework for community interactions under climate change. *Trends in ecology & evolution*, 25  
768 (6):325–331, 2010.
- 769 Gustaf Granath, Mark Vicari, Dawn R Bazely, John P Ball, Adriana Puentes, and Tomo Rakoce-  
770 vic. Variation in the abundance of fungal endophytes in fescue grasses along altitudinal and  
771 grazing gradients. *Ecography*, 30(3):422–430, 2007.
- 772 Andrin Gross, Célia Petitcollin, Cyril Dutech, Bayo Ly, Marie Massot, Julie Faivre d’Arcier, Laure  
773 Dubois, Gilles Saint-Jean, and Marie-Laure Desprez-Loustau. Hidden invasion and niche con-  
774 traction revealed by herbaria specimens in the fungal complex causing oak powdery mildew  
775 in europe. *Biological Invasions*, 23:885–901, 2021.
- 776 Pedro Emilio Gundel, Iñigo Zabalgoitia, and BR Vázquez De Aldana. Interaction between  
777 plant genotype and the symbiosis with epichloë fungal endophytes in seeds of red fescue  
778 (*festuca rubra*). *Crop and Pasture Science*, 62(11):1010–1016, 2011.
- 779 Edmund M. Hart and Kendon Bell. prism: Download data from the oregon prism project. 2015.  
780 doi: 10.5281/zenodo.33663. URL <https://github.com/ropensci/prism>. R package version  
781 0.0.6.

- 782 J Mason Heberling and David J Burke. Utilizing herbarium specimens to quantify historical  
783 mycorrhizal communities. *Applications in plant sciences*, 7(4):e01223, 2019.
- 784 Robert J Hijmans, Steven Phillips, John Leathwick, Jane Elith, and Maintainer Robert J Hijmans.  
785 Package ‘dismo’. *Circles*, 9(1):1–68, 2017.
- 786 Janneke HilleRisLambers, Melanie A Harsch, Ailene K Ettinger, Kevin R Ford, and Elinore J  
787 Theobald. How will biotic interactions influence climate change-induced range shifts? *Annals*  
788 *of the New York Academy of Sciences*, 1297(1):112–125, 2013.
- 789 IPCC. Climate change 2021: The physical science basis, 2021. URL [https://www.ipcc.ch/  
790 report/ar6/wg1/](https://www.ipcc.ch/report/ar6/wg1/).
- 791 Nick JB Isaac, Marta A Jarzyna, Petr Keil, Lea I Dambly, Philipp H Boersch-Supan, Ella Browning,  
792 Stephen N Freeman, Nick Golding, Gurutzeta Guillera-Arroita, Peter A Henrys, et al. Data  
793 integration for large-scale models of species distributions. *Trends in ecology & evolution*, 35(1):  
794 56–67, 2020.
- 795 Alberto Jiménez-Valverde. Insights into the area under the receiver operating characteristic curve  
796 (auc) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeog-  
797 raphy*, 21(4):498–507, 2012.
- 798 David Kahle, Hadley Wickham, and Maintainer David Kahle. Package ‘ggmap’. *Retrieved Septem-  
799 ber*, 5:2021, 2019.
- 800 Melanie R Kazenel, Catherine L Debban, Luciana Ranelli, Will Q Hendricks, Y Anny Chung,  
801 Thomas H Pendergast IV, Nikki D Charlton, Carolyn A Young, and Jennifer A Rudgers. A  
802 mutualistic endophyte alters the niche dimensions of its host plant. *AoB plants*, 7:plv005, 2015.
- 803 Roland A Knapp, Gary M Fellers, Patrick M Kleeman, David AW Miller, Vance T Vredenburg,  
804 Erica Bree Rosenblum, and Cheryl J Briggs. Large-scale recovery of an endangered amphibian

- despite ongoing exposure to multiple stressors. *Proceedings of the National Academy of Sciences*, 113(42):11889–11894, 2016.
- Mikhail V Kozlov, Irina V Sokolova, Vitali Zverev, Alexander A Egorov, Mikhail Y Goncharov, and Elena L Zvereva. Biases in estimation of insect herbivory from herbarium specimens. *Scientific Reports*, 10(1):12298, 2020.
- Benjamin R Lee, Evelyn F Alecrim, Tara K Miller, Jessica RK Forrest, J Mason Heberling, Richard B Primack, and Risa D Sargent. Phenological mismatch between trees and wildflowers: Reconciling divergent findings in two recent analyses. *Journal of Ecology*, 112(6):1184–1199, 2024.
- James Lendemer, Barbara Thiers, Anna K Monfils, Jennifer Zaspel, Elizabeth R Ellwood, Andrew Bentley, Katherine LeVan, John Bates, David Jennings, Dori Contreras, et al. The extended specimen network: A strategy to enhance us biodiversity collections, promote research and education. *BioScience*, 70(1):23–30, 2020.
- A Leuchtmann. Systematics, distribution, and host specificity of grass endophytes. *Natural Toxins*, 1(3):150–162, 1992.
- Adrian Leuchtmann, Charles W Bacon, Christopher L Schardl, James F White Jr, and Mariusz Tadych. Nomenclatural realignment of neotyphodium species with genus epichloë. *Mycologia*, 106(2):202–215, 2014.
- Finn Lindgren, Håvard Rue, and Johan Lindström. An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011.
- Canran Liu, Pam M Berry, Terence P Dawson, and Richard G Pearson. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28(3):385–393, 2005.
- Patricia D Mc Cargo, Leopoldo J Iannone, María Victoria Vignale, Christopher L Schardl, and

- 829 María Susana Rossi. Species diversity of epichloë symbiotic with two grasses from southern  
830 argentinean patagonia. *Mycologia*, 106(2):339–352, 2014.
- 831 Margaret McFall-Ngai, Michael G Hadfield, Thomas CG Bosch, Hannah V Carey, Tomislav  
832 Domazet-Lošo, Angela E Douglas, Nicole Dubilier, Gerard Eberl, Tadashi Fukami, Scott F  
833 Gilbert, et al. Animals in a bacterial world, a new imperative for the life sciences. *Proceedings*  
834 *of the National Academy of Sciences*, 110(9):3229–3236, 2013.
- 835 Timothy D Meehan, Nicole L Michel, and Håvard Rue. Spatial modeling of audubon christmas  
836 bird counts reveals fine-scale patterns and drivers of relative abundance trends. *Ecosphere*, 10  
837 (4):e02707, 2019.
- 838 Emily K Meineke and Barnabas H Daru. Bias assessments to expand research harnessing biological  
839 collections. *Trends in Ecology & Evolution*, 36(12):1071–1082, 2021.
- 840 Emily K Meineke, Charles C Davis, and T Jonathan Davies. The unrealized potential of herbaria  
841 for global change biology. *Ecological Monographs*, 88(4):505–525, 2018.
- 842 Emily K Meineke, Aimée T Classen, Nathan J Sanders, and T Jonathan Davies. Herbarium  
843 specimens reveal increasing herbivory over the past century. *Journal of Ecology*, 107(1):105–117,  
844 2019.
- 845 Abigail R Meyer, Maria Valentin, Laima Liulevicius, Tami R McDonald, Matthew P Nelsen, Jean  
846 Pengra, Robert J Smith, and Daniel Stanton. Climate warming causes photobiont degradation  
847 and c starvation in a boreal climate sentinel lichen. *American Journal of Botany*, 2022.
- 848 David AW Miller, Krishna Pacifici, Jamie S Sanderlin, and Brian J Reich. The recent past and  
849 promising future for data integration methods to estimate species' distributions. *Methods in  
850 Ecology and Evolution*, 10(1):22–37, 2019.
- 851 Daniel S Park, Ian Breckheimer, Alex C Williams, Edith Law, Aaron M Ellison, and Charles C  
852 Davis. Herbarium specimens reveal substantial and unexpected variation in phenological sen-

- 853 sitivity across the eastern united states. *Philosophical Transactions of the Royal Society B*, 374  
854 (1763):20170394, 2019.
- 855 Benjamin J Parker, Jan Hrček, Ailsa HC McLean, and H Charles J Godfray. Genotype specificity  
856 among hosts, pathogens, and beneficial microbes influences the strength of symbiont-mediated  
857 protection. *Evolution*, 71(5):1222–1231, 2017.
- 858 Martin Parniske. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nature Reviews  
859 Microbiology*, 6(10):763–775, 2008.
- 860 Anton Pauw and Julie A Hawkins. Reconstruction of historical pollination rates reveals linked  
861 declines of pollinators and plants. *Oikos*, 120(3):344–349, 2011.
- 862 Shilong Piao, Qiang Liu, Anping Chen, Ivan A Janssens, Yongshuo Fu, Junhu Dai, Lingli Liu,  
863 XU Lian, Miaogen Shen, and Xiaolin Zhu. Plant phenology and global climate change: Current  
864 progresses and challenges. *Global change biology*, 25(6):1922–1940, 2019.
- 865 Timothée Poisot, Gabriel Bergeron, Kevin Cazelles, Tad Dallas, Dominique Gravel, Andrew Mac-  
866 Donald, Benjamin Mercier, Clément Violet, and Steve Vissault. Global knowledge gaps in  
867 species interaction networks data. *Journal of Biogeography*, 48(7):1552–1563, 2021.
- 868 Nicole E Rafferty, Paul J CaraDonna, and Judith L Bronstein. Phenological shifts and the fate of  
869 mutualisms. *Oikos*, 124(1):14–21, 2015.
- 870 Christopher J Raxworthy and Brian Tilston Smith. Mining museums for historical dna: advances  
871 and challenges in museomics. *Trends in Ecology & Evolution*, 36(11):1049–1060, 2021.
- 872 François Renoz, Inès Pons, and Thierry Hance. Evolutionary responses of mutualistic insect–  
873 bacterial symbioses in a world of fluctuating temperatures. *Current opinion in insect science*, 35:  
874 20–26, 2019.
- 875 Elizabeth Lewis Roberts and Aileen Ferraro. Rhizosphere microbiome selection by epichloë en-  
876 dophytes of festuca arundinacea. *Plant and soil*, 396:229–239, 2015.

- 877 RJ Rodriguez, JF White Jr, Anne E Arnold, and a RS and Redman. Fungal endophytes: diversity  
878 and functional roles. *New phytologist*, 182(2):314–330, 2009.
- 879 Gregor Rolshausen, Francesco Dal Grande, Anna D Sadowska-Deś, Jürgen Otte, and Imke  
880 Schmitt. Quantifying the climatic niche of symbiont partners in a lichen symbiosis indicates  
881 mutualist-mediated niche expansions. *Ecography*, 41(8):1380–1392, 2018.
- 882 Jennifer A Rudgers and Angela L Swafford. Benefits of a fungal endophyte in *elymus virginicus*  
883 decline under drought stress. *Basic and Applied Ecology*, 10(1):43–51, 2009.
- 884 Jennifer A Rudgers, Michelle E Afkhami, Megan A Rúa, Andrew J Davitt, Samantha Hammer,  
885 and Valérie M Huguet. A fungus among us: broad patterns of endophyte distribution in the  
886 grasses. *Ecology*, 90(6):1531–1539, 2009.
- 887 Jennifer A Rudgers, Michelle E Afkhami, Lukas Bell-Dereske, Y Anny Chung, Kerri M Crawford,  
888 Stephanie N Kivlin, Michael A Mann, and Martin A Nuñez. Climate disruption of plant-  
889 microbe interactions. *Annual review of ecology, evolution, and systematics*, 51(1):561–586, 2020.
- 890 Håvard Rue, Sara Martino, and Nicolas Chopin. Approximate bayesian inference for latent gaus-  
891 sian models by using integrated nested laplace approximations. *Journal of the royal statistical  
892 society: Series b (statistical methodology)*, 71(2):319–392, 2009.
- 893 Rowan F Sage. Global change biology: a primer. *Global Change Biology*, 26(1):3–30, 2020.
- 894 Kari Saikkonen, Pedro E Gundel, and Marjo Helander. Chemical ecology mediated by fungal  
895 endophytes in grasses. *Journal of chemical ecology*, 39:962–968, 2013.
- 896 Christopher L Schardl, Carolyn A Young, Jerome R Faulkner, Simona Florea, and Juan Pan.  
897 Chemotypic diversity of epichloae, fungal symbionts of grasses. *fungal ecology*, 5(3):331–344,  
898 2012.
- 899 María Semmartin, Marina Omacini, Pedro E Gundel, and Ignacio M Hernández-Agramonte.

