

<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, *Epichloë*, herbarium, INLA, museum specimen, plant-microbe sym-  
<sup>10</sup> biosis, *Poaceae*, spatially-varying coefficients model

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

## 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 promote 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) has 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 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 from across 63 populations, a rare extra step in collections-based research. 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 cryptic biological response to climate change that may contribute to the resilience of host-microbe symbiosis through fitness benefits to symbiotic hosts.

## Introduction

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

46 Natural history specimens, which were originally collected to study and preserve taxonomic di-  
47 versity, present a unique opportunity to explore long-term changes in ecological interactions across  
48 broad spatial and temporal scales (Meineke et al., 2018). Natural history collections, built and  
49 maintained by the efforts of thousands of scientists, are invaluable time machines, primarily com-  
50 prised of physical specimens of organisms along with information about the time and place of their  
51 collection. These specimens often preserve physical legacies of ecological processes and species' in-  
52 teractions from dynamically changing environments across time and space. For example, previous  
53 researchers have examined the flowers, pollen grains, and leaves of specimens within plant collec-  
54 tions (herbaria) collected across time to document shifts in reproductive phenology (Berg et al.,  
55 2019; Park et al., 2019; Willis et al., 2017), pollination (Duan et al., 2019; Pauw and Hawkins,  
56 2011), and herbivory (Meineke et al., 2019) related to anthropogenic climate change. However, few  
57 previous studies have leveraged biological collections to examine climate change-related shifts in a  
58 particularly common type of interaction: microbial symbiosis.

59 Microbial symbionts are common to all macroscopic organisms and can have important effects  
60 on their hosts' survival, growth and reproduction (McFall-Ngai et al., 2013; Rodriguez et al., 2009).  
61 Many microbial symbionts act as mutualists, engaging in reciprocally beneficial interactions 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 in ways that can cause 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 host's 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.

*Epichloë* fungal endophytes are specialized symbionts of cool-season grasses, which have been

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

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

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

## 131 Methods

### 132 *Focal species*

133 Our surveys focused on three native North American grasses: *Agrostis hyemalis*, *Agrostis perennans*,  
134 and *Elymus virginicus* that host *Epichloë* symbionts. These cool-season grass species are commonly  
135 represented in natural history collections with broad distributions covering much the eastern United  
136 States (Fig. 1). Cool-season grasses grow during the cooler temperatures of spring and autumn  
137 due to their reliance on  $C_3$  photosynthesis. *A. hyemalis* is a small short-lived perennial species  
138 that germinates in late winter winter and typically flowers between March and July (most common  
139 collection month: May). *A. perennans* is of similar stature but is longer lived than *Agrostis hyemalis*  
140 and flowers in late summer and early autumn (most common collection month: September). *A.*

<sup>141</sup> *perennans* is more sparsely distributed, tending to be found in shadier and moister habitats, while  
<sup>142</sup> *A. hyemalis* is commonly found in open and recently disturbed habitats. Both *Agrostis* species are  
<sup>143</sup> recorded from throughout the Eastern US, but *A. perennans* has a slightly more northern distri-  
<sup>144</sup> bution, whereas *A. hyemalis* is found rarely as far north as Canada and is listed as a rare plant  
<sup>145</sup> in Minnesota. *E. virginicus* is a larger and relatively longer-lived species that is more broadly dis-  
<sup>146</sup> tributed than the *Agrostis* species. It begins flowering as early as March or April but continues  
<sup>147</sup> throughout the summer (most common collection month: July).

<sup>148</sup> Both *Agrostis* species host *Epichloë amarillans* (Craven et al., 2001; Leuchtmann et al., 2014),  
<sup>149</sup> while *Elymus virginicus* typically hosts *Epichloë elymi* (Clay and Schardl, 2002). The fungal sym-  
<sup>150</sup> bionts primarily reproduce asexually and are passed from mother to offspring by vertical transmis-  
<sup>151</sup> sion through seeds. These traits contribute to highly specialized interactions between symbiont and  
<sup>152</sup> host. Some host species have been shown to partner with multiple symbiont species or strains, and in  
<sup>153</sup> some cases multiple symbiont lineages can co-exist within a host population (Mc Cargo et al., 2014).  
<sup>154</sup> However, surveys have typically found limited *Epichloë* genotypic diversity within host populations  
<sup>155</sup> (Treindl et al., 2023). Across host populations, concentrations of biologically-active alkaloids and  
<sup>156</sup> the genes associated with their production vary substantially (Schardl et al., 2012). In this analysis,  
<sup>157</sup> we focus on the presence/absence of *Epichloë* symbionts, and we discuss potential implications of  
<sup>158</sup> symbiont genotypic diversity in the Discussion.

### <sup>159</sup> *Herbarium surveys*

<sup>160</sup> We visited nine herbaria between 2019 and 2022 (see Table A1 for a summary of specimens included  
<sup>161</sup> from each collection). With permission from herbarium staff, we acquired seed samples from 1135  
<sup>162</sup> *A. hyemalis* specimens collected between 1824 and 2019, 357 *A. perennans* specimens collected  
<sup>163</sup> between 1863 and 2017, and 854 *E. virginicus* specimens collected between 1839 and 2019 (Fig. 1,  
<sup>164</sup> Fig. 2A, Fig. A1). We chose our focal species in part because they are commonly represented in  
<sup>165</sup> herbarium collections and produce many seeds, meaning that small samples would not diminish the  
<sup>166</sup> value of the specimens for future studies. We collected 5-10 seeds per specimen after examining the

167 herbarium sheet under a dissecting microscope to ensure that we collected mature seeds, not florets  
168 or unfilled seeds, fit for our purpose of identifying fungal endophytes with microscopy. We excluded  
169 specimens for which information about the collection location and date were unavailable.

170 Each specimen was assigned geographic coordinates based on collection information recorded on  
171 the herbarium sheet using the geocoding functionality of the ggmap R package (Kahle and Wickham,  
172 2019). Many specimens had digitized collection information readily available, but for those that did  
173 not, we transcribed information printed on the herbarium sheet. Collections were geo-referenced  
174 to the nearest county centroid, or nearest municipality when that information was available. For  
175 fifteen of the oldest specimens, only information at the state level was available, and so we used the  
176 state centroid. The median pairwise distance between georeferenced coordinate points was 841 km.  
177 The median longitudinal width of the bounding boxes generated to geocode municipality, county, or  
178 state centroids was 44.7 km. Among those specimens geo-referenced at the state level, the largest  
179 bounding box, spanning the state of Texas, was 1233 km wide. The smallest bounding boxes were  
180 less than 1 km across for small municipalities (while this suggests high precision, we note that some  
181 specimens were collected in natural habitat nearby to small municipalities not encompassed by this  
182 bounding boxes).

183 Our visits focused on herbaria with historic strengths in *Poaceae* collections (e.g. Texas A&M,  
184 Missouri Botanic Garden) and other herbaria in the Southern Great Plains region of the United  
185 States. While these nine herbaria garnered specimens that span the focal species' ranges, our dataset  
186 unevenly samples across the study region (Fig. 1). Texas, Oklahoma, Louisiana, and Missouri are  
187 the most represented states. Uneven sampling was most pronounced for *A. perennans*, which has  
188 much of its range in the northeastern US. We explore the potential influence of spatial bias in  
189 sampling on our results through a simulation analysis (Appendix A - Supporting Methods).

190 After collecting seed samples, we quantified the presence or absence of *Epichloë* fungal hyphae  
191 in each specimen using microscopy. We first softened seeds with a 10% NaOH solution, then stained  
192 the seeds with aniline blue-lactic acid stain and squashed them under a microscope cover slip. We  
193 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. 89% of scored plants had matching liberal and conservative scores, reflecting high confidence in endophyte status. The following analyses used the liberal status, however repeating all analyses with the conservative status 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 the spatial extent of each host’s distribution. To account for the spatial non-independence of geo-referenced 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

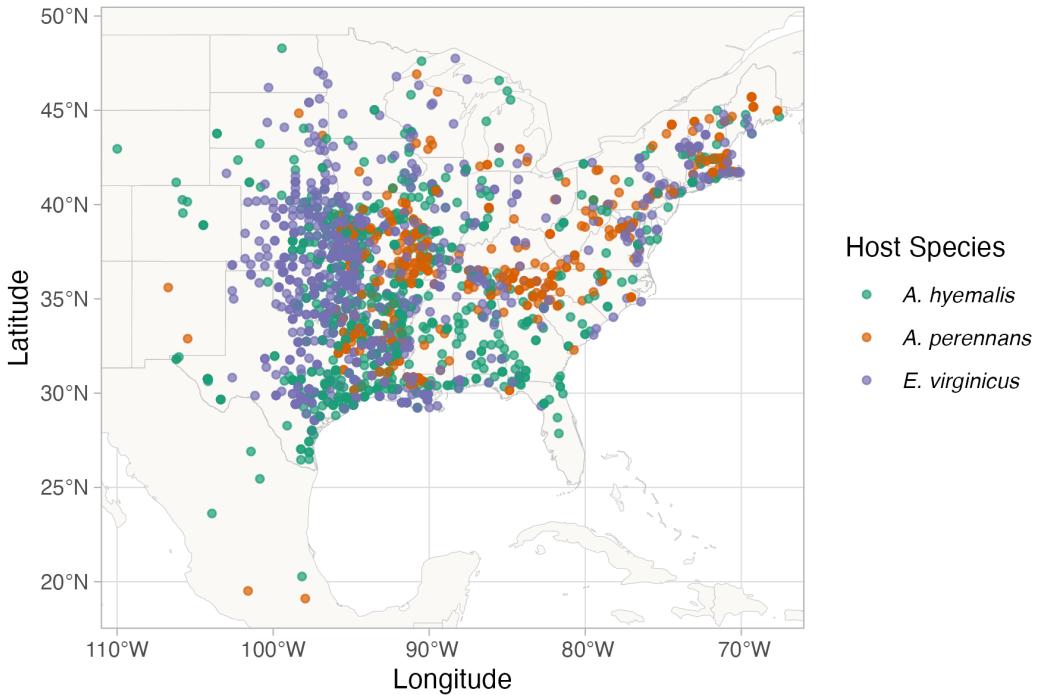


Figure 1: Collection locations of herbarium specimens sampled for *Epichloë* endophytes.

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.

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

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

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

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

246 We performed model fitting using the inlabru R package (Bachl et al., 2019). Global intercept  
 247 and slope parameters,  $A$  and  $T$ , were given vague priors. Collector and scorer random effects,  $\chi$  and  
 248  $\omega$  respectively, were centered at 0 with precision parameters assigned penalized complexity (PC)  
 249 priors with parameter values  $U_{PC} = 1$  and  $a_{PC} = 0.01$  (Simpson et al., 2017). Each spatially-  
 250 structured parameter depended on a covariance matrix according to the proximity of each pair  
 251 of collection locations (Bakka et al., 2018; Lindgren et al., 2011). The covariance matrix was  
 252 approximated using a Matérn covariance function, with each data point assigned a location according  
 253 to the nodes of a mesh of non-overlapping triangles encompassing the study area (Fig. A2). We  
 254 assessed model fit with visual posterior predictive checks (A3) and measurements of AUC (Figs.

255 A4-A5) (Gelman and Hill, 2006). Priors for the Matérn covariance function, termed "range" and  
256 "variance", define how proximity effects decay with distance. Results presented in the main text  
257 reflect a prior range of 342 kilometers (i.e. a 50% probability of estimating a range less than  
258 342 kilometers). We tested a range of values (from 68 kilometers to 1714 kilometers) and meshes  
259 (presented in the Supporting Methods), finding that while the magnitude and uncertainty of effects  
260 varied, model results were qualitatively similar, i.e. the same direction of effects across space.  
261 Through results and discussion that follow, we refer to the model described in this section as the  
262 "endophyte prevalence model".

263 *Modeling distributions of host species*

264 The herbarium records did not encompass the entirety of each host species' range. Therefore, we  
265 used additional data sources to model the geographic distribution of each host species, with two  
266 goals: (1) generate realistic maps on which we could project the predictions of the INLA model,  
267 and (2) use the geographic distributions to test for relationships between climate change drivers and  
268 trends in endophyte prevalence. We followed the ODMAP (overview, data, model, assessment,  
269 prediction) protocol (Crossley et al., 2022) (see Supporting Methods). In short, we used presence-  
270 only observations of each host species from Global Biodiversity Information Facility (GBIF) between  
271 1990 to 2020 (713 occurrence records for *A. hyemalis* (GBIF.Org, 2025a), 656 occurrence records  
272 for *A. perennans* (GBIF.Org, 2025b), and 2338 occurrence records for *A. virginicus* (GBIF.Org,  
273 2025c)). We fit maximum entropy (MaxEnt) models using the maxent function in the R package  
274 dismo (Hijmans et al., 2017) using following seasonal climate predictors (1990-2020 climate normals):  
275 mean and standard deviation of spring, summer, and autumn temperature, and mean and standard  
276 deviation of spring, summer, and autumn cumulative precipitation. We generated 10,000 pseudo-  
277 absences as background points, and split the occurrence data into 75% for model training and 25%  
278 for model testing. The performance of models was evaluated with AUC (Jiménez-Valverde, 2012).  
279 We found AUC values of 0.862, 0.838, 0.821 respectively for *Agrostis hyemalis*, *Agrostis perennans*,  
280 and *Elymus virginicus* indicating good model fit to data. We used the training sensitivity (true

281 positive rate) and specificity (true negative rate) to set a threshold for transforming the continuous  
282 predicted probabilities into binary presence - absence host distribution maps on which we projected  
283 INLA predictions of endophyte prevalence (Liu et al., 2005).

284 *Assessing the role of climate drivers*

285 We assessed how the magnitude of climate change may have driven changes in endophyte prevalence  
286 by assessing correlations between changes in climate and changes in endophyte prevalence predicted  
287 from our spatial model at evenly spaced pixels across the study area. We first downloaded monthly  
288 temperature and precipitation rasters from the PRISM climate group (Daly and Bryant, 2013) cov-  
289 ering the time period between 1895 and 2020 using the 'prism' R package (Hart and Bell, 2015).  
290 Prism provides reconstructions of historic climate variables across the United States by spatially  
291 interpolating weather station data (Di Luzio et al., 2008). We calculated 30-year climate normals for  
292 seasonal mean temperature and cumulative precipitation for the recent (1990 to 2020) and historic  
293 (1895 to 1925) periods. We used three four-month seasons within the year (Spring: January, Febru-  
294 ary, March, April; Summer: May, June, July, August; Autumn: September, October, November,  
295 December). This division of seasons allowed us to quantify differences in climate associated with  
296 the two "cool" seasons, when we expected our focal species to be most active (*A. hyemalis* flowering  
297 phenology: spring; *E. virginicus*: spring and summer; *A. perennans*: autumn). In addition to mean  
298 climate conditions, environmental variability itself can influence population dynamics (Tuljapurkar,  
299 1982) and changes in variability are a key prediction of climate change models (IPCC, 2021; Stocker  
300 et al., 2013). Therefore, we calculated the standard deviation for each annual and seasonal climate  
301 driver across each 30-year period. We then took the difference between recent and historic periods  
302 for the mean and standard deviation for each climate driver (Figs. A13-A15). All together, we  
303 assessed twelve potential climate drivers: the mean and standard deviation of spring, summer, and  
304 autumn temperature, as well as the mean and standard deviation of spring, summer, and autumn  
305 cumulative precipitation (the same climate covariates used in the MaxEnt models).

306 We then evaluated whether areas that have experienced the greatest changes in endophyte preva-

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

321 *Validating model performance with in-sample and out-of-sample tests*

322 We evaluated the predictive ability of the endophyte prevalence model using both in-sample train-  
323 ing data from the herbarium surveys, and with out-of-sample test data, an important but rarely  
324 used strategy in ecological studies (Lee et al., 2024; Tredennick et al., 2021). We generated out-of-  
325 sample test data from contemporary surveys of endophyte prevalence in natural populations of *A.*  
326 *hyemalis* and *E. virginicus* in Texas and the southern US. Surveys of *E. virginicus* were conducted  
327 in 2013 as described in Sneck et al. (2017), and surveys of *A. hyemalis* took place between 2015 and  
328 2020. Population surveys of *A. hyemalis* were initially designed to cover longitudinal variation in  
329 endophyte prevalence towards its range edge, while surveys of *E. virginicus* were designed to cover  
330 latitudinal variation. In total, we visited 43 populations of *A. hyemalis* and 20 populations of *E.*  
331 *virginicus* across the south-central US, with emphasis on Texas and neighboring states (Fig A12).  
332 During surveys, we collected seeds from up to 30 individuals per population (average number of

333 plants sampled per population: 22.9); note that this sampling design provided greater local depth  
334 of information than the herbarium records, where only one plant was sampled at each locality. We  
335 quantified the endophyte status of each individual with microscopy as described for the herbarium  
336 surveys (with 5-10 seeds scored per individual), and calculated the prevalence of endophytes within  
337 the population (proportion of plants that were endophyte-symbiotic). For each population, we com-  
338 pared the observed fraction of endophyte-symbiotic hosts to the predicted probability of endophyte  
339 occurrence  $\hat{P}$  derived from the model for that location and year. The contemporary survey period  
340 (2013-2020) is at the most recent edge of the time period encompassed by the historical specimens  
341 used for model fitting.