- 900 Broad-scale variation of fungal-endophyte incidence in temperate grasses. *Journal of Ecology*,  
901 103(1):184–190, 2015.
- 902 Sonia I Seneviratne, Xuebin Zhang, Muhammad Adnan, Wafae Badi, Claudine Dereczynski, A Di  
903 Luca, Subimal Ghosh, Iskhaq Iskandar, James Kossin, Sophie Lewis, et al. Weather and climate  
904 extreme events in a changing climate. 2021.
- 905 Daniel Simpson, Håvard Rue, Andrea Riebler, Thiago G Martins, and Sigrunn H Sørbye. Penal-  
906 ising model component complexity: A principled, practical approach to constructing priors.  
907 2017.
- 908 Michelle E Sneek, Jennifer A Rudgers, Carolyn A Young, and Tom EX Miller. Variation in the  
909 prevalence and transmission of heritable symbionts across host populations in heterogeneous  
910 environments. *Microbial Ecology*, 74:640–653, 2017.
- 911 Thomas F Stocker, Dahe Qin, G-K Plattner, Lisa V Alexander, Simon K Allen, Nathaniel L Bindoff,  
912 F-M Bréon, John A Church, Ulrich Cubasch, Seita Emori, et al. Technical summary. In *Climate  
913 change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment  
914 Report of the Intergovernmental Panel on Climate Change*, pages 33–115. Cambridge University  
915 Press, 2013.
- 916 Peter A Stott, Nathan P Gillett, Gabriele C Hegerl, David J Karoly, Dáithí A Stone, Xuebin Zhang,  
917 and Francis Zwiers. Detection and attribution of climate change: a regional perspective. *Wiley  
918 interdisciplinary reviews: climate change*, 1(2):192–211, 2010.
- 919 S Sully, DE Burkepile, MK Donovan, G Hodgson, and R Van Woesik. A global analysis of coral  
920 bleaching over the past two decades. *Nature communications*, 10(1):1–5, 2019.
- 921 E Toby Kiers, Todd M Palmer, Anthony R Ives, John F Bruno, and Judith L Bronstein. Mutualisms  
922 in a changing world: an evolutionary perspective. *Ecology letters*, 13(12):1459–1474, 2010.
- 923 Andrew T Tredennick, Giles Hooker, Stephen P Ellner, and Peter B Adler. A practical guide to

- 924 selecting models for exploration, inference, and prediction in ecology. *Ecology*, 102(6):e03336,  
925 2021.
- 926 Artemis D Treindl, Jessica Stapley, and Adrian Leuchtmann. Genetic diversity and population  
927 structure of epichloë fungal pathogens of plants in natural ecosystems. *Frontiers in Ecology and*  
928 *Evolution*, 11:1129867, 2023.
- 929 Kevin E Trenberth, John T Fasullo, and Theodore G Shepherd. Attribution of climate extreme  
930 events. *Nature climate change*, 5(8):725–730, 2015.
- 931 Amy M Truitt, Martin Kapun, Rupinder Kaur, and Wolfgang J Miller. Wolbachia modifies thermal  
932 preference in drosophila melanogaster. *Environmental microbiology*, 21(9):3259–3268, 2019.
- 933 Shripad D. Tuljapurkar. Population dynamics in variable environments. III. Evolutionary dynam-  
934 ics of r-selection. *Theoretical Population Biology*, 21(1):141–165, February 1982. ISSN 0040-5809.  
935 doi: 10.1016/0040-5809(82)90010-7. URL <http://www.sciencedirect.com/science/article/pii/0040580982900107>.
- 937 Arantxa Urdangarin, Tomás Goicoa, and María Dolores Ugarte. Evaluating recent methods to  
938 overcome spatial confounding. *Revista Matemática Complutense*, 36(2):333–360, 2023.
- 939 Veronika Vikuk, Carolyn A Young, Stephen T Lee, Padmaja Nagabhyru, Markus Krischke, Mar-  
940 tin J Mueller, and Jochen Krauss. Infection rates and alkaloid patterns of different grass species  
941 with systemic epichloë endophytes. *Applied and Environmental Microbiology*, 85(17):e00465–19,  
942 2019.
- 943 Maria von Cräutlein, Marjo Helander, Helena Korpelainen, Päivi Helena Leinonen, Beatriz R  
944 Vázquez de Aldana, Carolyn Anne Young, Iñigo Zabalgoitia, and Kari Saikonen. Ge-  
945 netic diversity of the symbiotic fungus epichloë festucae in naturally occurring host grass  
946 populations. *Frontiers in microbiology*, 12:756991, 2021.
- 947 Zhengfeng Wang, Chunjie Li, and James White. Effects of epichloë endophyte infection on

- 948 growth, physiological properties and seed germination of wild barley under saline conditions.
- 949 *Journal of Agronomy and Crop Science*, 206(1):43–51, 2020.
- 950 Robert J Warren and Mark A Bradford. Mutualism fails when climate response differs between  
951 interacting species. *Global Change Biology*, 20(2):466–474, 2014.
- 952 Nicole S Webster, Rose E Cobb, and Andrew P Negri. Temperature thresholds for bacterial  
953 symbiosis with a sponge. *The ISME journal*, 2(8):830–842, 2008.
- 954 James F White and Garry T Cole. Endophyte-host associations in forage grasses. i. distribution  
955 of fungal endophytes in some species of lolium and festuca. *Mycologia*, 77(2):323–327, 1985.
- 956 Franziska M Willems, JF Scheepens, and Oliver Bossdorf. Forest wildflowers bloom earlier as  
957 europe warms: lessons from herbaria and spatial modelling. *New Phytologist*, 235(1):52–65,  
958 2022.
- 959 Charles G Willis, Elizabeth R Ellwood, Richard B Primack, Charles C Davis, Katelin D Pearson,  
960 Amanda S Gallinat, Jenn M Yost, Gil Nelson, Susan J Mazer, Natalie L Rossington, et al. Old  
961 plants, new tricks: Phenological research using herbarium specimens. *Trends in ecology &*  
962 *evolution*, 32(7):531–546, 2017.
- 963 Chao Xia, Nana Li, Yawen Zhang, Chunjie Li, Xingxu Zhang, and Zhibiao Nan. Role of epichloë  
964 endophytes in defense responses of cool-season grasses to pathogens: A review. *Plant disease*,  
965 102(11):2061–2073, 2018.
- 966 Kentaro Yoshida, Eriko Sasaki, and Sophien Kamoun. Computational analyses of ancient  
967 pathogen dna from herbarium samples: challenges and prospects. *Frontiers in plant science*,  
968 6:771, 2015.

## Appendix A

060

970

971 Appendix to "Increasing Prevalence of plant-fungal symbiosis across two  
972 centuries of environmental change"

031

972

973 Authors:

974 Joshua C. Fowler<sup>1,2\*</sup>

975 Jacob Moutouama<sup>1</sup>

976 Tom E. X. Miller<sup>1</sup>

977

978 1. Rice University, Department of BioSciences, Houston, Texas 77006

979 2. University of Miami, Department of Biology, Miami, Florida

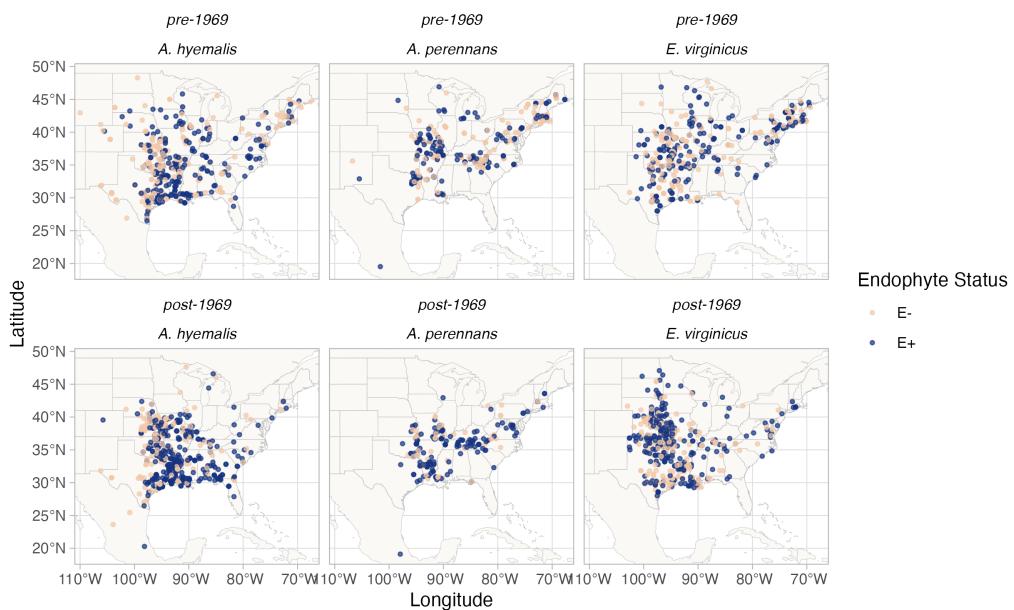
<sup>980</sup> \* Corresponding author; e-mail: jcf221@miami.edu.

981 **Contents:**

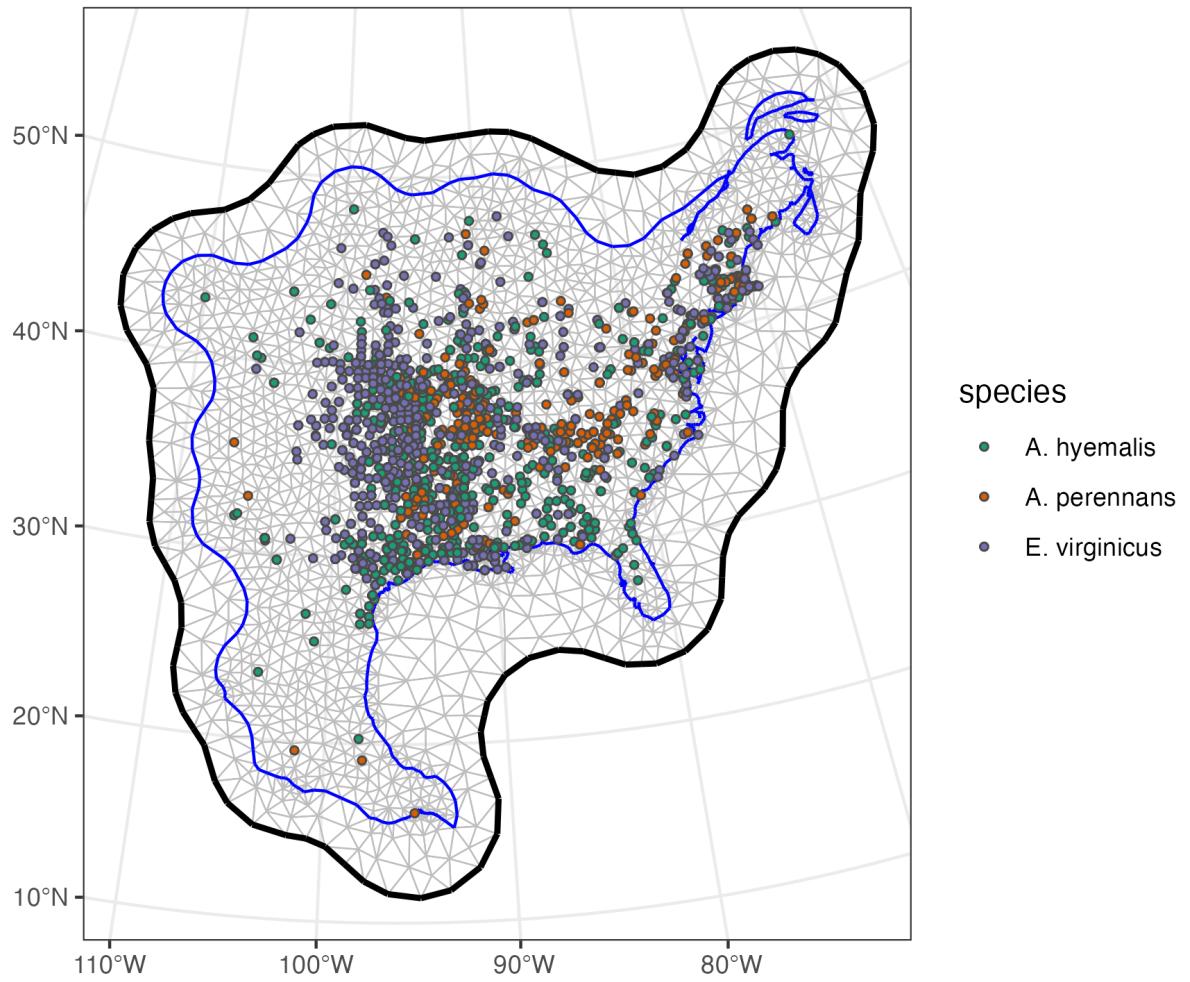
**982 Appendix A includes: Figure A1 - Figure A15, Table A1, and Supporting Methods).**

983

## Supplemental Figures



**Figure A1: Endophyte presence/absence in specimens of each host species.** Points show collection locations colored according to whether the specimen contained endophytes (E+; blue points) or did not contain endophytes (E-, tan points). To visualize temporal change, the data are faceted before and after the median year of collection. Map lines delineate study areas and do not necessarily depict accepted national boundaries.



**Figure A2: Triangulation mesh used to estimate spatial dependence between data points.** Grey lines indicate edges of triangles used to define distances between observations. Colored points indicate locations of sampled herbarium specimens for each host species, and the blue line shows the convex hull and coastline used to define the edge of the mesh around the data points. The thick black line shows the convex hull defining a buffer space around the edge of the mesh to reduce the influence of edge effects on model estimates.

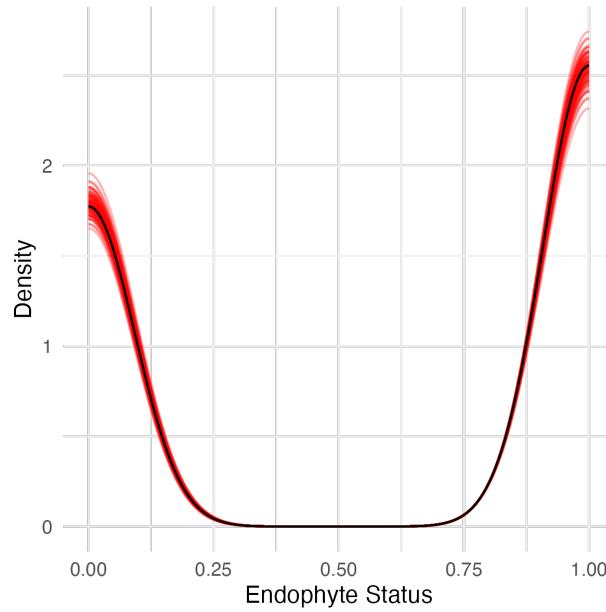


Figure A3: **Graphical posterior predictive check of the endophyte prevalence model fit.** Consistency between **observed** data and **predicted** values indicate that the fitted model accurately describes the data. Graph shows density curves for the observed data (black) along with 100 **predicted** datasets (red).

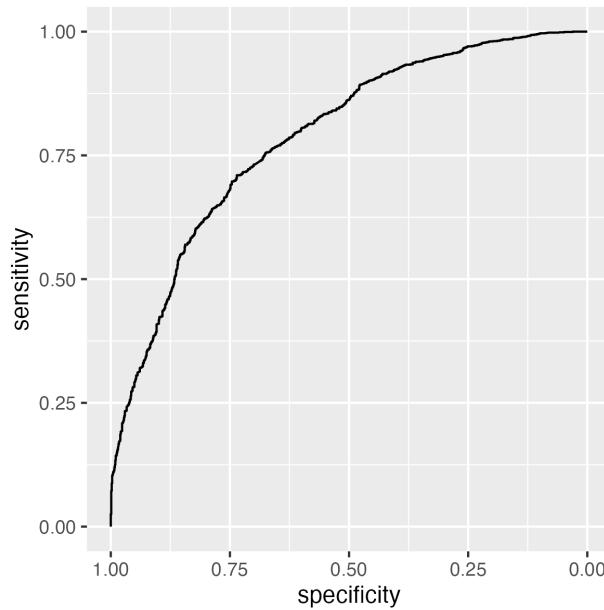


Figure A4: **ROC plot showing performance of the endophyte prevalence model in classifying observations according to endophyte status within the in-sample training data from herbarium collections.** The curves show adequate model performance for observed data. The AUC value is 0.79.

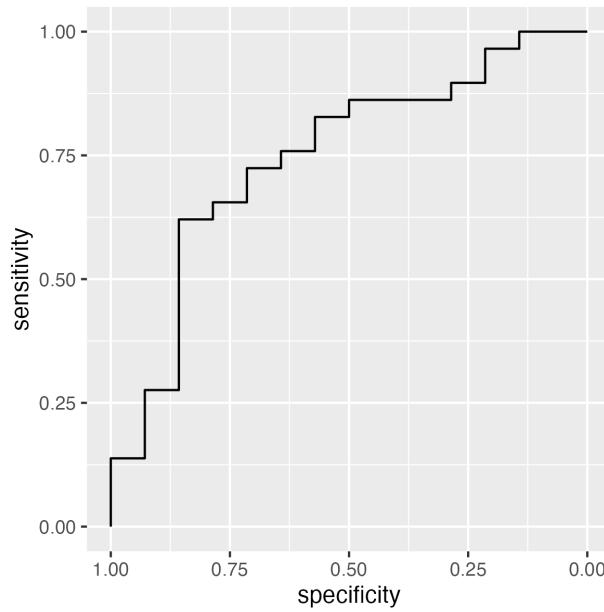
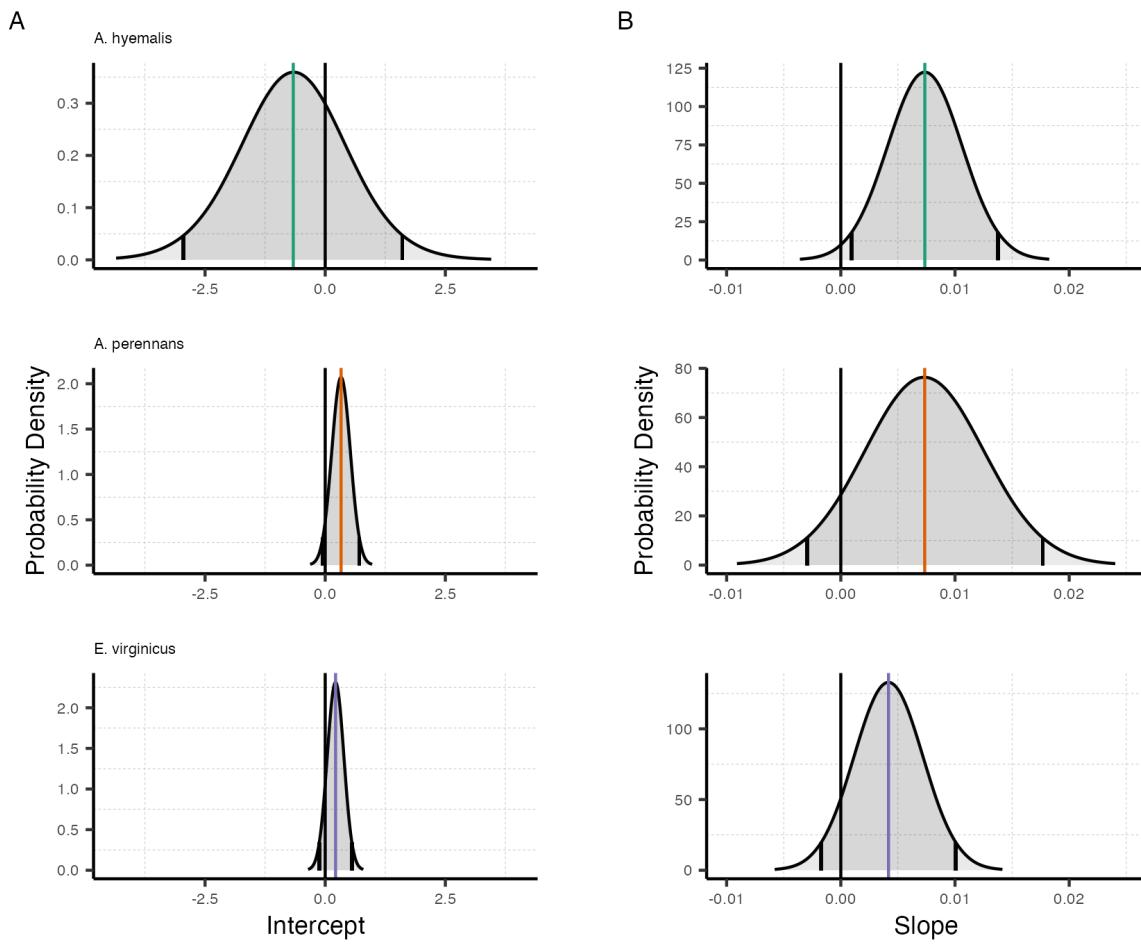
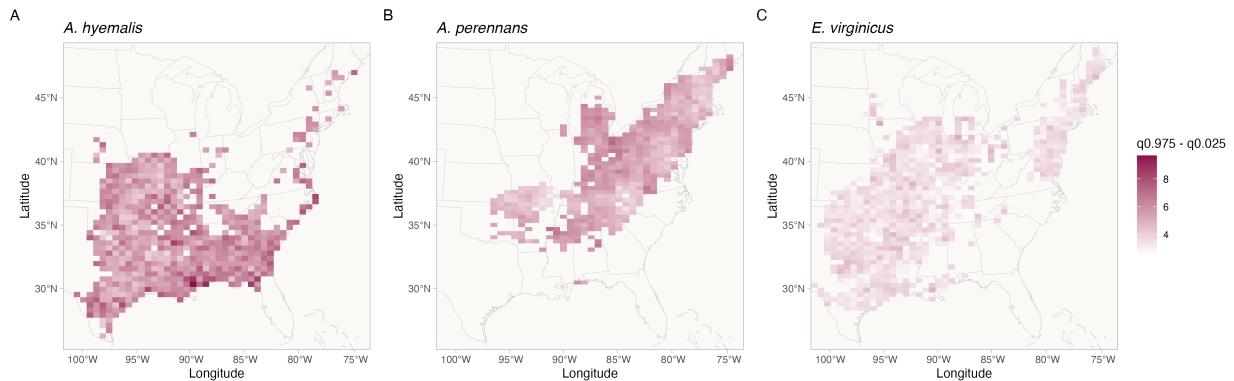