342

## Results

343 *How has endophyte prevalence changed over time?*

344 Across more than 2300 herbarium specimens dating back to 1824, we found that prevalence of  
345 *Epichloë* endophytes increased over the last two centuries for all three grass host species (Fig. 2).  
346 On average, endophytes of *A. perennans* and *E. virginicus* increased from ~ 40 % to 70% prevalence  
347 across the study region, and *A. hyemalis* increased from ~ 25% to over 50% prevalence. Our model  
348 indicates high confidence that overall temporal trends are positive across species (99% probability  
349 of a positive overall year slope in *A. hyemalis*, 92% probability of a positive overall year slope in *A.*  
350 *perennans*, and 91% probability of a positive overall year slope in *E. virginicus*) (Fig. A6).

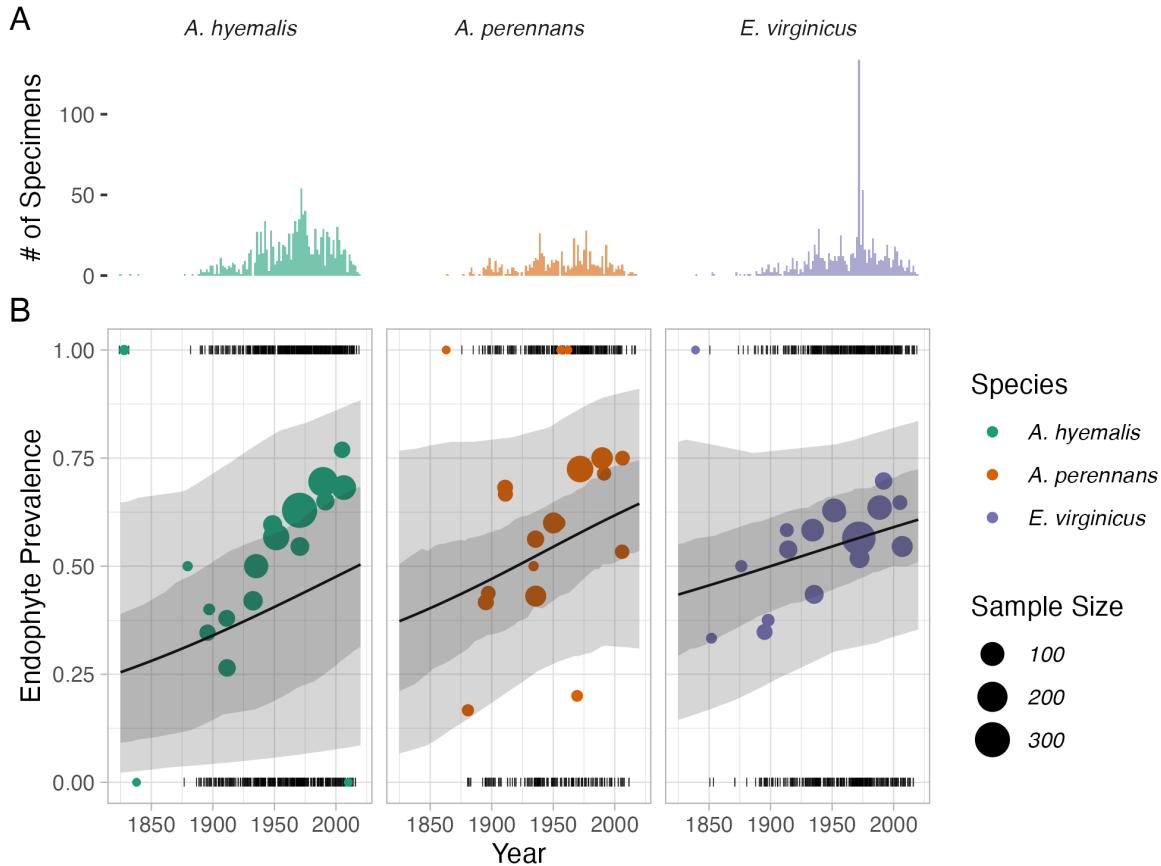


Figure 2: **Temporal trends in endophyte prevalence.** (A) Histograms show the frequency of scored specimens through time for each host species. (B) Lines show mean endophyte prevalence predicted by the endophyte prevalence model over the study period along with the 50% and 95% CI bands incorporating parameter uncertainty and variation associated with collector and scorer random effects. 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.

351 The model appears to under-predict the observed increase in endophyte prevalence relative to  
 352 the data, particularly for *A. hyemalis* (Fig. 2B), but the model is accounting for random effects  
 353 and spatial non-independence that are not readily seen in the figure. We found no evidence that  
 354 collector biases influenced our results. Collector random effects were consistently small (Fig. A9),

355 and models fit with and without this random effect provide qualitatively similar results. The identity  
356 of individual scorers, the researchers who identified endophyte status microscopically, did contribute  
357 to observed patterns in endophyte prevalence. For example, 3 of the 25 scorers were significantly  
358 more likely than average to assign positive endophyte status, as indicated by 95% credible intervals  
359 greater than zero, while 4 of the 25 had 95% credible intervals below zero (Fig. A10).

360 *How spatially variable are temporal trends in endophyte prevalence?*

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

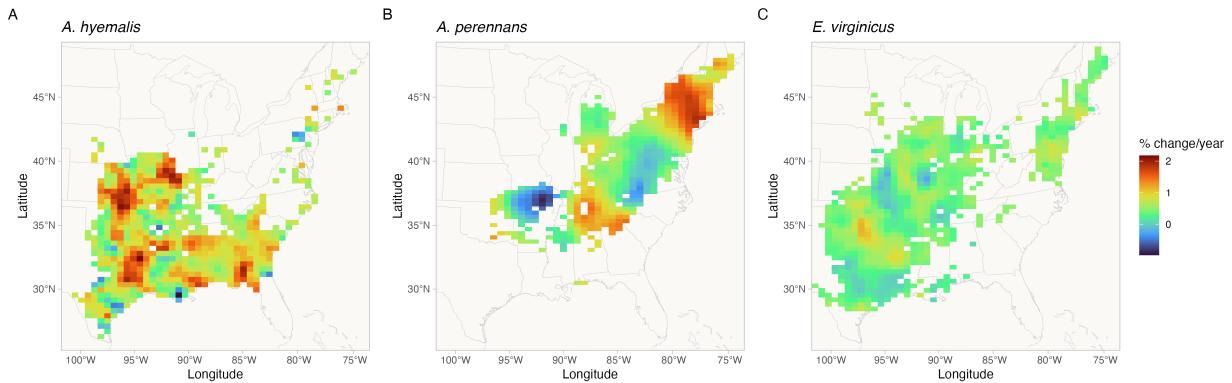


Figure 3: Predicted posterior mean of spatially-varying slopes representing change in endophyte prevalence for each host species (A, *A. hyemalis*; B, *A. perennans*; C, *E. virginicus*). Spatially-varying trends are estimated from the endophyte prevalence model. Color indicates the relative change in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

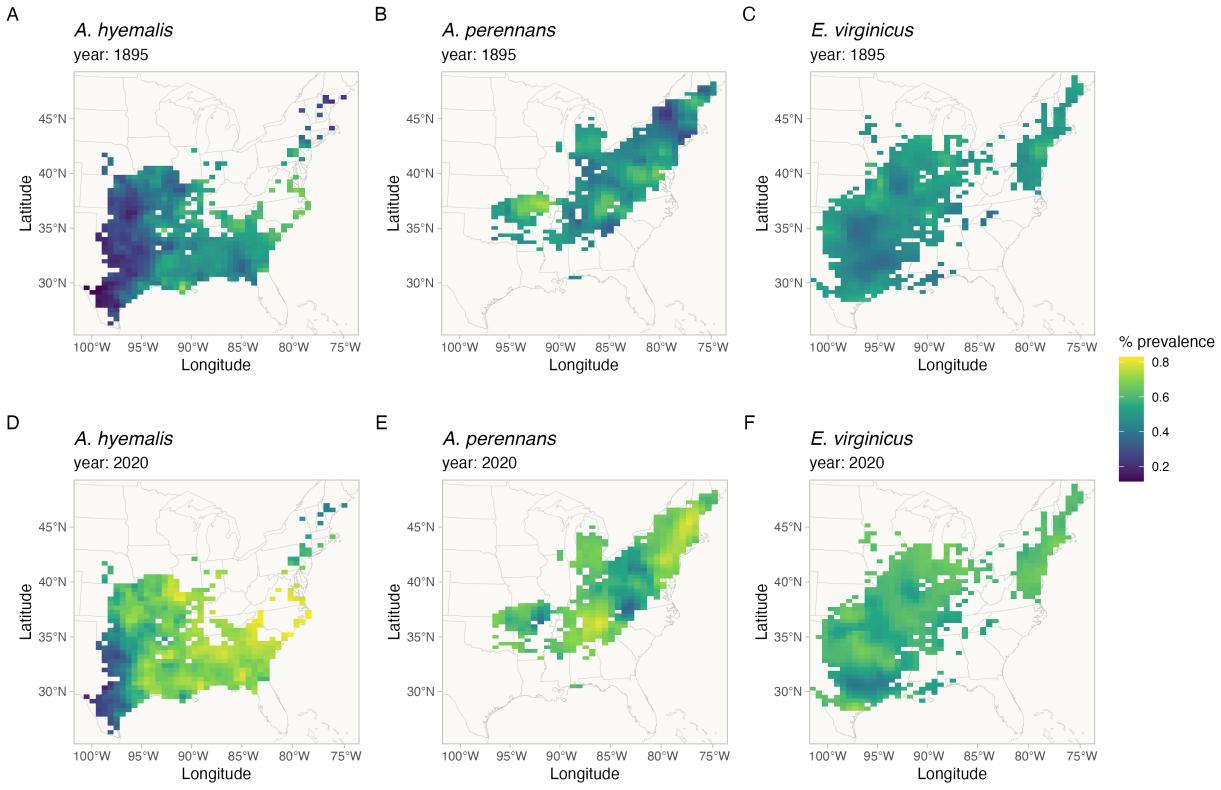


Figure 4: Predicted endophyte prevalence for each host species in 1895 and 2020. Predictions of prevalence come from the endophyte prevalence model. Color indicates the posterior mean endophyte prevalence for *A. hyemalis* (A,D), *A. perennans* (B,E), and *E. virginicus* (C,F). Map lines delineate study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

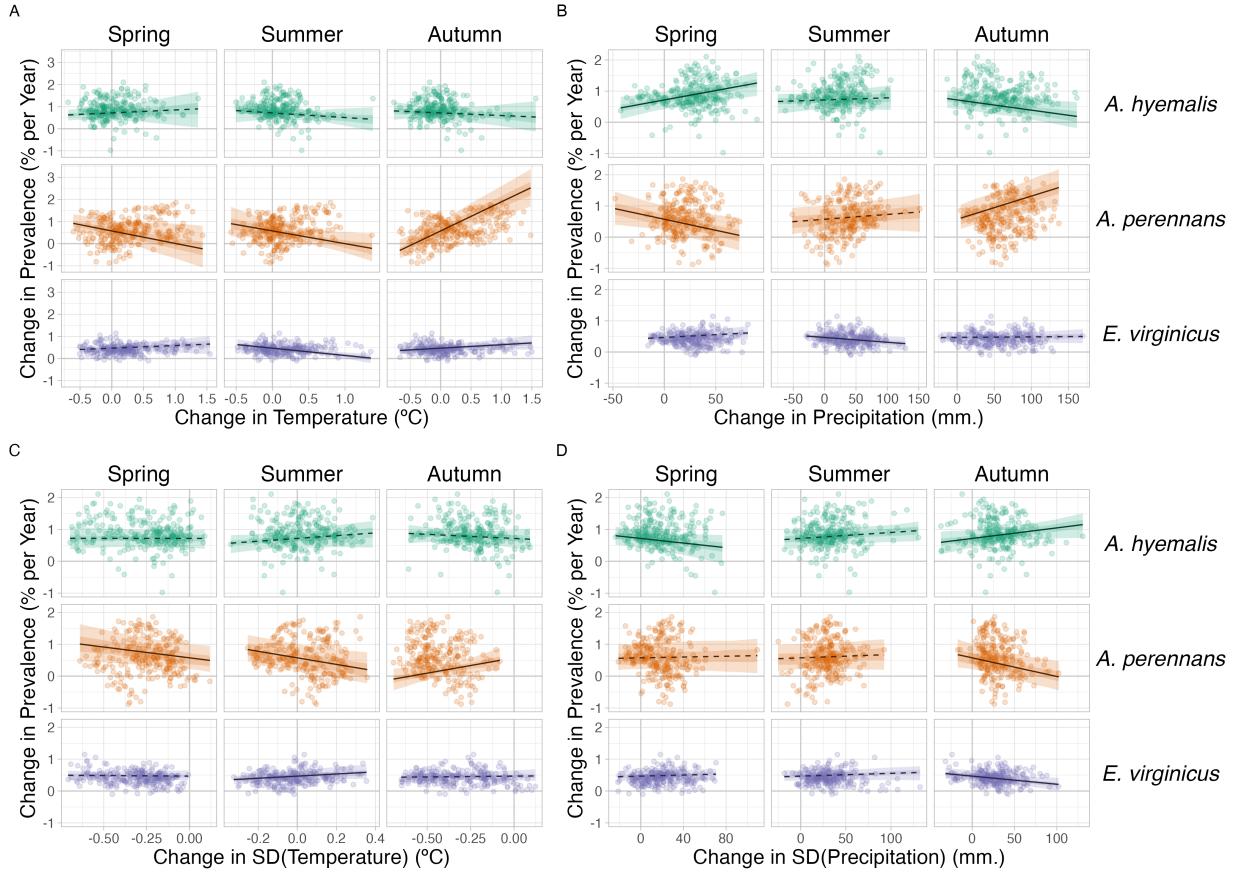
379     What is the relationship between variation in temporal trends in endophyte  
 380                          prevalence and changes in climate drivers?

381     We found that trends in endophyte prevalence were strongly associated with one or more seasonal  
 382     climate change drivers (Fig. 5). For the majority of the study region, the climate has become wetter  
 383     (an average increase in annual precipitation of 60 mm) with relatively little temperature warming  
 384     (an average increase in annual temperature of 0.02 °C) over the last century (Fig. A13-A15), a  
 385     consequence of regional variation in global climate change (IPCC, 2021). Within the region, climate

386 changes were spatially variable; certain locations experienced increases in annual precipitation as  
387 large as 375 mm or decreases up to 54 mm across the last century, while annual temperature changes  
388 ranged from warming as great as 1.4 °C to cooling by 0.46 °C. Spatially variable climate trends were  
389 predictive of trends in endophyte prevalence. For example, strong increases in endophyte prevalence  
390 for *A. perennans* were most strongly associated with increasing autumn precipitation and with  
391 increasing mean and variability in autumn temperature (greater than 97% posterior probabilities  
392 of positive slopes). For this species, each 1 °C increase in autumn temperature was associated with  
393 a 1.07 % greater increase per year in endophyte prevalence (Fig. 5A) and a 100 mm increase in  
394 precipitation was associated with a 0.8% greater increase per year in endophyte prevalence (Fig.  
395 5B). This result aligns with the species' autumn active growing season, however other seasonal  
396 climate drivers were also associated with increasing endophyte prevalence in this host species. In  
397 particular, we found cooler and drier springs and cooler summers to be associated with increasing  
398 endophyte prevalence (greater than 99% posterior probabilities of negative slopes), though these  
399 slopes were generally of smaller magnitude than those for autumn climate drivers.

400 Changes in endophyte prevalence across the ranges of *A. hyemalis* and *E. virginicus* were less  
401 strongly driven by changes in climate. Like *A. perennans*, climate during peak growing season  
402 (spring for *A. perennans* and summer for *E. virginicus*) emerged most commonly as drivers of  
403 changes in endophyte prevalence. Increases in mean spring precipitation were the strongest predictor  
404 of increasing trends in endophyte prevalence for *A. hyemalis* (Fig. 5B) (greater than 99% posterior  
405 probability of a positive slope). For this species, an increase of 100 mm in spring precipitation was  
406 associated with 0.6% per year stronger increases in endophyte prevalence relative to regions with  
407 no change in precipitation. The next greatest slopes were those associated with variability in spring  
408 precipitation (greater than 96% posterior probability of a negative slope), as well as in the mean  
409 and variability of autumn climate (greater than 98% probability of negative and positive slopes,  
410 respectively). Changes in endophyte prevalence in *E. virginicus* were not strongly associated with  
411 changes in most climate drivers, but regions with reduced variability in autumn precipitation (Fig.  
412 5B) and with cooler and more variable summer temperatures (Fig. 5A,C) experienced the largest

413 increases in endophyte prevalence. Our analysis indicated relatively high confidence that these  
414 climate drivers influence endophyte prevalence shifts in *E. virginicus*(greater than 99% posterior  
415 probability of either negative or positive slopes respectively), however they translate, for example,  
416 to less than a 0.4% decrease in endophyte prevalence per year for each 1°C of summer warming  
417 over the century. Repeating this analysis using all pixels across each species' distribution were  
418 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% posterior 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.

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

420      Tests of the endophyte prevalence model's predictive performance as quantified by AUC and by  
421      visual posterior predictive checks, indicated good predictive ability. Model performance was similar  
422      between historic herbarium specimens used as training data and the out-of-sample test data from  
423      contemporary surveys (AUC = 0.79 and 0.77 respectively; Fig. A5-A4). The model successfully  
424      captured broad regional trends in endophyte prevalence seen in the contemporary survey data,  
425      such as decline endophyte prevalence in *A. hyemalis* towards western longitudes (Fig. 6A) and an  
426      increase towards northern latitudes (Fig. 6B). It is noteable that model predictions for endophyte  
427      prevalence exhibited relatively little local geographic variation, whereas the out-of-sample survey  
428      data were highly variable with populations spanning 0% to 100% endophyte-symbiotic plants (Fig.  
429      6C), indicating population-to-population variation not captured in the endophyte prevalence model.

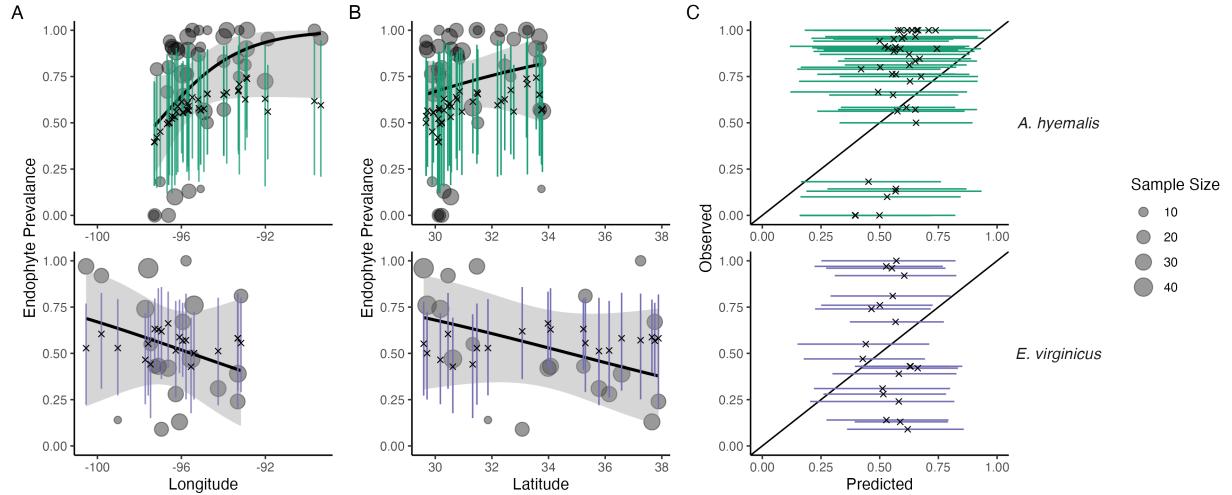


Figure 6: **Predictive performance for contemporary test data.** (A) The endophyte prevalence model, trained on historic herbarium collection data, performed modestly at predicting prevalence in contemporary population surveys. The model captured regional trends across (A) longitude and (B) latitude. 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 contemporary observed population prevalence vs. predicted endophyte prevalence shows that contemporary test data had more variance between populations than in model predictions.

## Discussion

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

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

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

443 Differences across host species underscore that while all of these  $C_3$  grasses share similar broad-scale  
444 distributions, each engages in unique biotic interactions and has unique responses to environmental  
445 drivers. We identified hotspots of change for *A. perennans*, which was the species whose endophyte  
446 prevalence was most responsive to changes in climate drivers (Fig. 5). Predicted declines of 0.9%  
447 per year in the southern portion of its range and predicted increases of up to 2% per year in the  
448 north suggest a potential poleward range shift of endophyte-symbiotic plants (Fig. 3B); whether  
449 the overall host distribution is shifting in parallel is an exciting next question.

450 Based on previous work demonstrating that endophytes can shield their hosts from drought  
451 stress (reviewed in Decunta et al. (2021)), we generally predicted that drought conditions would be  
452 a driver of increasing endophyte prevalence. In contrast to this expectation, increasing prevalence  
453 for *A. perennans* was associated with both increasing autumn temperature and precipitation (Fig.  
454 5). To our knowledge, the response of the symbiosis in *A. perennans* to drought has not been  
455 examined experimentally, but in a greenhouse experiment, endophytes had a positive effect on host  
456 reproduction under shaded, low-light conditions (Davitt et al., 2010). Our results also hint that it  
457 may be useful to investigate whether lagged climate effects are important predictors of host fitness  
458 in this system (Evers et al., 2021). Endophyte prevalence of the autumn-flowering *A. perennans* was  
459 strongly linked with decreasing spring precipitation, and that of the spring-flowering *A. hyemalis*  
460 was associated with decreasing autumn precipitation (Fig. 5B). For *A. hyemalis*, endophytes could  
461 be playing a role helping hosts weather autumn-season droughts, which is likely also an important  
462 time for the species' germination. Previous work demonstrated drought benefits in a greenhouse

manipulation with this host-symbiont pair (Davitt et al., 2011), and early life stages may be particularly vulnerable to prolonged droughts. For *E. virginicus*, which experienced the weakest changes in endophyte prevalence overall (ranging between 1.1% increases and 0.2% decreases), we only found modest associations with changes in climate drivers. Surveys by Sneck et al. (2017), used as part of the test data in this study, identified a drought index (SPEI) that integrates precipitation with estimated evapotranspiration as an important predictor of contemporary endophyte prevalence in this species. The diverse relationships we detect between trends in endophyte prevalence and climate drivers suggest a more complicated picture than the simple explanation that drought alone, as measured through changes in annual precipitation, causes increasing endophyte prevalence through context-dependent fitness benefits.

While we show consistent increasing trends in prevalence between the three species, the mechanisms that explain these changes may be diverse and idiosyncratic. First, climate change responses may depend on genotype-specific responses that are not considered in our current analysis. While *Epichloë* symbioses are highly specialized, surveys have demonstrated genotypic and chemotypic diversity of the symbionts among and within populations (Treindl et al., 2023; von Cräutlein et al., 2021). Genotypic variation in *Epichloë* endophytes, particularly in genes responsible for alkaloid production, produces "chemotypes" with differing benefits for hosts against insect or mammalian herbivores mediated by environmental conditions (Saikkonen et al., 2013; Schardl et al., 2012). Genotypic variation of the hosts themselves can also influence interaction outcomes (Gundel et al., 2011; Parker et al., 2017). Whether hotspots of change in endophyte prevalence reflect selection for genotype-pairings with particularly strong fitness benefits is an unanswered question. Additionally, *Epichloë* endophytes have been connected to a suite of non-drought related fitness benefits including herbivory defense (Brem and Leuchtmann, 2001), salinity resistance (Wang et al., 2020), and mediation of pathogens (Vikuk et al., 2019) and the soil microbiome (Roberts and Ferraro, 2015). Broad changes in the distribution and abundance of natural enemies (Côté et al., 2004), along with stresses from anthropogenic changes in landcover and pollution (Sage, 2020) likely influence the benefits of symbiosis (Rudgers et al., 2020). The historic trends that we observed result from the combination

490 of these fitness benefits playing out across the heterogeneous map of a changing climate and other  
491 anthropogenic drivers.

492 Our results indicate that *Epichloë* symbiosis has likely improved host fitness in stressful envi-  
493 ronments leading to increasing prevalence. What is less clear is how this will influence future range  
494 shifts. Based on our analysis, it is likely that the symbiosis will facilitate range shifts for hosts by  
495 improving population growth at range edges. Previous population surveys (Rudgers and Swafford,  
496 2009; Semmarin et al., 2015; Sneck et al., 2017) attributed environment-dependent gradients in  
497 endophyte prevalence to symbiont-derived fitness benefit's allowing hosts to persist in environments  
498 where they otherwise could not (Afkhami et al., 2014; Kazenel et al., 2015). However, symbiont-  
499 facilitated range shifts require that endophytes be present in the populations to be able to support  
500 population growth. The arid western range edge of *A. hyemalis* has had historically low endophyte  
501 prevalence (Fig. 4), and while prevalence has increased most quickly in the regions with histor-  
502 ically low endophyte prevalence (Fig. A11), the complete absence of endophytes at range edges  
503 in this species would make it impossible for prevalence to increase without dispersal of symbiotic  
504 seeds (Fowler et al., 2023). These factors potentially contribute to the ability of the host species  
505 to track its environmental niche. Another interesting question is the degree to which symbiotic  
506 and non-symbiotic hosts, which occupy overlapping but distinct niches, are likely to experience  
507 distribution shifts in tandem or at different rates in future. More extreme climate stresses, which  
508 are expected more frequently in the future (Seneviratne et al., 2021), have the potential to alter  
509 the costs and benefits of the interaction. The past indicates a resilient symbiosis, but it will be  
510 crucial to understand whether more extreme future climate conditions could tip this interaction to  
511 deteriorate.

512 *Steps towards forecasts of host-microbe symbioses*

513 The combination of a spatially-explicit model and historic herbarium specimens allowed us to  
514 identify regions of both increasing and decreasing endophyte prevalence, however we see several  
515 next steps toward the goal of predicting host and symbiont niche-shifts in response to future cli-

516 mate change. While the model recreated the large-scale spatial trends observed in contemporary  
517 population surveys, test data contained more population-to-population variability in prevalence than  
518 could be explained by the model. We interpret this to mean that the model captures coarse-scale  
519 spatial and temporal trends reasonably well, but is not equipped to capture local-scale nuances that  
520 generate population-to-population differences. Validating our model predictions with this test, a  
521 rare extra step in collections-based studies, allows us to identify ways in which the model's out-of-  
522 sample predictive ability could be improved. Lack of information on local variability in symbiont  
523 prevalence may simply be a feature of data derived from herbarium specimens. Natural history  
524 collectors sample one or a few specimens from local populations, and these observations are aggre-  
525 gated by the model to derive broad-scale estimates. This suggests that increasing local replication  
526 should be a factor considered in future collection efforts of natural history specimens, balancing the  
527 required time and effort along with limitations on storage space within collections. Poor predictive  
528 ability at local scales in this grass-endophyte system is not surprising, as previous studies have found  
529 that local variation, even to the scale of hundreds of meters can structure endophyte-host niches  
530 (Gundel et al., 2024; Kazenel et al., 2015). An important step would be integrating data from local  
531 and regional scales through modeling to constrain estimates of local and regional variation.

532 Predicting future niche-shifts of hosts and symbionts will require considering the coupled dynam-  
533 ics of host-symbiont dispersal in addition to fitness benefits. For example, transplanting symbiotic  
534 and non-symbiotic plants beyond the range edge of *A. hyemalis* could tell us whether low endophyte  
535 prevalence in that area (Fig. 4A) is a result of environmental conditions that lead the symbiosis  
536 to have negative fitness consequences, or is a result of some historical contingency or dispersal lim-  
537 itation that has thus far limited the presence of symbiotic hosts from a region where they would  
538 otherwise flourish and provide resilience. Incorporating available climatic and soil layers as covari-  
539 ates is another obvious step that could improve predictions. These steps will bridge gaps that often  
540 exist between large but broad bioclimatic and biodiversity data and small but high-resolution data  
541 on biotic interactions, and move towards the goal of predicting the dynamics of microbial symbioses  
542 under climate change (Isaac et al., 2020; Miller et al., 2019).

543 *Herbaria for global change research*

544 Our analysis advances the use of herbarium specimens in global change biology in two ways.  
545 First and foremost, this is one of a growing number of studies to examine microbial symbiosis using  
546 specimens from natural history collections, and the first, to our knowledge, to link long-term changes  
547 in the symbioses to changes in climate. The responses of microbial symbioses are a rich target for  
548 future studies within historic specimens, particularly those that take advantage of advances in  
549 sequencing technology. While we used relatively coarse presence/absence data based on fungal  
550 morphology, other studies have examined historic plant microbiomes using molecular sequencing  
551 and sophisticated bioinformatics techniques, but these studies have so far been limited to relatively  
552 few specimens at limited spatial extents (Bieker et al., 2020; Bradshaw et al., 2021; Gross et al.,  
553 2021; Heberling and Burke, 2019; Yoshida et al., 2015). Continued advances in capturing historic  
554 DNA and in filtering out potential contamination during specimen storage (Bakker et al., 2020;  
555 Daru et al., 2019; Raxworthy and Smith, 2021) will be imperative in the effort to scale up these  
556 efforts. This scaling up will be essential to be able to quantify changes not just in the prevalence  
557 of symbionts, but also in symbionts' intraspecific variation and evolutionary responses to climate  
558 change, as well as in changes in the wider host microbiome. With improved molecular insights from  
559 historic specimens, we could ask whether the broad increases in endophytes that we have identified  
560 reflect selection for particular genetic strains or chemotypes and how this selection varies across  
561 space. Answering these questions as well as the unknown questions that future researchers may  
562 ask also reiterates the value in capturing meta-information during ongoing digitization efforts at  
563 herbaria around the world and during the accession of newly collected specimens (Edwards et al.;  
564 Lendemer et al., 2020).

565 The second major advance in this analysis is in accounting for several potential biases in the data  
566 observation process that may be common to many collections-based research questions by using a  
567 spatially-explicit random effects model. Potential biases introduced by the sampling habits of col-  
568 lectors (Daru et al., 2018), and variation between contemporary researchers during the collection of

569 trait data, if not corrected for could lead to over-confident inference about the strength and direction  
570 of historic change (Fig. 2). Previous studies that have quantified the effects of collector biases typ-  
571 ically find them to be small (Davis et al., 2015; Meineke et al., 2019), and we similarly did not find  
572 that collector has a strong effect on the results of our analysis, but that scorer identity did impact  
573 results. It is difficult to distinguish whether the impact of scorers was driven by true differences  
574 in scorers' biases or by unintended spatial or temporal clustering of the specimens examined by  
575 each scorer (Clayton et al., 1993; Urdangarin et al., 2023). By under-weighting endophyte-positive  
576 samples that are clustered spatially or by collector or observer, the endophyte prevalence model is  
577 appropriately accounting for nuisance variables and providing a conservative inference of endophyte  
578 change relative to the raw data. Spatial autocorrelation is another phenomenon likely common in  
579 data derived from herbarium specimens (Willems et al., 2022), which our spatially-explicit analysis  
580 models among samples. Beyond spatial autocorrelation of outcomes, systematic differences in sam-  
581 pling across space can result in spatial bias. One strength of herbaria as vehicles for global change  
582 research is the relative ease with which specimens from many distinct geographic locations can be  
583 examined. We visited just nine institutions in the central southern United States, and we were able  
584 to sample seeds from specimens across an area spanning over 300,000 sq. km, including specimens  
585 from Mexico and Canada. Despite this advantage, the specimens we examined are concentrated in  
586 the south-central United States, with fewer specimens in the rapidly warming northeastern United  
587 States reflecting the regional focus of herbaria. We provide a simulation analysis exploring the  
588 potential impact of spatially and temporally biased sampling (Appendix A - Supporting Methods).  
589 We found that the spatially-varying coefficient model had a strong ability to re-capitulate temporal  
590 trends across space in simulated data, and that this result was robust to relatively high levels of  
591 spatial bias (80% of data missing from one spatial region). Simulation analyses that extend this  
592 work to consider the myriad ways herbarium data may be biased (i.e. testing different spatial ar-  
593 rangements and scales of spatial bias, or testing different sample sizes) would be extremely valuable  
594 (Daru et al., 2018; Erickson and Smith, 2021; Gaul et al., 2020; Meineke and Daru, 2021).