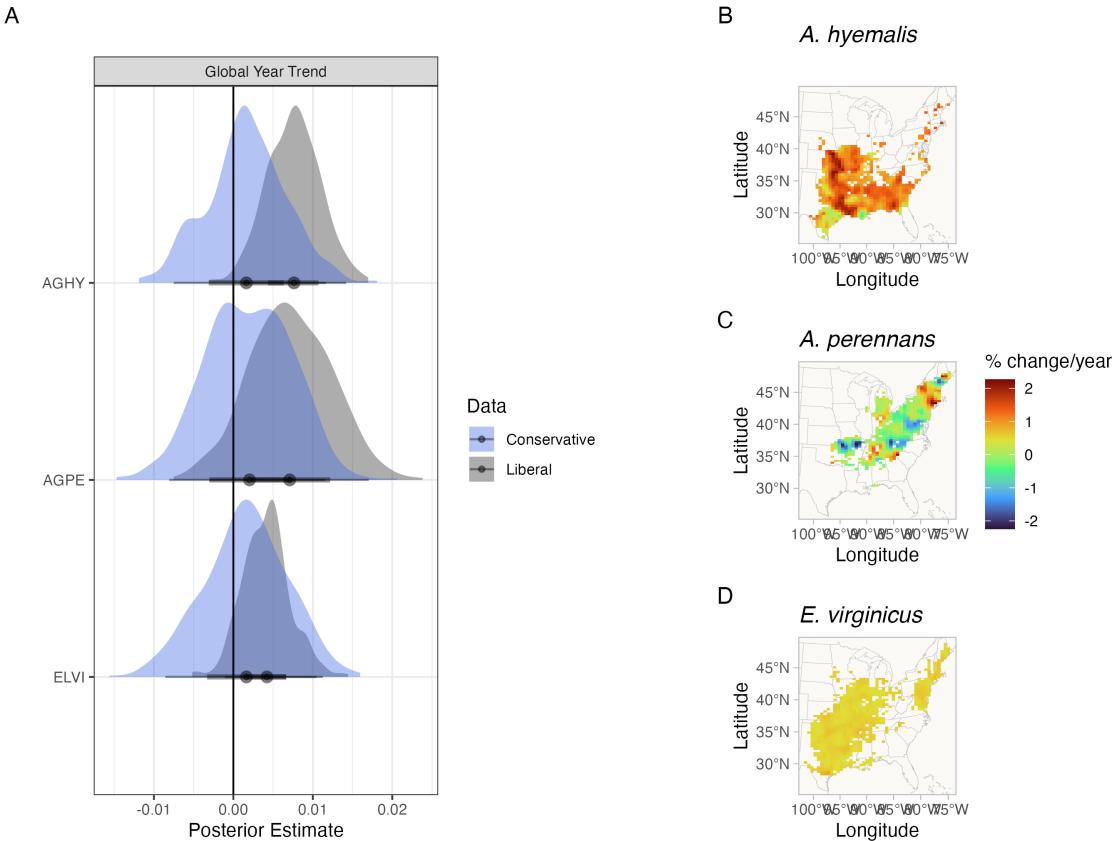
Figure A5: **ROC plot showing performance of the endophyte prevalence model in classifying observations according to endophyte status within the out-of-sample test data from contemporary surveys.** The curves show adequate model performance for test data. The AUC value is 0.77.



**Figure A6: Posterior estimates of parameters describing global intercept and temporal trends from the endophyte prevalence model.** Density curves show the probability density along with mean (colored line) and 95% CI (black lines) for the (A) intercept and (B) slope terms,  $\mathbf{A}$  and  $\mathbf{T}$  respectively from Eqn. 1. Colors represent each host species



**Figure A7: Credible interval width of temporal trends in endophyte prevalence across the distribution of each host species estimated from the endophyte prevalence model.** Shading represents the range of the 95% posterior credible interval given in units of % change in prevalence/year for spatially varying slopes,  $\tau$  from Eqn. 1. Map lines delineate study areas and do not necessarily depict accepted national boundaries.



**Figure A8: Comparison of endophyte prevalence model estimates fit to data with liberal versus conservative endophyte scores.** Liberal and conservative scores document uncertainty in the endophyte identification process. Each specimen was given both a liberal and conservative scores. In cases of uncertain identification, the liberal status assumed a potential endophyte identification was more likely to be endophyte-positive while the conservative status assumed that the potential endophyte identification was less likely to be endophyte-positive. (A) Posterior estimates of global temporal trend ( $T$  from Eqn. 1) for the endophyte prevalence model fit to liberal scores (grey) and to conservative scores (blue). Maps show the spatially varying temporal trend estimates ( $\tau$  from Eqn. 1) from the endophyte prevalence model fit to conservative scores for (B) *A. hyemalis*, (C) *A. perennans*, and (D) *E. virginicus*. Note that the color scale differs between this visualization and Fig. 3 that shows estimates fit using liberal endophyte scores.

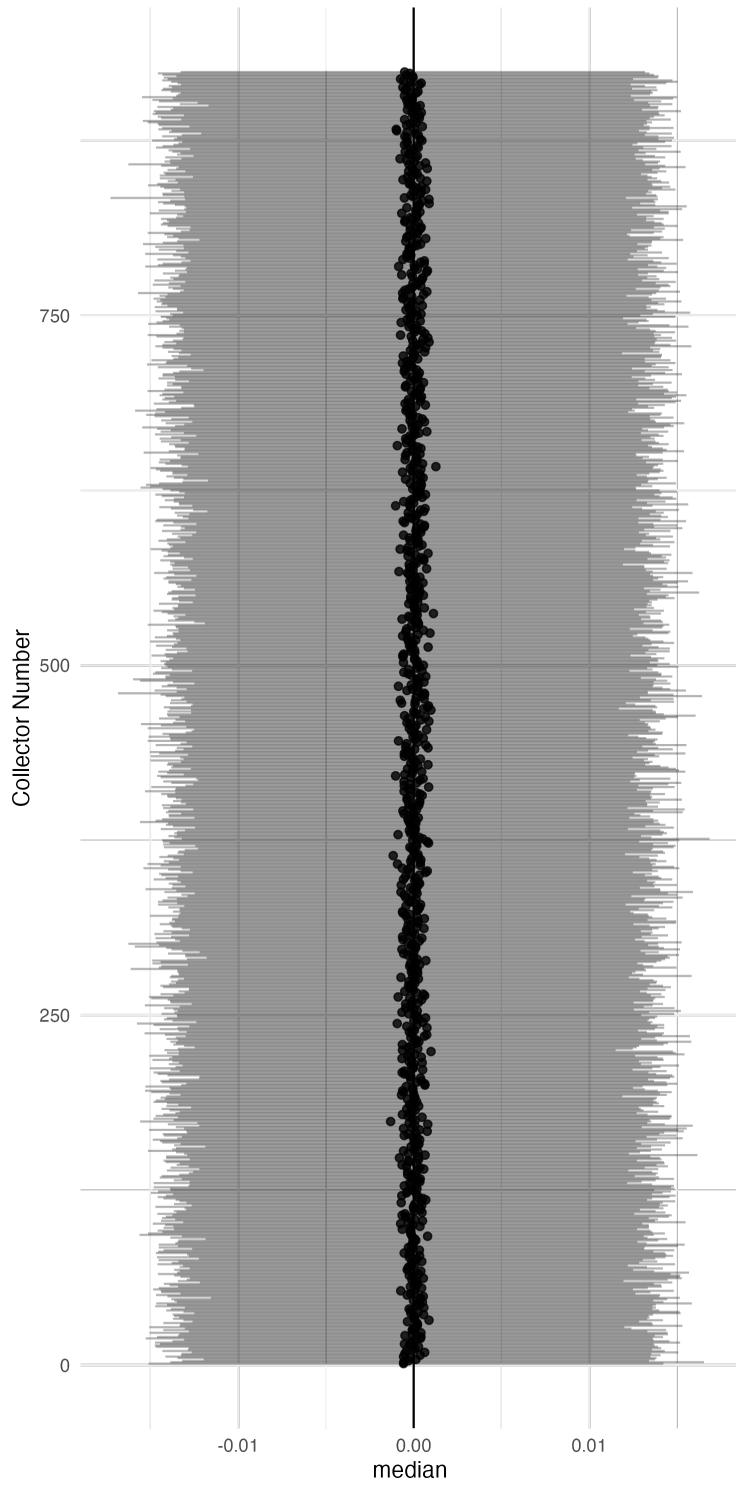


Figure A9: Posterior estimates of collector random effects from endophyte prevalence model.

Collector random effects are denoted  $\chi$  in Eqn. 1 and represent variance associated with researchers who collected historic herbarium specimens. Points show posterior median along with 95% CI for each of 924 individual collectors.