595 *Conclusion*

596 Ultimately, a central goal of global change biology is to generate predictive insights into the  
597 future of natural systems on a rapidly changing planet. Beyond host-microbe symbioses, detect-  
598 ing ecological responses to anthropogenic global change and attributing their causes would inform  
599 public policy decision-makers and adaptive management strategies. This survey of historic endo-  
600 phyte prevalence is necessarily correlative, yet it serves as a foundation to develop better predictive  
601 models of the response of microbial symbioses to climate change. By comparing detected ecologi-  
602 cal responses with alternative mechanistic simulations of the past, we could attribute their cause,  
603 in a manner similar to methods from climate science and economics (Carleton and Hsiang, 2016;  
604 Stott et al., 2010; Trenberth et al., 2015). Combining the insights from this type of regional-scale  
605 survey with field experiments and physiological performance data could be invaluable to identify  
606 mechanisms driving shifts in host-symbiont dynamics. Evidence is strong that certain dimensions  
607 of climate change correlated with endophytes' temporal responses, however we do not know why  
608 trends in prevalence were weak in some areas or how endophytes would respond to more extreme  
609 changes in climate. The "time machine" of natural history collections revealed evidence of mutual-  
610 ism resilience for grass-endophyte symbioses in the face of environmental change, but more extreme  
611 changes could potentially push one or both partners beyond their physiological limits, leading to  
612 the collapse of the mutualism; more research is needed to understand what those limits might be.

613 **Acknowledgments**

614 We thank Dr. Jessica Budke for help in drafting our initial destructive sampling plan, and to the  
615 many staff members of herbaria who facilitated our research visits, as well as to the hundreds of  
616 collectors who contributed to the natural history collections. Several high school and undergraduate  
617 researchers contributed to data collection, including A. Appio-Riley, P. Bilderback, E. Chong, K.  
618 Dickens, L. Dufresne, B. Gutierrez, A. Johnson, S. Linder, E. Scales, B. Scherick, K. Schrader, E.  
619 Segal , G. Singla, and M. Tucker. This research was supported by funding from National Science  
620 Foundation (grants 1754468 and 2208857) and by funding from the Texas Ecolab Program. Two

621 anonymous reviewers greatly improved an earlier version of this manuscript.

622 **Statement of Authorship**

623 J.C.F. contributed to research conception, data collection, data analysis, and led manuscript draft-  
624 ing. J.M. contributed to data analysis and manuscript revisions. T.E.X.M. contributed to research  
625 conception, data collection, data analysis, and manuscript revisions.

626 **Data and Code Availability**

627 Data from this publication can be found through a publicly available repository  
628 (<https://doi.org/10.5061/dryad.rn8pk0pn0>). Code for analyses can be found through a publicly  
629 available repository (<https://github.com/joshuacfowler/EndoHerbarium>) that will be permanently  
630 archived upon publication.

631

## Literature Cited

- 632 M E Afkhami. Fungal endophyte–grass symbioses are rare in the California floristic province and  
633 other regions with mediterranean-influenced climates. *Fungal Ecology*, 5(3):345–352, 2012.
- 634 M E Afkhami and J A Rudgers. Symbiosis lost: Imperfect vertical transmission of fungal endophytes  
635 in grasses. *The American Naturalist*, 172(3):405–416, 2008.
- 636 M E Afkhami, P J McIntyre, and S Y Strauss. Mutualist-mediated effects on species’ range limits  
637 across large geographic scales. *Ecology Letters*, 17(10):1265–1273, 2014.
- 638 S N Aitken, S Yeaman, J A Holliday, T Wang, and S Curtis-McLane. Adaptation, migration  
639 or extirpation: Climate change outcomes for tree populations. *Evolutionary Applications*, 1(1):  
640 95–111, 2008.
- 641 C E Aslan, E S Zavaleta, B Tershy, and D Croll. Mutualism disruption threatens global plant  
642 biodiversity: a systematic review. *PloS one*, 8(6):e66993, 2013.
- 643 F E Bachl, F Lindgren, D L Borchers, and J B Illian. inlabru: an R package for Bayesian spatial  
644 modelling from ecological survey data. *Methods in Ecology and Evolution*, 10(6):760–766, 2019.
- 645 C W Bacon and J F White. Stains, media, and procedures for analyzing endophytes. In *Biotech-*  
646 *nology of endophytic fungi of grasses*, pages 47–56. CRC Press, 2018.
- 647 H Bakka, H Rue, G-A Fuglstad, A Riebler, D Bolin, J Illian, E Krainski, D Simpson, and F Lind-  
648 gren. Spatial modeling with r-inla: A review. *Wiley Interdisciplinary Reviews: Computational*  
649 *Statistics*, 10(6):e1443, 2018.
- 650 F T Bakker, V C Bieker, and M D Martin. Herbarium collection-based plant evolutionary genetics  
651 and genomics. *Frontiers in Ecology and Evolution*, 8:603948, 2020.
- 652 D R Bazely, J P Ball, M Vicari, A J Tanentzap, M Bérenger, T Rakocevic, and S Koh. Broad-

- 653 scale geographic patterns in the distribution of vertically-transmitted, asexual endophytes in four  
654 naturally-occurring grasses in sweden. *Ecography*, 30(3):367–374, 2007.
- 655 J Beguin, S Martino, H Rue, and S G Cumming. Hierarchical analysis of spatially autocorrelated  
656 ecological data using integrated nested laplace approximation. *Methods in Ecology and Evolution*,  
657 3(5):921–929, 2012.
- 658 C S Berg, J L Brown, and J J Weber. An examination of climate-driven flowering-time shifts at  
659 large spatial scales over 153 years in a common weedy annual. *American Journal of Botany*, 106  
660 (11):1435–1443, 2019.
- 661 V C Bieker, F Sánchez Barreiro, J A Rasmussen, M Brunier, N Wales, and M D Martin. Metage-  
662 nomic analysis of historical herbarium specimens reveals a postmortem microbial community.  
663 *Molecular Ecology Resources*, 20(5):1206–1219, 2020.
- 664 J L Blois, P L Zarnetske, M C Fitzpatrick, and S Finnegan. Climate change and the past, present,  
665 and future of biotic interactions. *Science*, 341(6145):499–504, 2013.
- 666 M Bradshaw, U Braun, M Elliott, J Kruse, S-Y Liu, G Guan, and P Tobin. A global genetic analysis  
667 of herbarium specimens reveals the invasion dynamics of an introduced plant pathogen. *Fungal*  
668 *Biology*, 125(8):585–595, 2021.
- 669 D Brem and A Leuchtmann. *Epichloë* grass endophytes increase herbivore resistance in the woodland  
670 grass *Brachypodium sylvaticum*. *Oecologia*, 126(4):522–530, 2001.
- 671 T A Carleton and S M Hsiang. Social and economic impacts of climate. *Science*, 353(6304):aad9837,  
672 2016.
- 673 S Cheng, Y-N Zou, K Kuča, A Hashem, E F Abd\_Allah, and Q-S Wu. Elucidating the mecha-  
674 nisms underlying enhanced drought tolerance in plants mediated by arbuscular mycorrhizal fungi.  
675 *Frontiers in Microbiology*, 12:4029, 2021.

- 676 K Clay and C Schardl. Evolutionary origins and ecological consequences of endophyte symbiosis  
677 with grasses. *The American Naturalist*, 160(S4):S99–S127, 2002.
- 678 D G Clayton, L Bernardinelli, and C Montomoli. Spatial correlation in ecological analysis. *Inter-*  
679 *national Journal of Epidemiology*, 22(6):1193–1202, 1993.
- 680 S D Côté, T P Rooney, J-P Tremblay, C Dussault, and D M Waller. Ecological impacts of deer  
681 overabundance. *Annu. Rev. Ecol. Evol. Syst.*, 35(1):113–147, 2004.
- 682 KD Craven, PTW Hsiao, A Leuchtmann, W Hollin, and CL Schardl. Multigene phylogeny of  
683 *Epichloë* species, fungal symbionts of grasses. *Annals of the Missouri Botanical Garden*, pages  
684 14–34, 2001.
- 685 K M Crawford, J M Land, and J A Rudgers. Fungal endophytes of native grasses decrease insect  
686 herbivore preference and performance. *Oecologia*, 164:431–444, 2010.
- 687 M S Crossley, T D Meehan, M D Moran, J Glassberg, W E Snyder, and A K Davis. Opposing  
688 global change drivers counterbalance trends in breeding North American monarch butterflies.  
689 *Global Change Biology*, 28(15):4726–4735, 2022.
- 690 C Daly and K Bryant. The PRISM climate and weather system — An introduction. *Corvallis, OR:*  
691 *PRISM climate group*, 2, 2013.
- 692 B H Daru, D S Park, R B Primack, C G Willis, D S Barrington, T J S Whitfeld, T G Seidler, P W  
693 Sweeney, D R Foster, A M Ellison, et al. Widespread sampling biases in herbaria revealed from  
694 large-scale digitization. *New Phytologist*, 217(2):939–955, 2018.
- 695 B H Daru, E A Bowman, D H Pfister, and A E Arnold. A novel proof of concept for capturing  
696 the diversity of endophytic fungi preserved in herbarium specimens. *Philosophical Transactions*  
697 *of the Royal Society B*, 374(1763):20170395, 2019.
- 698 C C Davis, C G Willis, B Connolly, C Kelly, and A M Ellison. Herbarium records are reliable sources

- 699 of phenological change driven by climate and provide novel insights into species' phenological  
700 cueing mechanisms. *American Journal of Botany*, 102(10):1599–1609, 2015.
- 701 A J Davitt, M Stansberry, and H A Rudgers. Do the costs and benefits of fungal endophyte symbiosis  
702 vary with light availability? *New Phytologist*, 188(3):824–834, 2010.
- 703 A J Davitt, C Chen, and J A Rudgers. Understanding context-dependency in plant–microbe sym-  
704 biosis: The influence of abiotic and biotic contexts on host fitness and the rate of symbiont  
705 transmission. *Environmental and Experimental Botany*, 71(2):137–145, 2011.
- 706 F A Decunta, L I Pérez, D P Malinowski, M A Molina-Montenegro, and P E Gundel. A systematic  
707 review on the effects of *Epichloë* fungal endophytes on drought tolerance in cool-season grasses.  
708 *Frontiers in Plant Science*, 12:644731, 2021.
- 709 M Di Luzio, G L Johnson, C Daly, J K Eischeid, and J G Arnold. Constructing retrospective  
710 gridded daily precipitation and temperature datasets for the conterminous United States. *Journal*  
711 *of Applied Meteorology and Climatology*, 47(2):475–497, 2008.
- 712 M L Donald, T F Bohner, K M Kolis, R A Shadow, J A Rudgers, and TEX Miller. Context-  
713 dependent variability in the population prevalence and individual fitness effects of plant–fungal  
714 symbiosis. *Journal of Ecology*, 109(2):847–859, 2021.
- 715 AE Douglas. Host benefit and the evolution of specialization in symbiosis. *Heredity*, 81(6):599–603,  
716 1998.
- 717 Y-W Duan, H Ren, T Li, L-L Wang, Z-Q Zhang, Y-L Tu, and Y-P Yang. A century of pollination  
718 success revealed by herbarium specimens of seed pods. *New Phytologist*, 224(4):1512–1517, 2019.
- 719 E J Edwards, B D Mishler, and C D Davis. University herbaria are uniquely important. *Trends in*  
720 *Plant Science*.
- 721 M Engel, T Mette, and W Falk. Spatial species distribution models: Using Bayes inference with

- 722 INLA and SPDE to improve the tree species choice for important European tree species. *Forest*  
723 *Ecology and Management*, 507:119983, 2022.
- 724 K D Erickson and A B Smith. Accounting for imperfect detection in data from museums and herbaria  
725 when modeling species distributions: combining and contrasting data-level versus model-level bias  
726 correction. *Ecography*, 44(9):1341–1352, 2021.
- 727 S M Evers, T M Knight, D W Inouye, TEX Miller, R Salguero-Gómez, A M Iler, and A Compagnoni.  
728 Lagged and dormant season climate better predict plant vital rates than climate during the  
729 growing season. *Global Change Biology*, 27(9):1927–1941, 2021.
- 730 P E M Fine. Vectors and vertical transmission: an epidemiologic perspective. *Annals of the New*  
731 *York Academy of Sciences*, 266(1):173–194, 1975.
- 732 J C Fowler, M L Donald, J L Bronstein, and TEX Miller. The geographic footprint of mutualism:  
733 How mutualists influence species' range limits. *Ecological Monographs*, 93(1):e1558, 2023.
- 734 J C Fowler, S Ziegler, K D Whitney, J A Rudgers, and TEX Miller. Microbial symbionts buffer  
735 hosts from the demographic costs of environmental stochasticity. *Ecology Letters*, 27(5):e14438,  
736 2024.
- 737 PR Frade, F De Jongh, F Vermeulen, J Van Bleijswijk, and RPM Bak. Variation in symbiont  
738 distribution between closely related coral species over large depth ranges. *Molecular Ecology*, 17  
739 (2):691–703, 2008.
- 740 M E Frederickson. Mutualisms are not on the verge of breakdown. *Trends in Ecology & Evolution*,  
741 32(10):727–734, 2017.
- 742 W Gaul, D Sadykova, H J White, L Leon-Sanchez, P Caplat, M C Emmerson, and J M Yearsley.  
743 Data quantity is more important than its spatial bias for predictive species distribution modelling.  
744 *PeerJ*, 8:e10411, 2020.

- 745 GBIF.Org. Occurrence download, 2025a. URL <https://www.gbif.org/occurrence/download/>
- 746 0018831-250127130748423.
- 747 GBIF.Org. Occurrence download, 2025b. URL <https://www.gbif.org/occurrence/download/>
- 748 0018850-250127130748423.
- 749 GBIF.Org. Occurrence download, 2025c. URL <https://www.gbif.org/occurrence/download/>
- 750 0018851-250127130748423.
- 751 A Gelman and J Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, 2006.
- 752
- 753 S E Gilman, M C Urban, J Tewksbury, G W Gilchrist, and R D Holt. A framework for community
- 754 interactions under climate change. *Trends in Ecology & Evolution*, 25(6):325–331, 2010.
- 755 G Granath, M Vicari, D R Bazely, J P Ball, A Puentes, and T Rakocevic. Variation in the abundance
- 756 of fungal endophytes in fescue grasses along altitudinal and grazing gradients. *Ecography*, 30(3):
- 757 422–430, 2007.
- 758 A Gross, C Petitcollin, C Dutech, B Ly, M Massot, J Faivre d’Arcier, L Dubois, G Saint-Jean, and
- 759 M-L Desprez-Loustau. Hidden invasion and niche contraction revealed by herbaria specimens
- 760 in the fungal complex causing oak powdery mildew in europe. *Biological Invasions*, 23:885–901,
- 761 2021.
- 762 P E Gundel, I Zabalgogeazcoa, and BRV De Aldana. Interaction between plant genotype and
- 763 the symbiosis with *Epichloë* fungal endophytes in seeds of red fescue (*Festuca rubra*). *Crop and*
- 764 *Pasture Science*, 62(11):1010–1016, 2011.
- 765 P E Gundel, A C Ueno, C Casas, TEX Miller, L I Pérez, R Cuyeu, and M Omacini. Temporal
- 766 host–symbiont dynamics in community contexts: Impacts of host fitness and vertical transmission
- 767 efficiency on symbiosis prevalence. *Functional Ecology*, 38(12):2610–2622, 2024.

- 768 E K Hart and K Bell. *prism: Download data from the Oregon prism project*, 2015. URL <https://github.com/ropensci/prism>. R package version 0.0.6.
- 770 J J Heberling and D J Burke. Utilizing herbarium specimens to quantify historical mycorrhizal  
771 communities. *Applications in plant sciences*, 7(4):e01223, 2019.
- 772 R J Hijmans, S Phillips, J Leathwick, J Elith, and Robert J Hijmans. Package ‘dismo’. *Circles*, 9  
773 (1):1–68, 2017.
- 774 J HilleRisLambers, M A Harsch, A K Ettinger, K R Ford, and E J Theobald. How will biotic  
775 interactions influence climate change-induced range shifts? *Annals of the New York Academy of  
776 Sciences*, 1297(1):112–125, 2013.
- 777 IPCC. Climate change 2021: The physical science basis, 2021. URL <https://www.ipcc.ch/report/ar6/wg1/>.
- 779 N J B Isaac, M A Jarzyna, P Keil, L I Dambl, P H Boersch-Supan, E Browning, S N Freeman,  
780 N Golding, G Guillera-Arroita, P A Henrys, et al. Data integration for large-scale models of  
781 species distributions. *Trends in Ecology & Evolution*, 35(1):56–67, 2020.
- 782 A Jiménez-Valverde. Insights into the area under the receiver operating characteristic curve (auc)  
783 as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*,  
784 21(4):498–507, 2012.
- 785 D Kahle and H Wickham. Package ‘ggmap’. *Retrieved September*, 5:2021, 2019.
- 786 M R Kazenel, C L Debban, L Ranelli, W Q Hendricks, Y A Chung, T H Pendergast IV, N D  
787 Charlton, C A Young, and J A Rudgers. A mutualistic endophyte alters the niche dimensions of  
788 its host plant. *AoB plants*, 7:plv005, 2015.
- 789 R A Knapp, G M Fellers, P M Kleeman, D AW Miller, V T Vredenburg, E B Rosenblum, and C J  
790 Briggs. Large-scale recovery of an endangered amphibian despite ongoing exposure to multiple  
791 stressors. *Proceedings of the National Academy of Sciences*, 113(42):11889–11894, 2016.

- 792 M V Kozlov, I V Sokolova, V Zverev, A A Egorov, M Y Goncharov, and E L Zvereva. Biases in  
793 estimation of insect herbivory from herbarium specimens. *Scientific Reports*, 10(1):12298, 2020.
- 794 B R Lee, E F Alecrim, T K Miller, J RK Forrest, J M Heberling, R B Primack, and R D Sargent.  
795 Phenological mismatch between trees and wildflowers: Reconciling divergent findings in two recent  
796 analyses. *Journal of Ecology*, 112(6):1184–1199, 2024.
- 797 J Lendemer, B Thiers, A K Monfils, J Zaspel, E R Ellwood, A Bentley, K LeVan, J Bates, D Jen-  
798 nings, D Contreras, et al. The extended specimen network: A strategy to enhance us biodiversity  
799 collections, promote research and education. *BioScience*, 70(1):23–30, 2020.
- 800 A Leuchtmann. Systematics, distribution, and host specificity of grass endophytes. *Natural toxins*,  
801 1(3):150–162, 1992.
- 802 A Leuchtmann, C W Bacon, C L Schardl, J F White Jr, and M Tadych. Nomenclatural realignment  
803 of *Neotyphodium* species with genus *Epichloë*. *Mycologia*, 106(2):202–215, 2014.
- 804 F Lindgren, H Rue, and J Lindström. An explicit link between gaussian fields and gaussian markov  
805 random fields: the stochastic partial differential equation approach. *Journal of the Royal Statis-  
806 tical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011.
- 807 C Liu, P M Berry, T P Dawson, and R G Pearson. Selecting thresholds of occurrence in the  
808 prediction of species distributions. *Ecography*, 28(3):385–393, 2005.
- 809 D D Mc Cargo, L J Iannone, M V Vignale, C L Schardl, and M S Rossi. Species diversity of  
810 *Epichloë* symbiotic with two grasses from southern Argentinean Patagonia. *Mycologia*, 106(2):  
811 339–352, 2014.
- 812 M McFall-Ngai, M G Hadfield, TCG Bosch, H V Carey, T Domazet-Lošo, A E Douglas, N Dubilier,  
813 G Eberl, T Fukami, S F Gilbert, et al. Animals in a bacterial world, a new imperative for the  
814 life sciences. *Proceedings of the National Academy of Sciences*, 110(9):3229–3236, 2013.

- 815 T D Meehan, N L Michel, and H Rue. Spatial modeling of Audubon Christmas Bird Counts reveals  
816 fine-scale patterns and drivers of relative abundance trends. *Ecosphere*, 10(4):e02707, 2019.
- 817 E K Meineke and B H Daru. Bias assessments to expand research harnessing biological collections.  
818 *Trends in Ecology & Evolution*, 36(12):1071–1082, 2021.
- 819 E K Meineke, C C Davis, and T J Davies. The unrealized potential of herbaria for global change  
820 biology. *Ecological Monographs*, 88(4):505–525, 2018.
- 821 E K Meineke, A T Classen, N J Sanders, and T J Davies. Herbarium specimens reveal increasing  
822 herbivory over the past century. *Journal of Ecology*, 107(1):105–117, 2019.
- 823 A R Meyer, M Valentin, L Liulevicius, T R McDonald, M P Nelsen, J Pengra, R J Smith, and  
824 D Stanton. Climate warming causes photobiont degradation and carbon starvation in a boreal climate  
825 sentinel lichen. *American Journal of Botany*, 2022.
- 826 D AW Miller, K Pacifici, J S Sanderlin, and B J Reich. The recent past and promising future for  
827 data integration methods to estimate species' distributions. *Methods in Ecology and Evolution*,  
828 10(1):22–37, 2019.
- 829 D S Park, I Breckheimer, A C Williams, E Law, A M Ellison, and C C Davis. Herbarium specimens  
830 reveal substantial and unexpected variation in phenological sensitivity across the eastern United  
831 States. *Philosophical Transactions of the Royal Society B*, 374(1763):20170394, 2019.
- 832 B J Parker, J Hrček, AHC McLean, and H C J Godfray. Genotype specificity among hosts,  
833 pathogens, and beneficial microbes influences the strength of symbiont-mediated protection. *Evo-  
834 lution*, 71(5):1222–1231, 2017.
- 835 M Parniske. Arbuscular mycorrhiza: The mother of plant root endosymbioses. *Nature Reviews  
836 Microbiology*, 6(10):763–775, 2008.
- 837 A Pauw and J A Hawkins. Reconstruction of historical pollination rates reveals linked declines of  
838 pollinators and plants. *Oikos*, 120(3):344–349, 2011.

- 839 S Piao, Q Liu, A Chen, I A Janssens, Y Fu, J Dai, L Liu, X Lian, M Shen, and X Zhu. Plant  
840 phenology and global climate change: Current progresses and challenges. *Global Change Biology*,  
841 25(6):1922–1940, 2019.
- 842 T Poisot, G Bergeron, K Cazelles, T Dallas, D Gravel, A MacDonald, B Mercier, C Violet, and  
843 S Vissault. Global knowledge gaps in species interaction networks data. *Journal of Biogeography*,  
844 48(7):1552–1563, 2021.
- 845 N E Rafferty, P J CaraDonna, and J L Bronstein. Phenological shifts and the fate of mutualisms.  
846 *Oikos*, 124(1):14–21, 2015.
- 847 C J Raxworthy and B T Smith. Mining museums for historical DNA: Advances and challenges in  
848 museomics. *Trends in Ecology & Evolution*, 36(11):1049–1060, 2021.
- 849 F Renoz, I Pons, and T Hance. Evolutionary responses of mutualistic insect–bacterial symbioses in  
850 a world of fluctuating temperatures. *Current Opinion in Insect Science*, 35:20–26, 2019.
- 851 E L Roberts and A Ferraro. Rhizosphere microbiome selection by *Epichloë* endophytes of *Festuca*  
852 *arundinacea*. *Plant and Soil*, 396:229–239, 2015.
- 853 R J Rodriguez, J F White Jr, Anne E Arnold, and a RS and Redman. Fungal endophytes: diversity  
854 and functional roles. *New Phytologist*, 182(2):314–330, 2009.
- 855 G Rolshausen, F Dal Grande, A D Sadowska-Deś, J Otte, and I Schmitt. Quantifying the climatic  
856 niche of symbiont partners in a lichen symbiosis indicates mutualist-mediated niche expansions.  
857 *Ecography*, 41(8):1380–1392, 2018.
- 858 J A Rudgers, M E Afkhami, M A Rúa, A J Davitt, S Hammer, and V M Huguet. A fungus among  
859 us: broad patterns of endophyte distribution in the grasses. *Ecology*, 90(6):1531–1539, 2009.
- 860 J A Rudgers, M E Afkhami, L Bell-Dereske, Y A Chung, K M Crawford, S N Kivlin, M A Mann,  
861 and M A Nuñez. Climate disruption of plant-microbe interactions. *Annual review of ecology,*  
862 *evolution, and systematics*, 51(1):561–586, 2020.

- 863 Jennifer A Rudgers and Angela L Swafford. Benefits of a fungal endophyte in *Elymus virginicus*  
864 decline under drought stress. *Basic and Applied Ecology*, 10(1):43–51, 2009.
- 865 H Rue, S Martino, and N Chopin. Approximate bayesian inference for latent gaussian models by  
866 using integrated nested laplace approximations. *Journal of the royal statistical society: Series b*  
867 (*statistical methodology*), 71(2):319–392, 2009.
- 868 R F Sage. Global change biology: a primer. *Global Change Biology*, 26(1):3–30, 2020.
- 869 K Saikkonen, P E Gundel, and M Helander. Chemical ecology mediated by fungal endophytes in  
870 grasses. *Journal of chemical ecology*, 39:962–968, 2013.
- 871 C L Schardl, C A Young, J R Faulkner, S Florea, and J Pan. Chemotypic diversity of *Epichloae*,  
872 fungal symbionts of grasses. *Fungal Ecology*, 5(3):331–344, 2012.
- 873 María Semmartin, Marina Omacini, Pedro E Gundel, and Ignacio M Hernández-Agramonte. Broad-  
874 scale variation of fungal-endophyte incidence in temperate grasses. *Journal of Ecology*, 103(1):  
875 184–190, 2015.
- 876 S I Seneviratne, X Zhang, M Adnan, W Badi, C Dereczynski, A D Luca, S Ghosh, I Iskandar,  
877 J Kossin, S Lewis, et al. Weather and climate extreme events in a changing climate. 2021.
- 878 D Simpson, H Rue, A Riebler, T G Martins, and S H Sørbye. Penalising model component com-  
879 plexity: A principled, practical approach to constructing priors. 2017.
- 880 Michelle E Sneck, Jennifer A Rudgers, Carolyn A Young, and Tom EX Miller. Variation in the  
881 prevalence and transmission of heritable symbionts across host populations in heterogeneous en-  
882 vironments. *Microbial Ecology*, 74:640–653, 2017.
- 883 T F Stocker, D Qin, G-K Plattner, L V Alexander, S K Allen, N L Bindoff, F-M Bréon, J A Church,  
884 U Cubasch, S Emori, et al. Technical summary. In *Climate Change 2013: The Physical Science*  
885 *Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental*  
886 *Panel on Climate Change*, pages 33–115. Cambridge University Press, 2013.

- 887 P A Stott, N P Gillett, G C Hegerl, D J Karoly, D A Stone, X Zhang, and F Zwiers. Detection and  
888 attribution of climate change: A regional perspective. *Wiley Interdisciplinary Reviews: Climate*  
889 *Change*, 1(2):192–211, 2010.
- 890 S Sully, DE Burkepile, MK Donovan, G Hodgson, and R Van Woesik. A global analysis of coral  
891 bleaching over the past two decades. *Nature communications*, 10(1):1–5, 2019.
- 892 E Toby Kiers, T M Palmer, A R Ives, J F Bruno, and J L Bronstein. Mutualisms in a changing  
893 world: an evolutionary perspective. *Ecology Letters*, 13(12):1459–1474, 2010.
- 894 A T Tredennick, G Hooker, S P Ellner, and P B Adler. A practical guide to selecting models for  
895 exploration, inference, and prediction in ecology. *Ecology*, 102(6):e03336, 2021.
- 896 A D Treindl, J Stapley, and A Leuchtmann. Genetic diversity and population structure of *Epichloë*  
897 fungal pathogens of plants in natural ecosystems. *Frontiers in Ecology and Evolution*, 11:1129867,  
898 2023.
- 899 K E Trenberth, J T Fasullo, and T G Shepherd. Attribution of climate extreme events. *Nature*  
900 *Climate Change*, 5(8):725–730, 2015.
- 901 A M Truitt, M Kapun, R Kaur, and W J Miller. Wolbachia modifies thermal preference in drosophila  
902 melanogaster. *Environmental microbiology*, 21(9):3259–3268, 2019.
- 903 Shripad D. Tuljapurkar. Population dynamics in variable environments. III. Evolutionary dynamics  
904 of r-selection. *Theoretical Population Biology*, 21(1):141–165, February 1982. ISSN 0040-5809.  
905 doi: 10.1016/0040-5809(82)90010-7. URL <http://www.sciencedirect.com/science/article/pii/0040580982900107>.
- 907 A Urdangarin, T Goicoa, and M D Ugarte. Evaluating recent methods to overcome spatial con-  
908 founding. *Revista Matemática Complutense*, 36(2):333–360, 2023.
- 909 V Vikuk, C A Young, S T Lee, P Nagabhyru, M Krischke, M J Mueller, and J Krauss. Infection

- 910 rates and alkaloid patterns of different grass species with systemic *Epichloë* endophytes. *Applied*  
911 *and Environmental Microbiology*, 85(17):e00465–19, 2019.
- 912 M von Cräutlein, M Helander, H Korpelainen, P H Leinonen, B R Vázquez de Aldana, C A Young,  
913 I Zabalgogeazcoa, and K Saikkonen. Genetic diversity of the symbiotic fungus *Epichloë festucae*  
914 in naturally occurring host grass populations. *Frontiers in Microbiology*, 12:756991, 2021.
- 915 Z Wang, C Li, and J White. Effects of *Epichloë* endophyte infection on growth, physiological  
916 properties and seed germination of wild barley under saline conditions. *Journal of Agronomy and*  
917 *Crop Science*, 206(1):43–51, 2020.
- 918 R J Warren and M A Bradford. Mutualism fails when climate response differs between interacting  
919 species. *Global Change Biology*, 20(2):466–474, 2014.
- 920 N S Webster, R E Cobb, and A P Negri. Temperature thresholds for bacterial symbiosis with a  
921 sponge. *The ISME journal*, 2(8):830–842, 2008.
- 922 J F White and G T Cole. Endophyte-host associations in forage grasses. i. Distribution of fungal  
923 endophytes in some species of *Lolium* and *Festuca*. *Mycologia*, 77(2):323–327, 1985.
- 924 F M Willems, JF Scheepens, and O Bossdorf. Forest wildflowers bloom earlier as europe warms:  
925 Lessons from herbaria and spatial modelling. *New Phytologist*, 235(1):52–65, 2022.
- 926 C G Willis, E R Ellwood, R B Primack, C C Davis, K D Pearson, A S Gallinat, J M Yost, G Nelson,  
927 S J Mazer, N L Rossington, et al. Old plants, new tricks: Phenological research using herbarium  
928 specimens. *Trends in Ecology & Evolution*, 32(7):531–546, 2017.
- 929 C Xia, N Li, Y Zhang, C Li, X Zhang, and Z Nan. Role of *Epichloë* endophytes in defense responses  
930 of cool-season grasses to pathogens: A review. *Plant Disease*, 102(11):2061–2073, 2018.
- 931 K Yoshida, E Sasaki, and S Kamoun. Computational analyses of ancient pathogen DNA from  
932 herbarium samples: Challenges and prospects. *Frontiers in Plant Science*, 6:771, 2015.