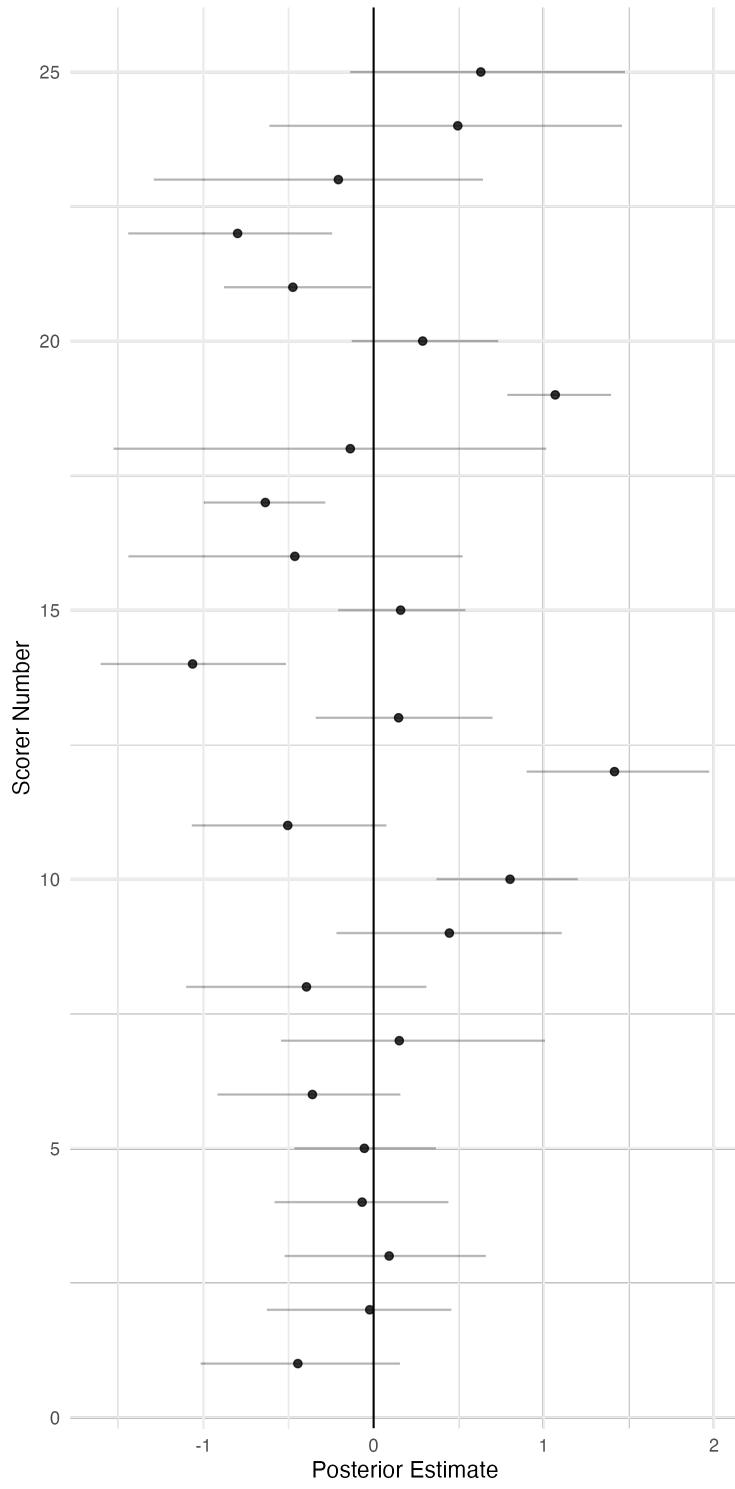
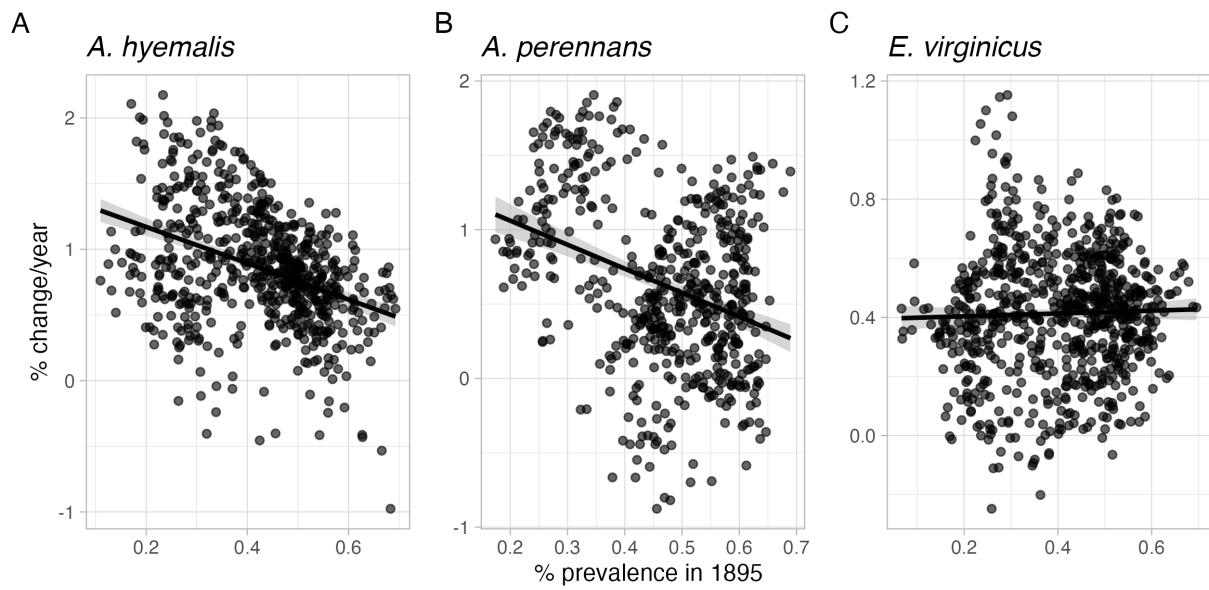
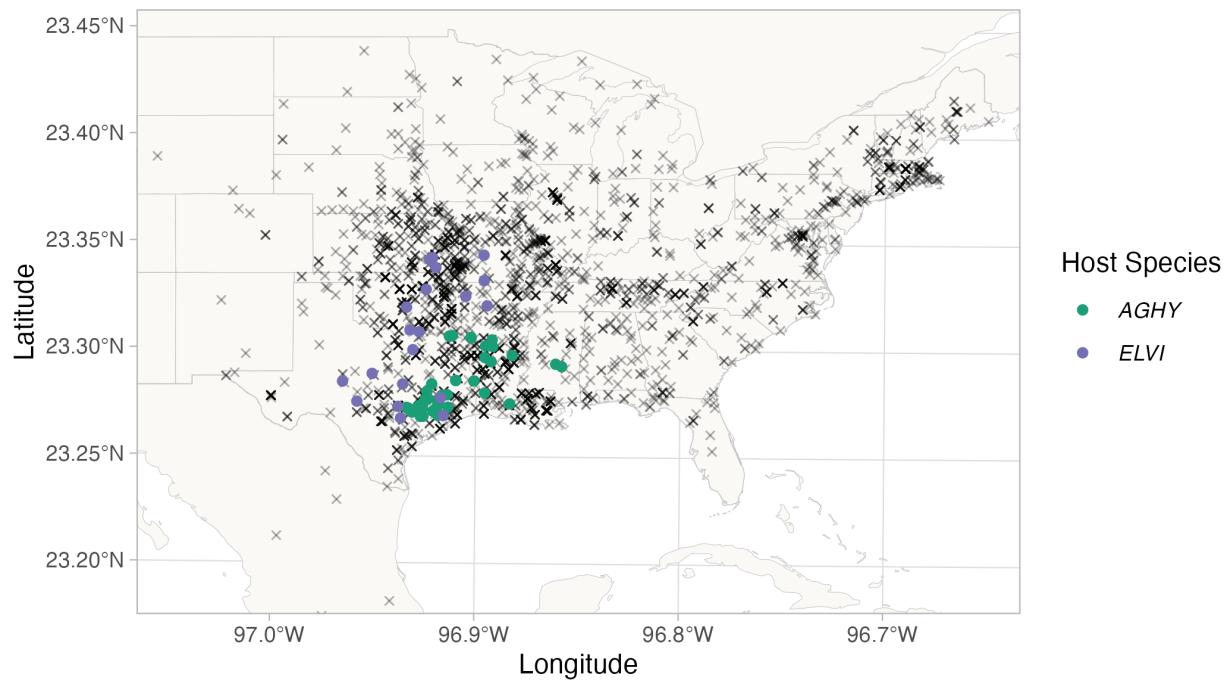


Figure A10: Posterior estimates of scorer random effects from endophyte prevalence model.

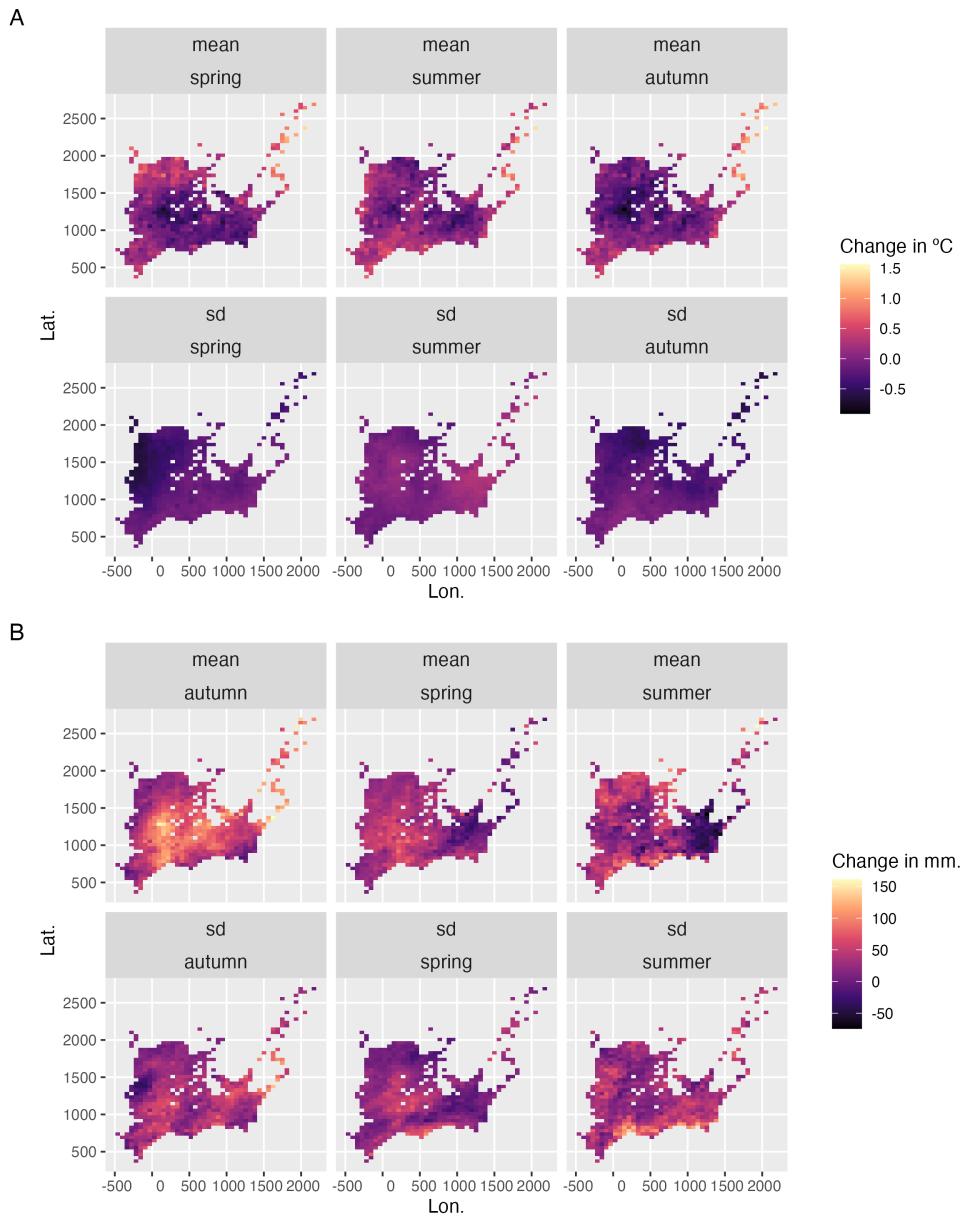
Scorer random effects are denoted  $\omega$  in Eqn. 1 and represent variance associated with researchers who identified *Epichloë* endophytes within herbarium specimen tissue samples. Points show posterior median along with 95% CI for each of 25 individual scorers.



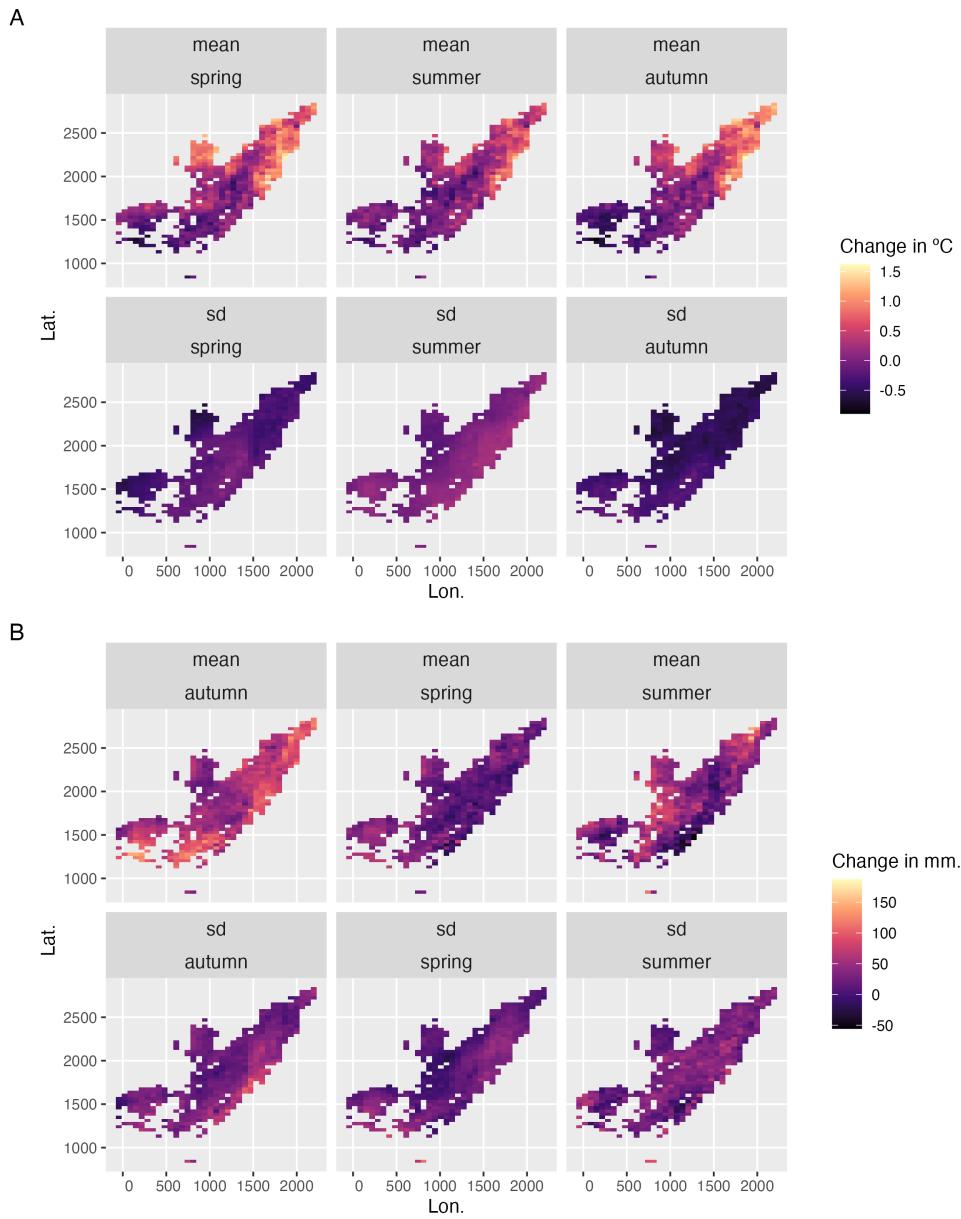
**Figure A11: Relationship between initial prevalence and temporal trends in prevalence estimated from the endophyte prevalence model.** Points show predicted posterior mean temporal trend for each species at pixels across each host distribution ((A) *A. hyemalis*, (B) *A. perennans*, and (C) *E. virginicus*). along with a linear regression and shaded ribbon showing 95% confidence interval.



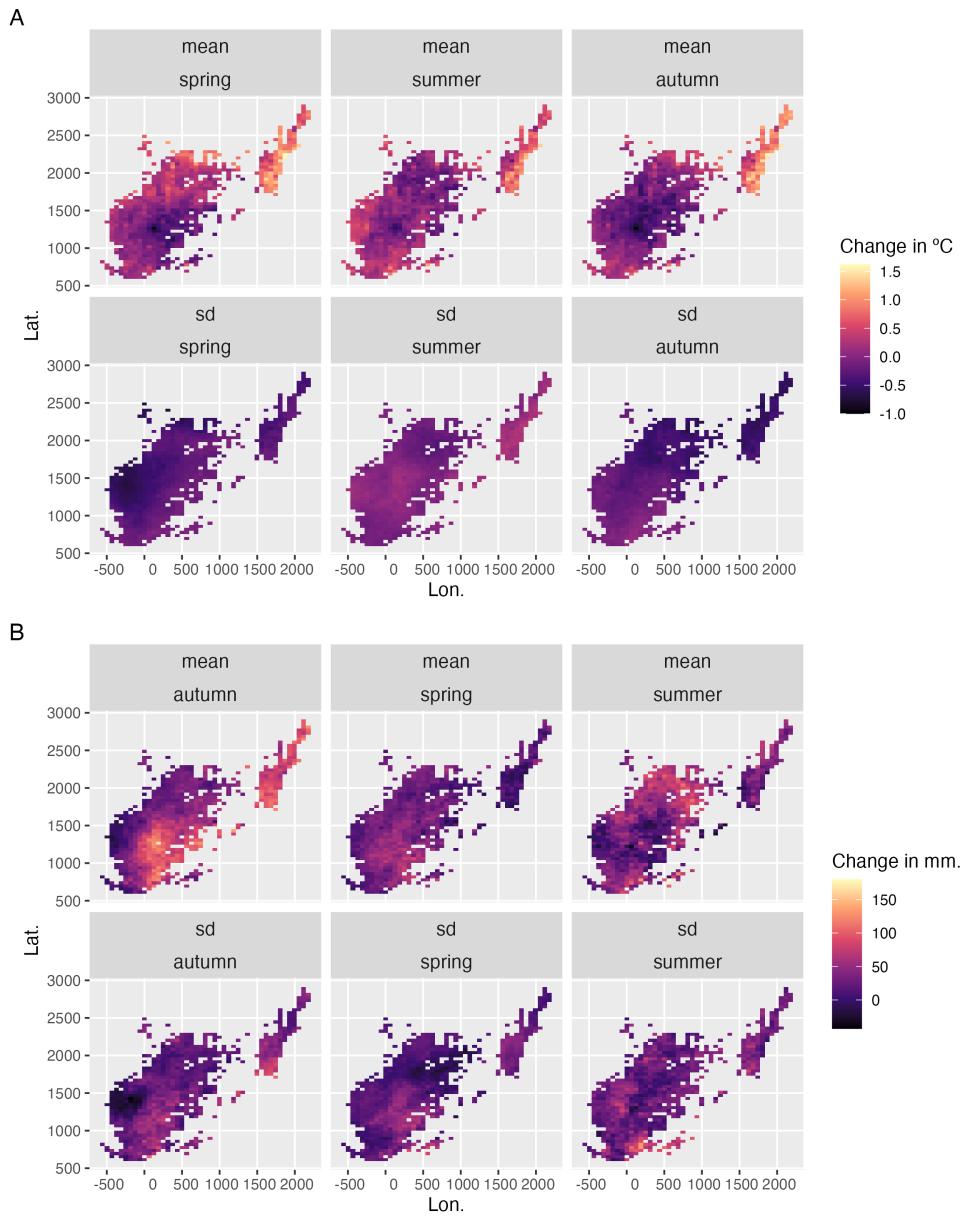
**Figure A12: Locations of contemporary surveys of endophytes used as "test" data to evaluate predictive ability of the endophyte prevalence model.** Points are locations of host populations surveyed between 2013 and 2019 for endophytes, colored by species (*A. hyemalis*: green, *E. virginicus*: purple). Black crosses show the historical herbarium collection locations used as "training" data for the endophyte prevalence model.



**Figure A13: Change in seasonal climate variables between the periods 1895-1925 and 1990-2020 across the distribution of *A. hyemalis*.** Color represents change in (A) seasonal temperature ( $^{\circ}\text{C}$ ) and (B) seasonal precipitation (mm.). Maps show pixels covering the modeled distribution of *A. hyemalis* used in *post hoc* climate regression analysis.



**Figure A14: Change in seasonal climate variables between the periods 1895-1925 and 1990-2020 across the distribution of *A. perennans*.** Color represents change in (A) seasonal temperature ( $^{\circ}\text{C}$ ) and (B) seasonal precipitation (mm.). Maps show pixels covering the modeled distribution of *A. perennans* used in *post hoc* climate regression analysis.



**Figure A15: Change in seasonal climate variables between the periods 1895-1925 and 1990-2020 across the distribution of *E. virginicus*.** Color represents change in (A) seasonal temperature ( $^{\circ}\text{C}$ ) and (B) seasonal precipitation (mm.). Maps show pixels covering the modeled distribution of *E. virginicus* used in *post hoc* climate regression analysis.

Table A1: Summary of herbarium samples across collections (no. of specimens)

Herbarium Collection	<i>A. hyemalis</i>	<i>A. perennans</i>	<i>E. virginicus</i>
Botanical Research Institute of Texas	350	190	198
Louisiana State University	72	38	62
Mercer Botanic Garden	3	–	6
Missouri Botanic Garden	210	205	122
Texas A&M	100	–	72
University of Kansas	134	34	197
University of Oklahoma	85	34	95
University of Texas & Lundell	183	91	102
Oklahoma State University	51	10	74

985

## Supporting Methods

986

### ODMAP Protocol

987 [Overview](#)

988 **Model purpose:** Mapping current distribution of *Epichloë* host species.

989 **Target species:** *Agrostis hyemalis*, *Agrostis perennans*, and *Elymus virginicus*.

990 **Study area:** Eastern North America

991 **Spatial extent:** -125.0208, -66.47917, 24.0625, 49.9375 (xmin, xmax, ymin, ymax).

992 **Spatial resolution:** 0.04166667, 0.04166667 (x, y).

993 **Temporal extent:** 1990 to 2020.

994 **Boundary:** Natural.

995 [Data](#)

996 **Observation type:** Occurrence records from Global Biodiversity Information Facility and

997 herbarium collection across eastern North America. We used 713 occurrences records for  
998 *Agrostis hyemalis*, 656 occurrence records for *Agrostis perennans* and 2338 for *Elymus virginicus*.

999 **Response data type:** occurrence record, presence-only.

1000 **Coordinate reference system:** WGS84 coordinate reference system (EPSG:4326 code)

1001 **Climatic data:** raster data extracted from PRISM

1002 **Model**

1003 **Model assumption:** We assumed that the target species are at equilibrium with their environment.  
1004

1005 **Algorithms:** Maximum entropy (maxent)

1006 **Workflow:** We described the workflow in the method section of the manuscript.

1007 **Software:** All statistics were performed using Maxent 3.3.4 and R4.3.1 with packages terra,  
1008 usdm, spThin and dismo.

1009 **Code availability:** Available through this link: <https://github.com/joshuacfowler/EndoHerbarium>