933

## Appendix A

934

935       *Appendix to "Increasing Prevalence of plant-fungal symbiosis across two*  
936                   *centuries of environmental change"*

937       **Authors:**

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

939       Jacob Moutouama<sup>1</sup>

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

941

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

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

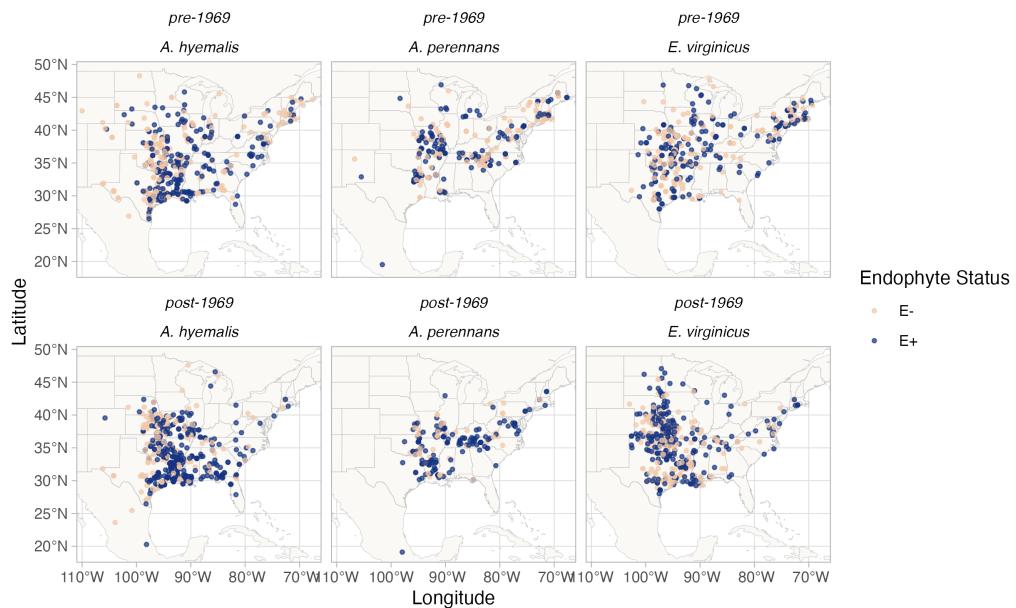
944       \* Corresponding author; e-mail: jcf221@miami.edu.

945       **Contents:**

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

947

## 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.

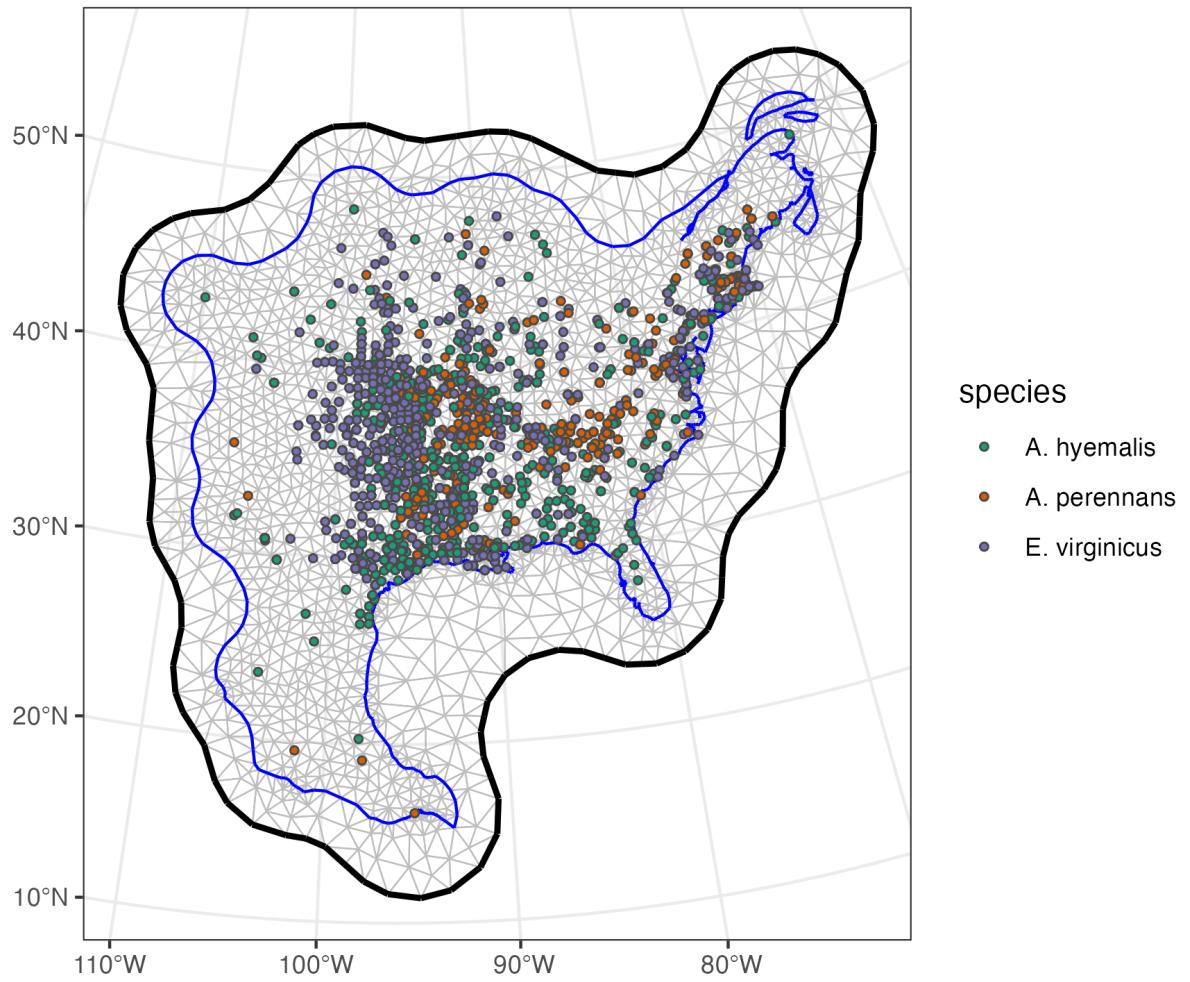


Figure A2: Triangulation mesh used to estimate spatial dependence between data points.

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.

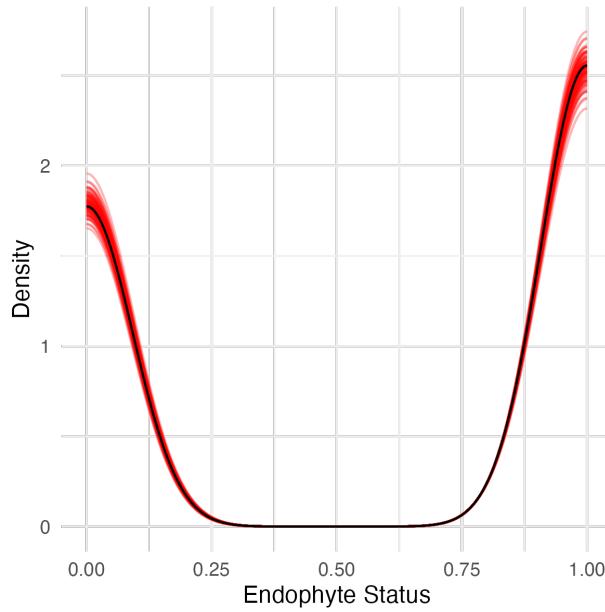
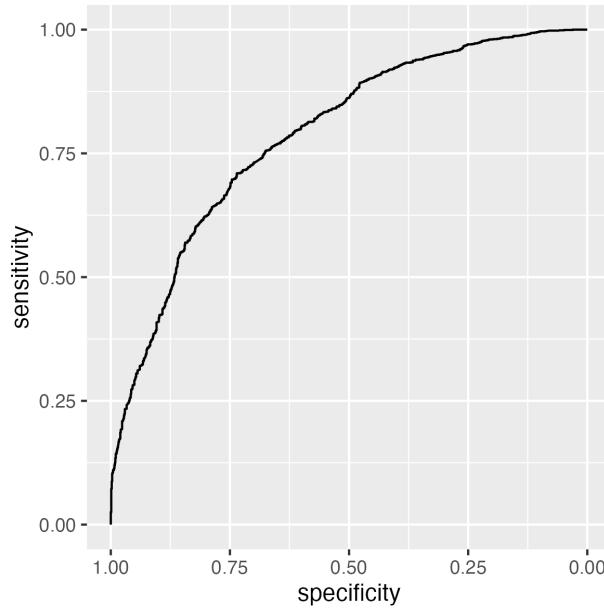
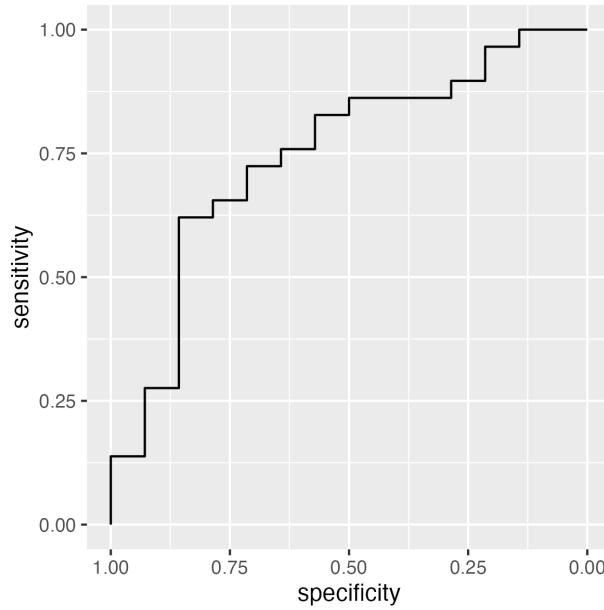


Figure A3: **Graphical posterior predictive check of the endophyte prevalence model fit.**

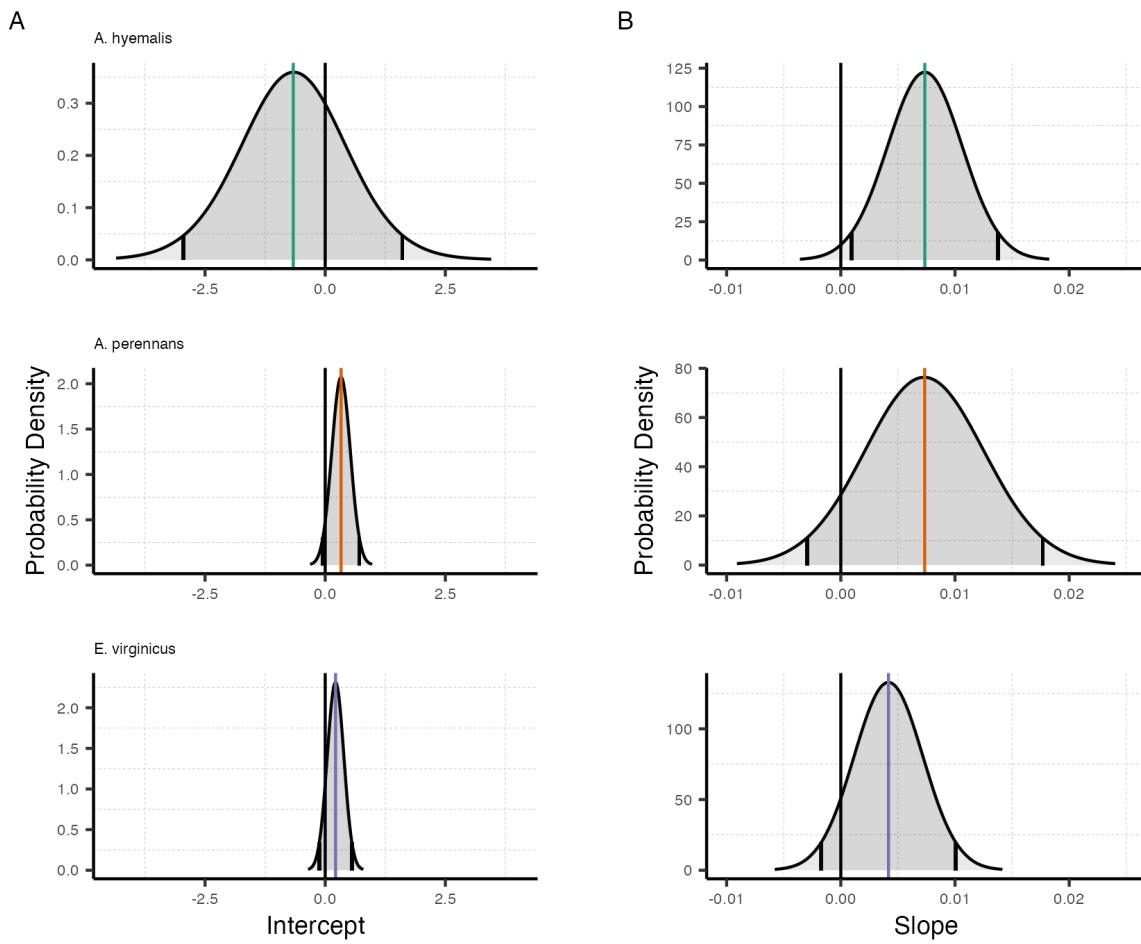
Consistency between observed data and predicted values indicate that the fitted model accurately describes the data. Graph shows density curves for the observed data (black) along with 100 predicted datasets (red).



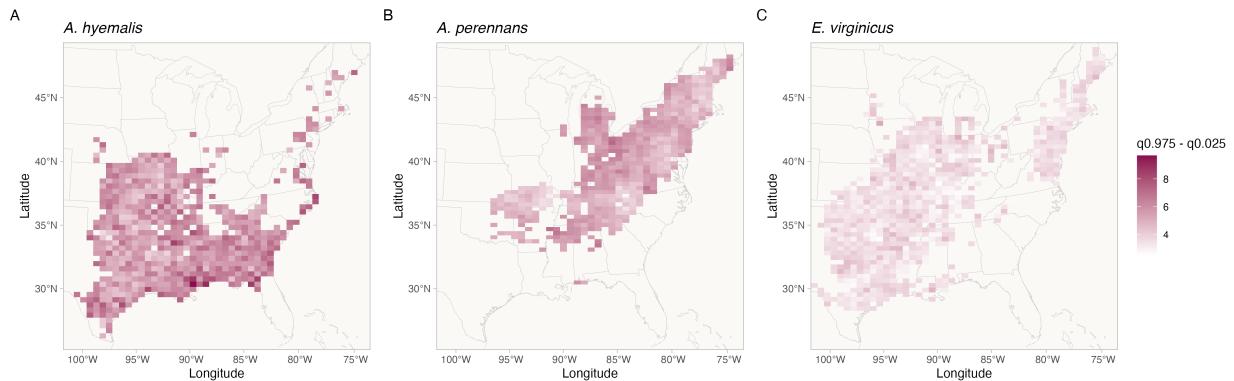
**Figure A4: ROC plot showing performance of the endophyte prevalence model in classifying observations according to endophyte status within the in-sample training data from herbarium collections.** The curves show adequate model performance for observed data. The AUC value is 0.79.



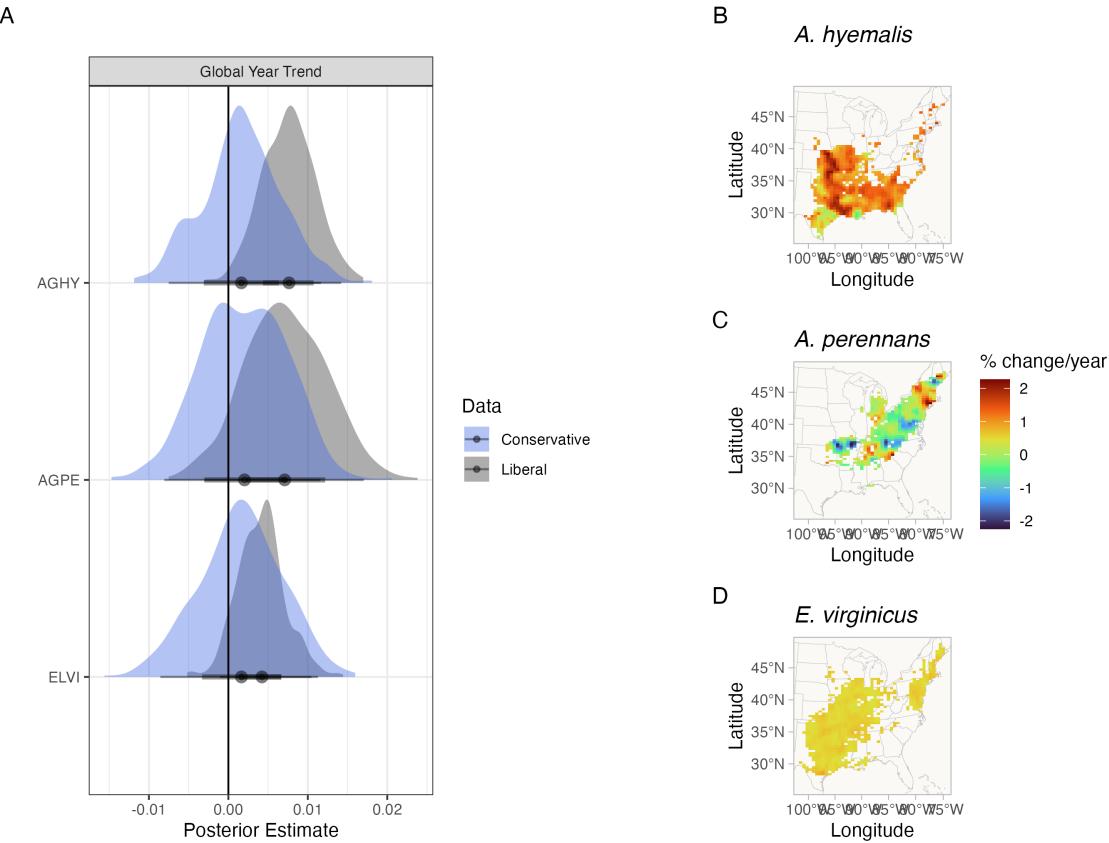
**Figure A5: ROC plot showing performance of the endophyte prevalence model in classifying observations according to endophyte status within the out-of-sample test data from contemporary surveys.** The curves show adequate model performance for test data. The AUC value is 0.77.