1010 **Data availability:** Will be available upon acceptance

1011 **Assessment**

1012 We used AUC to test model performance.

1013 **Prediction**

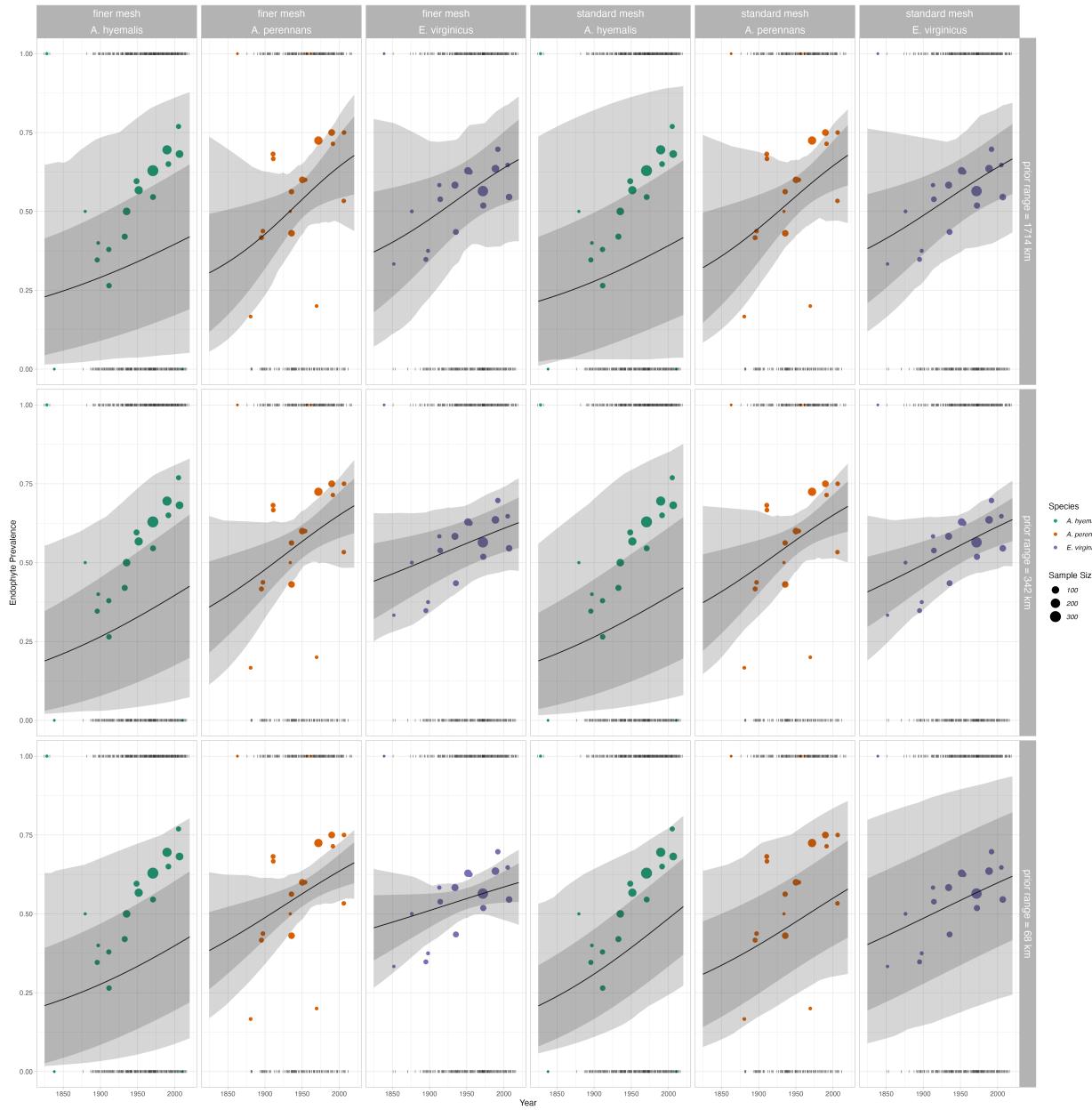
1014 We predicted the probability of presence of the host species as a binary maps (presence or  
1015 absence)

## 1016 *Mesh and Prior Sensitivity Analysis*

1017 To test the influence that the triangulation mesh and choice of priors has on results, we compared  
1018 model results across a range of meshes and priors. We re-ran our model for the mesh used in  
1019 main body of the text (Fig. A2), which we refer to as the "standard mesh", and with a mesh with  
1020 smaller minimum vertices (finer mesh). Finer scale meshes increase computation time. For each  
1021 of these meshes, we ran the model with a range of priors defining the spatial range of our spatial  
1022 random effects: 342km (the prior used for presented results), as well as ranges five times smaller

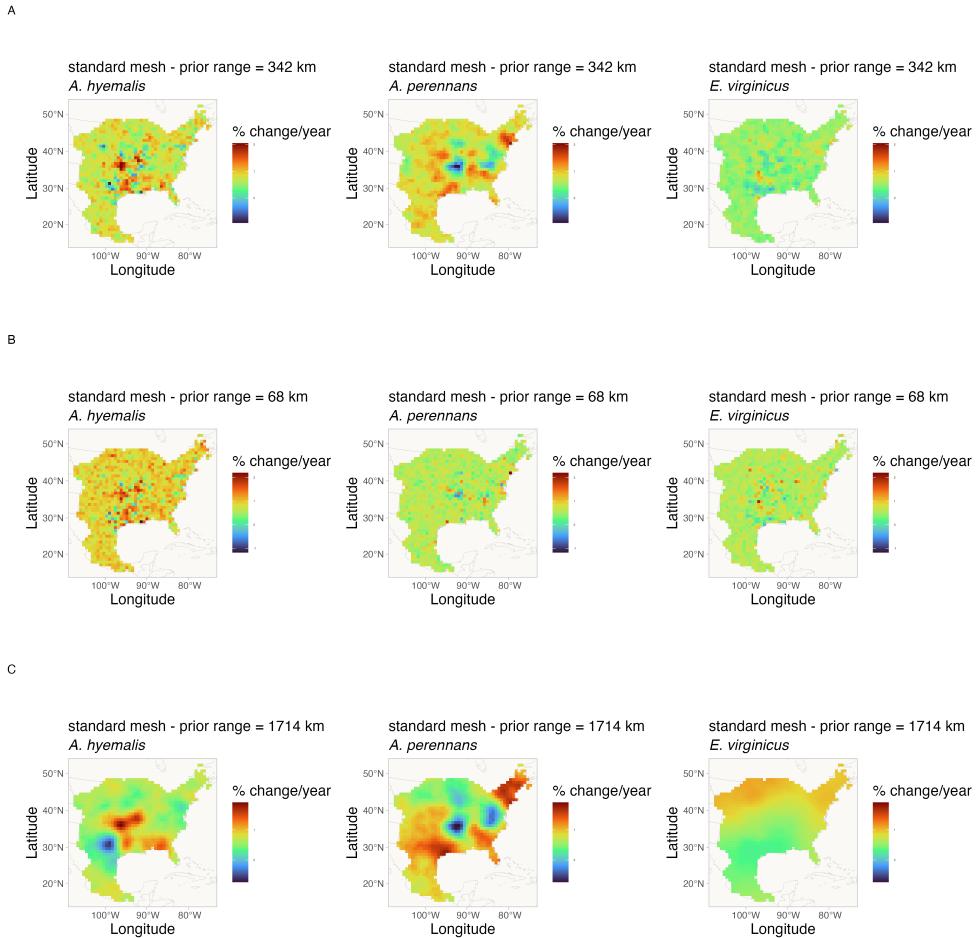
1023 (68 km) and five times larger (1714 km). We found generally that these choices did not alter the  
1024 direction of model predictions, but did influence the associated uncertainty and magnitude of  
1025 some effects.

1026 For overall temporal trends, we found that models with differing priors predicted consistently  
1027 positive relationships over time (Fig. A16).

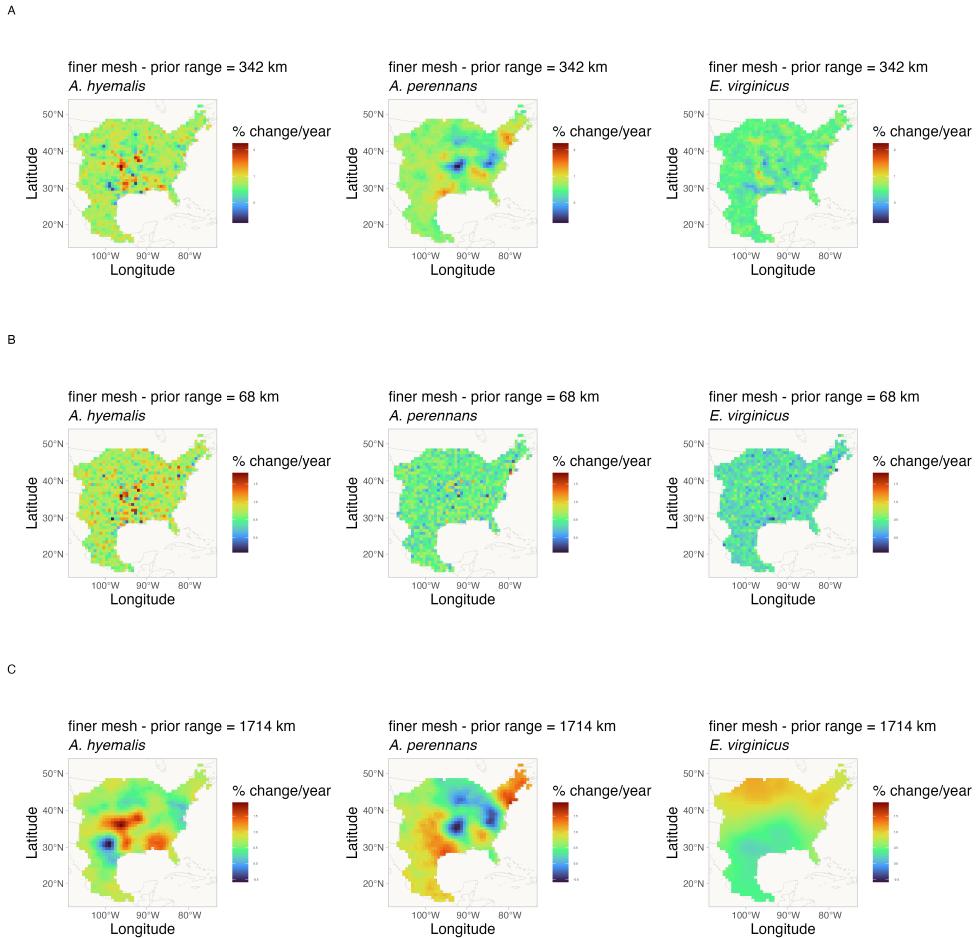


**Figure A16: Overall trend in endophyte prevalence evaluated for endophyte prevalence models with different range priors on spatially structured random effects, and for two different triangulation meshes.** Data used in model fitting is the same across all panels and as in the main text. Note that these plots, as compared to Fig. 2 in main text, show mean trends and do not incorporate variance associated with collector and scorer random effects.

1028 For spatially-varying temporal trends, we found that models with different priors predicted  
1029 consistent spatial patterns in temporal trends, although the range of this prediction varied de-  
1030 pending on the prior and mesh (Fig. A17 - A18). One noteworthy result of this analysis is that  
1031 combinations of prior choice and mesh can introduce instability in model fitting. This is evident  
1032 in A17 panel B and A18 panel B, where the prior range is smaller than the minimum vertex  
1033 length of the mesh. Model fitting takes an extended time period and the model struggles to  
1034 identify variation across space. Results with a set of prior ranges (Fig. A17 - A and C; Fig. A18  
1035 - A and C) result in models that estimate trends across space of the same direction and order of  
1036 magnitude, although the "smoothness" of these predictions vary.



**Figure A17: Spatially-varying trends in endophyte prevalence evaluated for the endophyte prevalence model with different range priors on spatially structured random effects, and for the "standard" mesh.** Data used in model fitting is the same across all panels and as in the main text. Shading indicates the magnitude and direction of predicted trends for each of three host species for each of three prior ranges (rows A-C). Note that each plot has an individual scale bar.



**Figure A18: Spatially-varying trends in endophyte prevalence evaluated for the endophyte prevalence model with different range priors on spatially structured random effects, and for the "finer" mesh.** Data used in model fitting is the same across all panels and as in the main text. Shading indicates the magnitude and direction of predicted trends for each of three host species for each of three prior ranges (rows A-C). Note that each plot has an individual scale bar.