**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,  $A$  and  $T$  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.

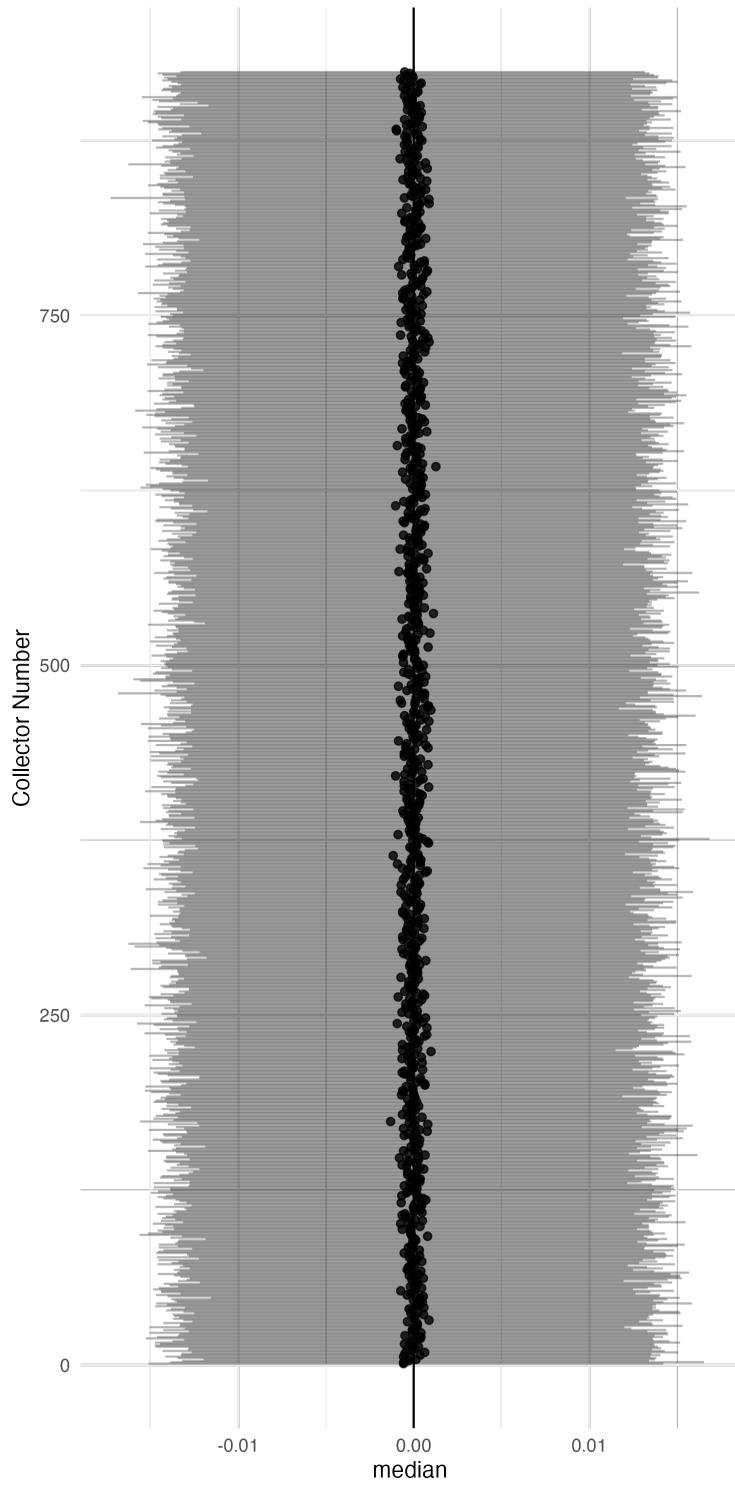
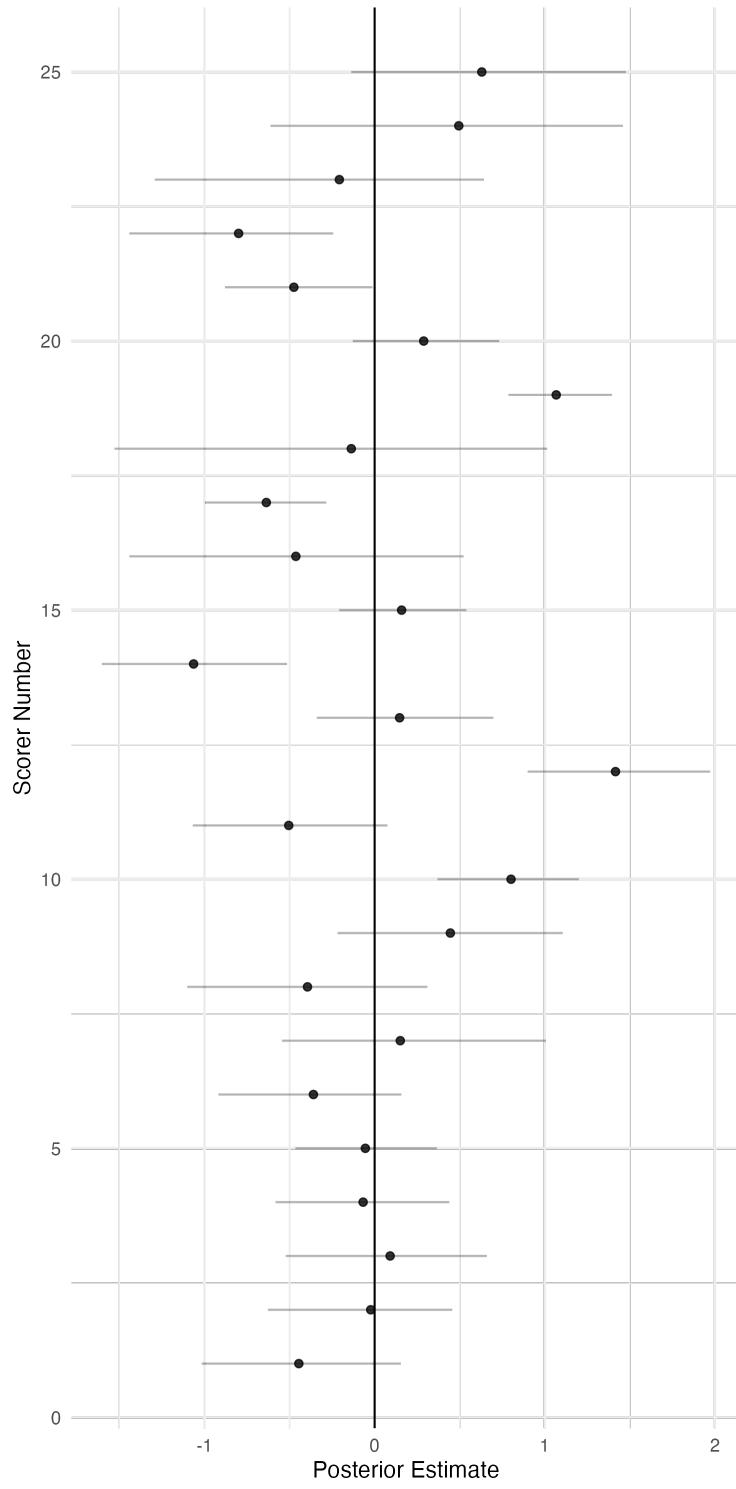
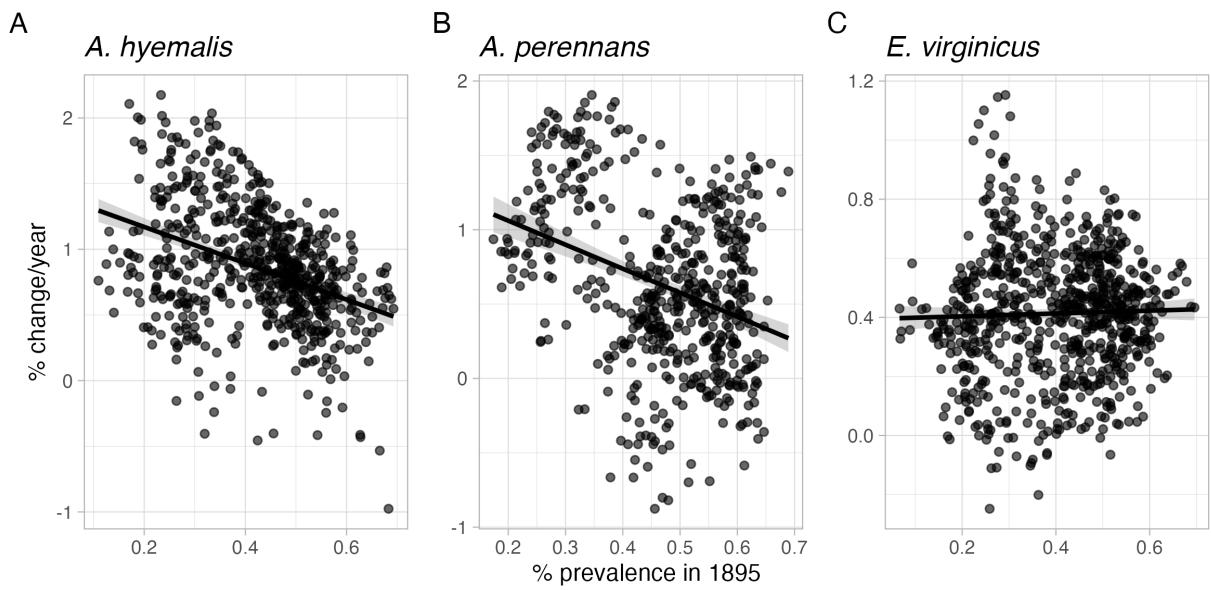


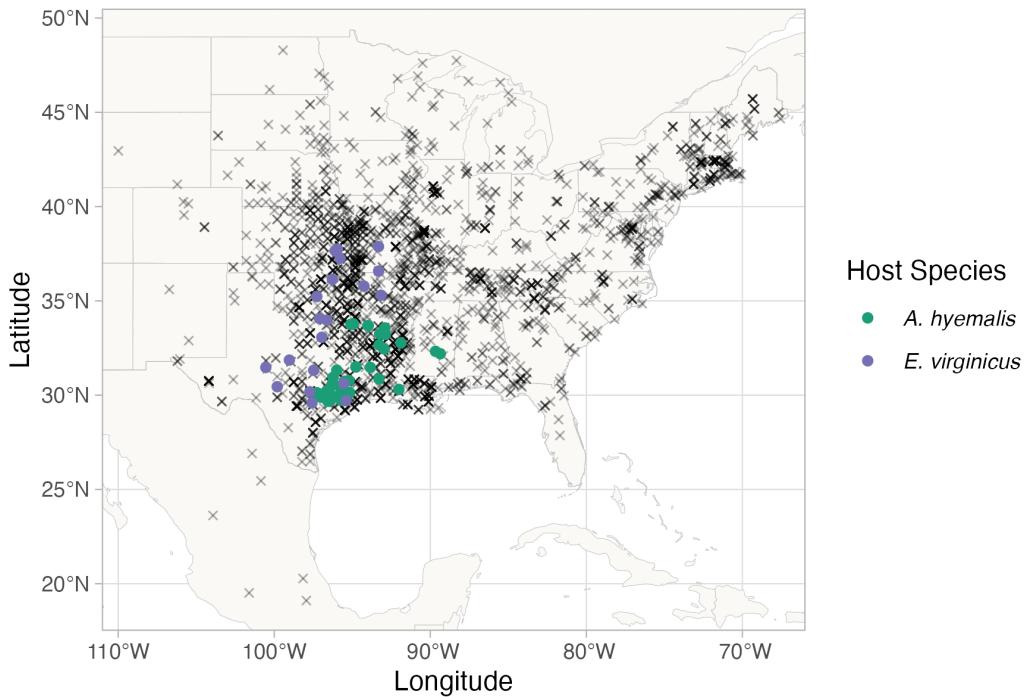
Figure A9: **Posterior estimates of collector random effects from endophyte prevalence model.** Collector random effects are denoted  $\chi$  in Eqn. 1 and represent variance associated with researchers who collected historic herbarium specimens. Points show posterior median along with 95% CI for each 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.