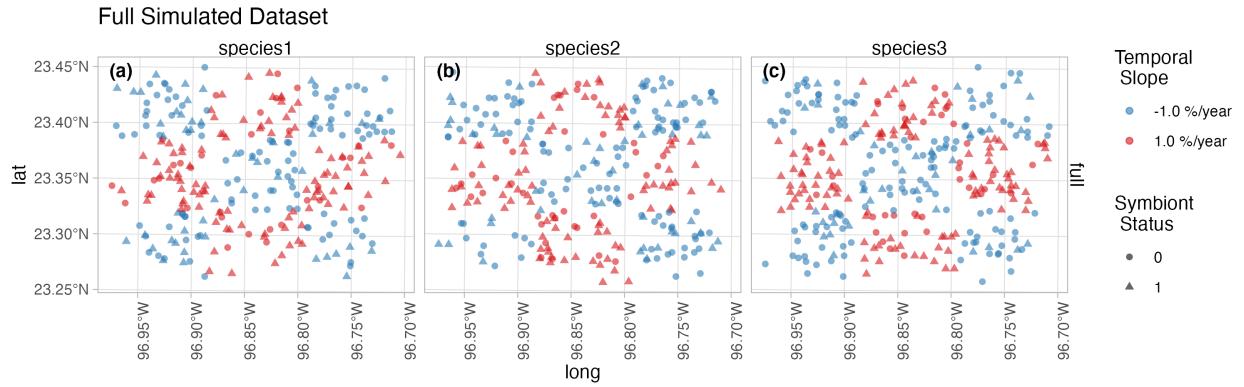
1037

## *Spatially-biased Sample Size Simulation Analysis*

1038 To examine how data that is unevenly distributed across host distributions may influence inter-  
1039 pretation of spatially-varying coefficients, we performed a simulation analysis. Our focal species,  
1040 *Agrostis hyemalis*, *Agrostis perennans*, and *Elymus virginicus*, are widely distributed grasses across  
1041 the eastern United States that host *Epichloë* fungal endophytes. For logistical reasons, our sam-  
1042 pling visits to herbaria focused on herbaria in the central southern U.S., which resulted in un-  
1043 evenly distributed data across each host species' range. This is particularly noteable for *Agrostis*  
1044 *perennans* which has the most northern distribution and relatively fewer total collected speci-  
1045 mens compared to the other focal species. Thus, a significant portion in the northeast of this  
1046 species' range is relatively sparsely sampled. Our analysis presented in the main text identified  
1047 this region as having strong increase in endophyte prevalence. Future visits to herbaria with re-  
1048 gional focuses in the Northeastern US would certainly garner new specimens that could provide  
1049 valuable insights into shifting host and symbiont distributions.

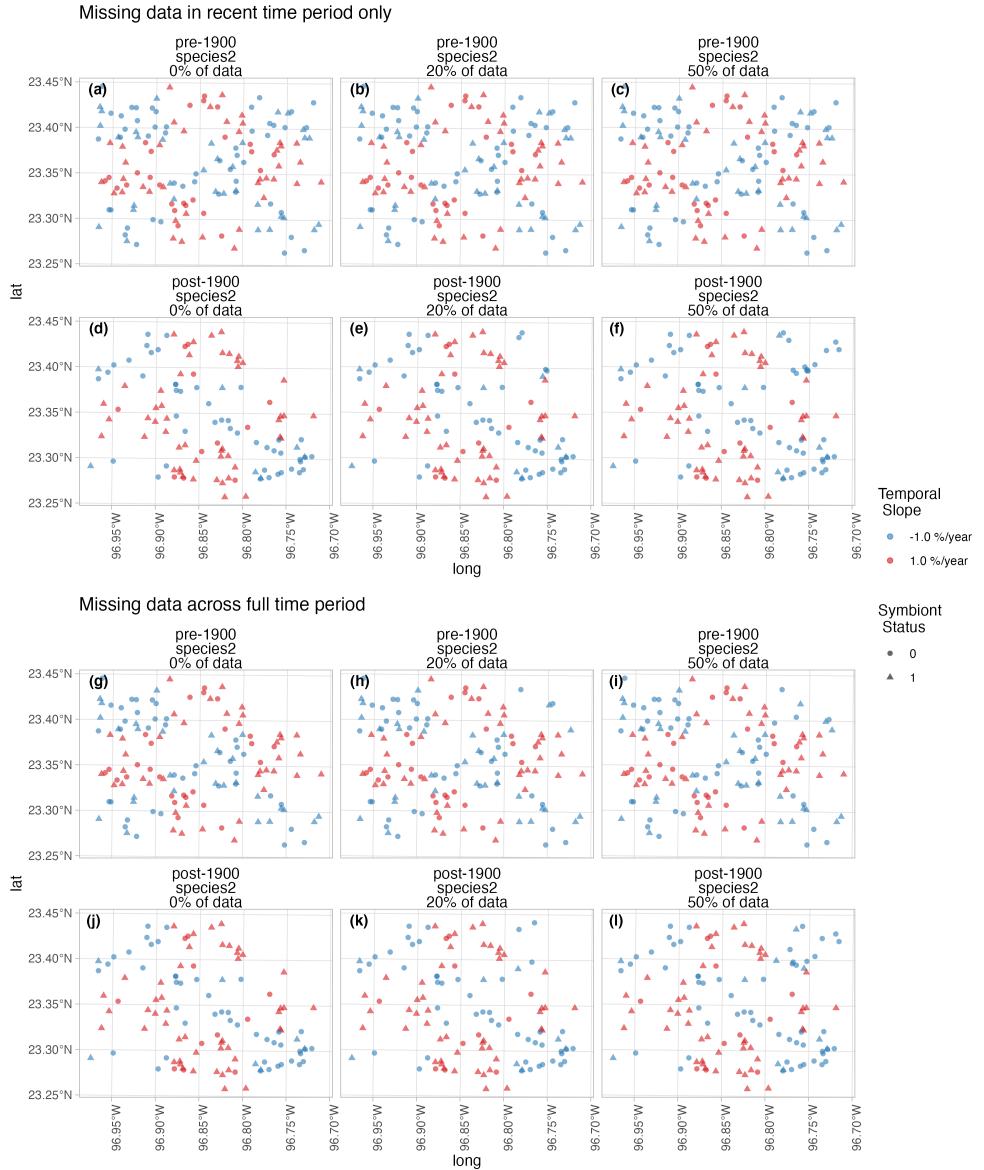
1050 *Simulation of spatially-biased symbiont occurrence data*

1051 We simulated datasets with varying levels of missing-ness to examine how this missing-ness  
1052 influenced the estimation of spatially-varying trend estimates. We first generated 300 data points  
1053 for each of three hypothetical species at random positions across an area approximating the scale  
1054 of our focal data. Each data point was randomly assigned a year of collection across 200 years.  
1055 We then simulated data from a Bernoulli process with trends alternating across nine regions (Fig.  
1056 A19) in a 3X3 grid pattern. This grid pattern was intended to create a complex spatial layout of  
1057 trends, where trends were either an increase of 1% per year or a decrease of 1% per year.



**Figure A19: Full simulated dataset of symbiotic association with spatially-varying temporal trends.** Color indicates the slope parameter used to simulate trends in endophyte status across nine "regions" for three species. Data are assigned collection years across a period of 200 years. Shape indicates the presence (1) or absence (0) of a symbiont.

1058 From this full data, we generated six additional datasets with missing-ness in the northeast  
 1059 region of the simulated data for hypothetical species 2. The data remained the same for Species 1  
 1060 and for species 3 across all datasets. For these six datasets, we removed data points at random in  
 1061 six ways: 0% of datapoints in northeast region, 0% of recent datapoints, only 20% of datapoints,  
 1062 only 20% of recent datapoints, only 50% of datapoints, and only 50% of recent datapoints (Fig.  
 1063 A20). We define the datapoints as part of the recent time period if they occur later than the  
 1064 median year. The result is 6 scenarios exploring degrees of spatial and temporal bias.



**Figure A20: Six simulated datasets representing scenarios of spatially-biased missingness for Species 2.** Missingness was imposed in the northeast region for six scenarios: 0% of recent datapoints available (a,d); only 20% of recent datapoints (b,e); only 50% of recent datapoints (c,f); 0% of datapoints across the full time period available (g,j); only 20% of datapoints across the full time period (h,k); and only 50% of datapoints across the full time period(i,l). Missingness was imposed only for hypothetical Species 2; Species 1 and 3 remain as in Figure A19. Color indicates the slope parameter used to simulate trends in endophyte status across 9 regions in a 3x3 grid. Shape indicates the presence (1) or absence (0) of a symbiont.

1065 *Statistical analysis*

1066 We analyzed each dataset with a model given by Eqn. A1 similar in construction to that used in  
1067 our central analysis.

$$\text{logit}(\hat{P}_{h,i}) = A_h + T_h * \text{year}_i + \alpha_{h,l_i} + \tau_{h,l_i} * \text{year}_i + \delta_{l_i} \quad (\text{A1})$$

1068 Where symbiont presence/absence of the  $i^{th}$  specimen ( $P_{h,i}$ ) was modeled as a Bernoulli re-  
1069 sponse variable with expected probability of symbiont occurrence  $\hat{P}_{h,i}$  for each host species  $h$ . We  
1070 modeled  $\hat{P}_{h,i}$  as a linear function of intercept  $A_h$  and slope  $T_h$  defining the global trend in en-  
1071 dophyte prevalence specific to each host species as well as with spatially-varying intercepts  $\alpha_{h,l_i}$   
1072 and slopes  $\tau_{h,l_i}$  associated with location ( $l_i$ , the unique latitude-longitude combination of the  $i$ th  
1073 observation). Similar to the SVC model of our central analysis (Eqn. 1), we estimated a shared  
1074 variance term with the spatially-dependent random effect  $\delta_{l_i}$ , intended to account for residual  
1075 spatial variation. However in this analysis we omit i.i.d.-random effects terms associated with  
1076 collector and scorer identity ( $\chi_{c_i}$  and  $\omega_{s_i}$  in Eqn. 1) for the sake of simplicity.

1077 *Influence of spatially-biased sampling on model interpretation*

1078 Our analysis of the full simulated data shows that our model is suitably flexible to capture com-  
1079 plex spatial patterns in temporal trends (Fig. A21 a-c). Beyond this, the model also qualitatively  
1080 captures the spatial patterns in temporal trends even with large amounts of data missingness (i.e  
1081 missing up to 80% of the datapoints (Fig. A21 p-r)).



```
./Plots/sim_svc_plot.png
```

Figure A21: **Mean predicted spatially-varying trend in symbiont prevalence across datasets with different levels of missingness.** Color indicates the estimated mean temporal trend within each pixel across the simulated data. Panels show estimates for models fit to different levels of missing data for species 2 in the northeast region ((a-c) the full dataset, (d-f) missing all datapoints across entire temporal period, (g-i) missing all datapoints only during the recent period, (j-l) missing 80% of the datapoints only during the recent period, (m-o) missing 50% of the datapoints only during the recent period, (p-r) missing 80% of the datapoints across the entire temporal period, (s-u) missing 50% of the datapoints across the entire temporal period). The mesh boundary that bounds the "full" simulated dataset is plotted in each panel.

1082 While this analysis is not an exhaustive examination of the influence of sampling bias on  
1083 our results for several reasons, including not examining how different strengths in temporal

1084 trends, different spatial arrangements of missing-ness influence model estimates, or different sam-  
1085 ple sizes, it demonstrates that the spatially-varying modelling framework implemented in INLA  
1086 we easily can suitably recover regional trends even with significant spatially-bias within data  
1087 collection, and further the analysis is likely robust to temporally-structured bias (missing data  
1088 within recent collection period). Future work could more fully explore the scenarios that cause  
1089 this ability to break down. We expect this simulation reflects what may be a common scenario  
1090 for research investigating global change using natural history specimens. Collection effort by  
1091 trained taxonomists and professional collectors peaked in the past, and collections contain rel-  
1092 atively fewer modern specimens in many regions. Additionally, most global change research  
1093 necessarily involves accessing many specimens across collections. Research efforts such as ours  
1094 will be unable to access every specimen from all possible collections. Ongoing digitization efforts  
1095 will make it possible to more clearly assess how much data is missing from a particular study  
1096 compared to the actual holdings of natural history collections, but ultimately, the decision of  
1097 what data and collections to include is a question of sample size and study design.