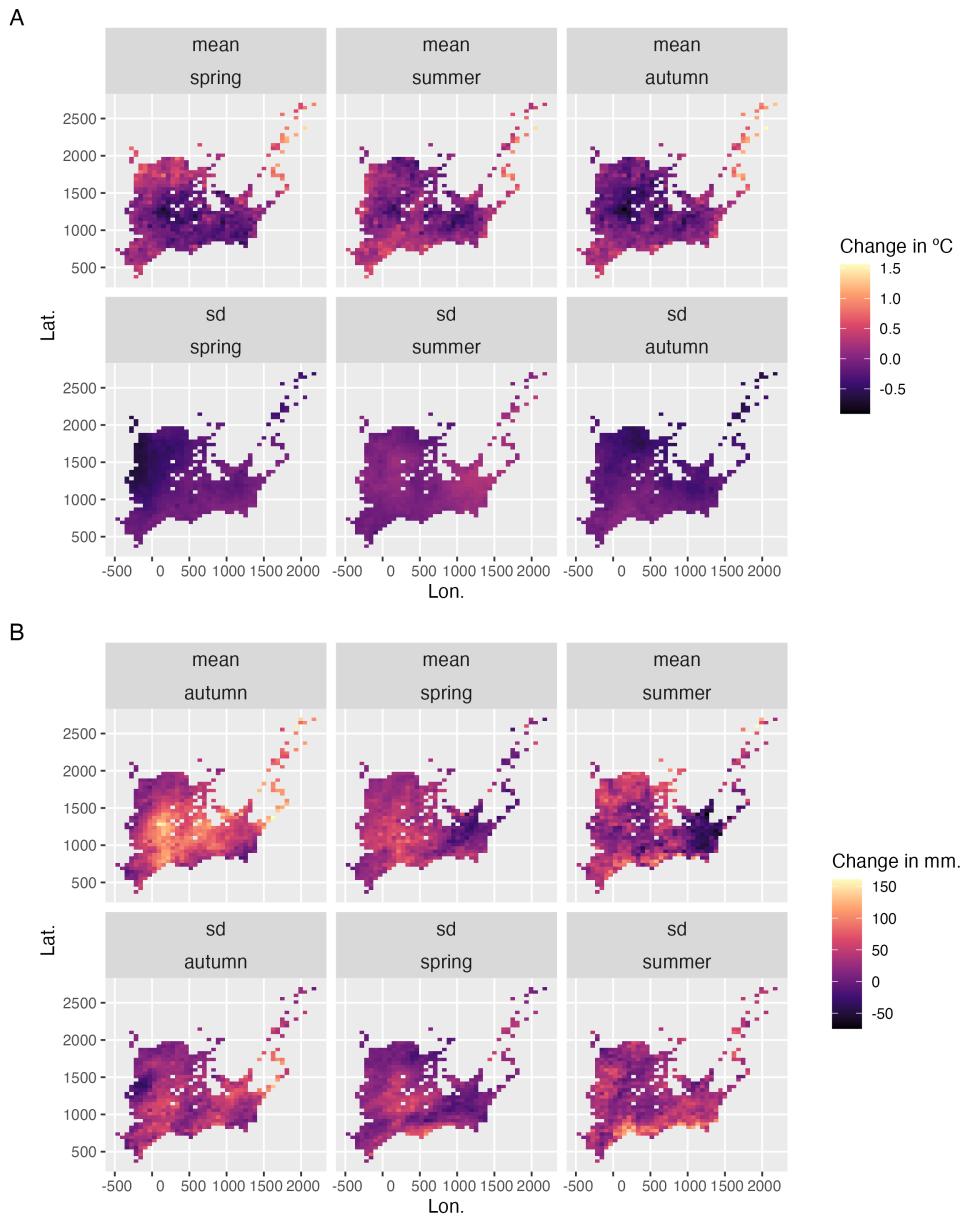
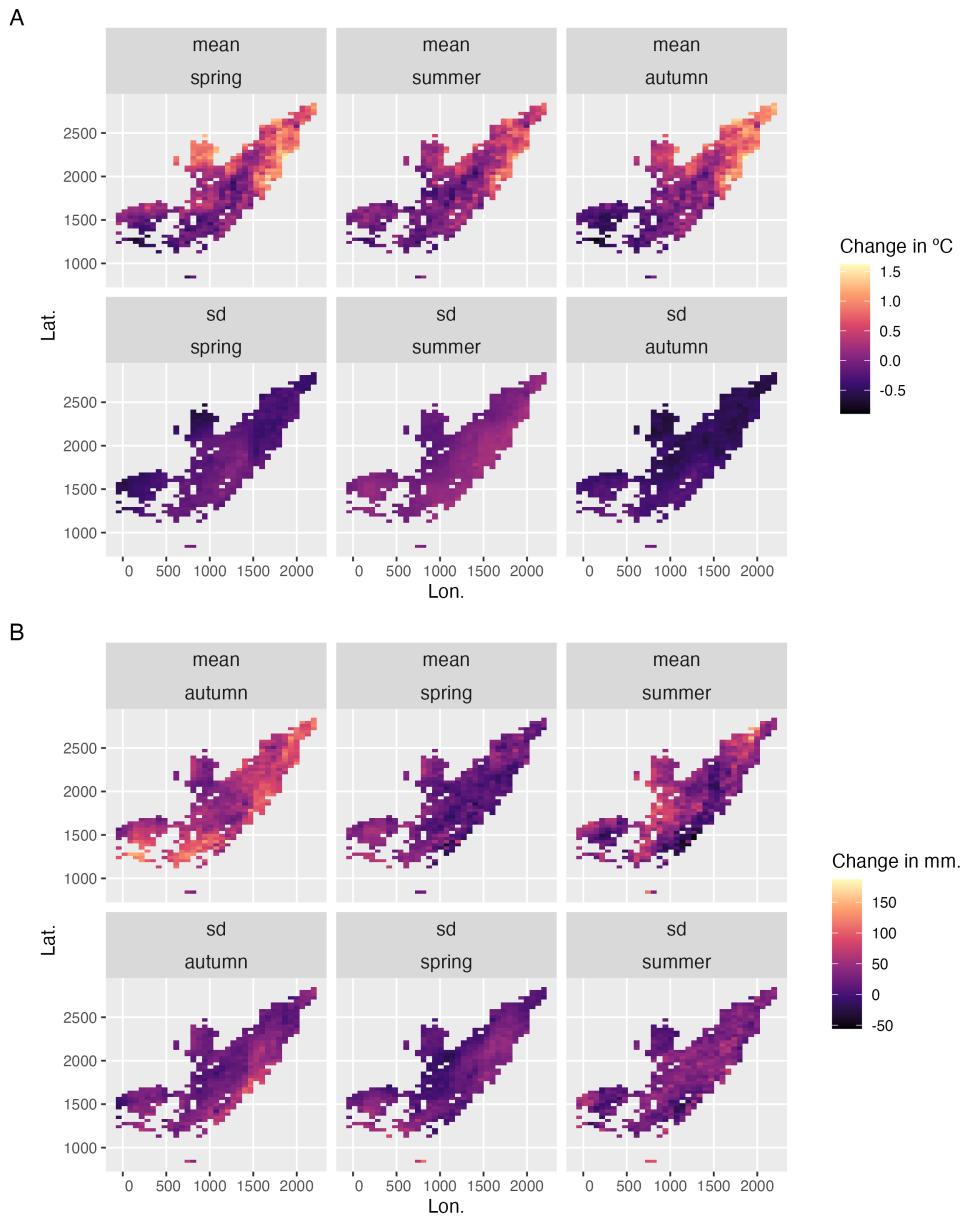


Figure A13: **Change in seasonal climate variables between the periods 1895-1925 and 1990-2020 across the distribution of *A. hyemalis*.** Color represents change in (A) seasonal temperature ( $^{\circ}\text{C}$ ) and (B) seasonal precipitation (mm.). Maps show pixels covering the modeled distribution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

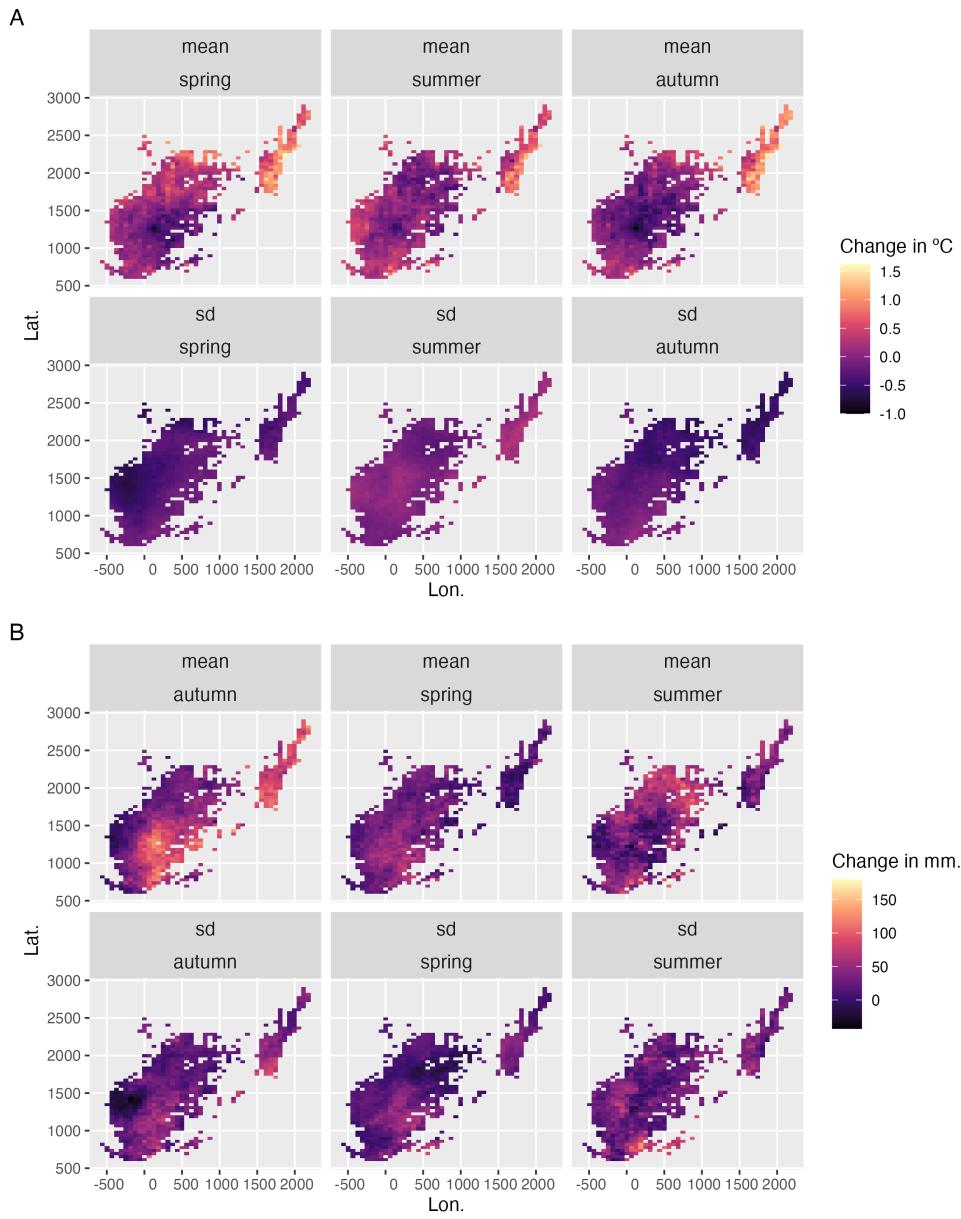


Figure A15: **Change in seasonal climate variables between the periods 1895-1925 and 1990-2020 across the distribution of *E. virginicus*.** Color represents change in (A) seasonal temperature ( $^{\circ}\text{C}$ ) and (B) seasonal precipitation (mm.). Maps show pixels covering the modeled distribution 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	0	6
Missouri Botanic Garden	210	205	122
Texas A&M	100	0	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

949

## Supporting Methods

950

### ODMAP Protocol

951 [Overview](#)

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

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

954 **Study area:** Eastern North America

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

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

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

958 **Boundary:** Natural.

959 [Data](#)

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

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

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

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

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

966 [Model](#)

967 **Model assumption:** We assumed that the target species are at equilibrium with their environment.  
968

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

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

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

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

974 **Data availability:** Data was accessed through open-source R packages *rgbif*. *A. hyemalis*  
975 (GBIF.Org, 2025a), *A. perennans* (GBIF.Org, 2025b), *E. virginicus* (GBIF.Org, 2025c)

976 [Assessment](#)

977 We used AUC to test model performance.

978 [Prediction](#)

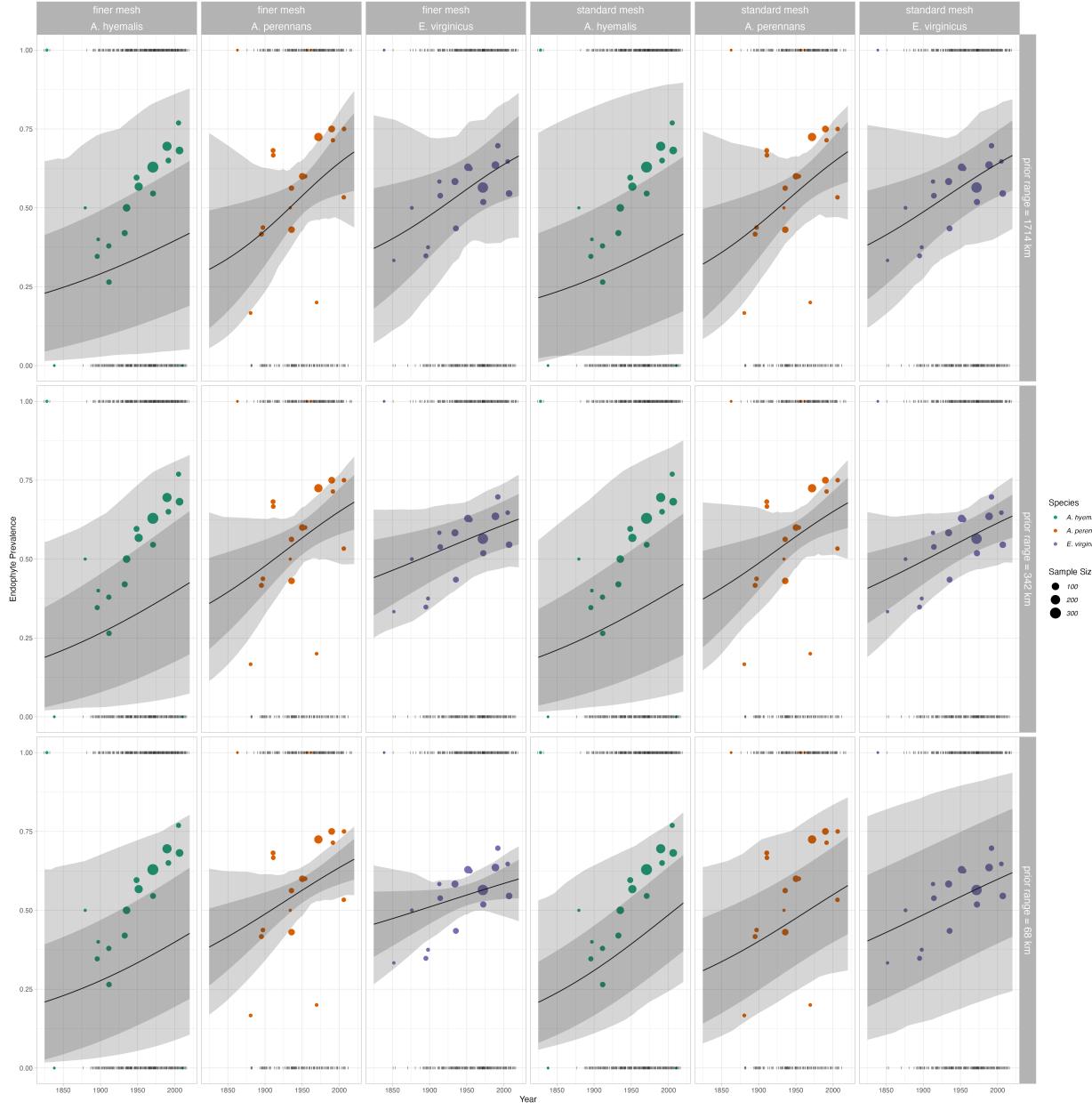
979 We predicted the probability of presence of the host species as a binary maps (presence or absence)

## 980 *Mesh and Prior Sensitivity Analysis*

981 To test the influence that the triangulation mesh and choice of priors has on results, we compared  
982 model results across a range of meshes and priors. We re-ran our model for the mesh used in main  
983 body of the text (Fig. A2), which we refer to as the "standard mesh", and with a mesh with smaller  
984 minimum vertices (finer mesh). Finer scale meshes increase computation time. For each of these  
985 meshes, we ran the model with a range of priors defining the spatial range of our spatial random  
986 effects: 342km (the prior used for presented results), as well as ranges five times smaller (68 km)

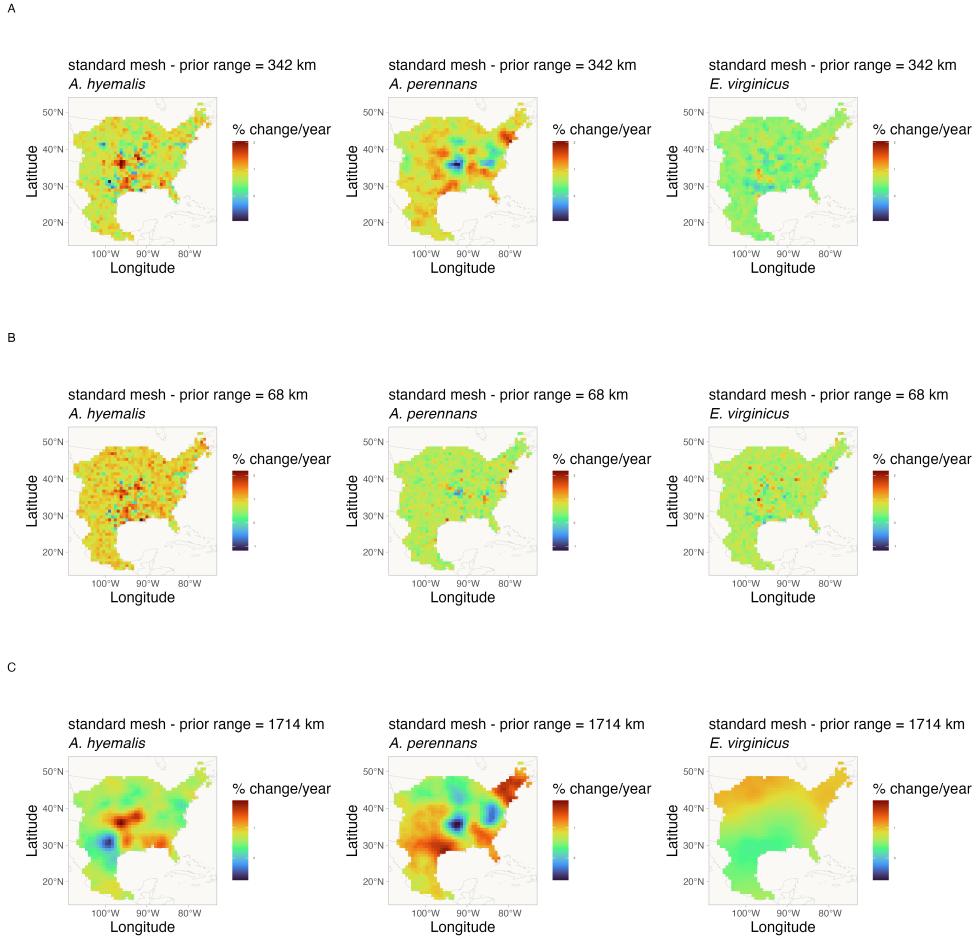
987 and five times larger (1714 km). We found generally that these choices did not alter the direction  
988 of model predictions, but did influence the associated uncertainty and magnitude of some effects.

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



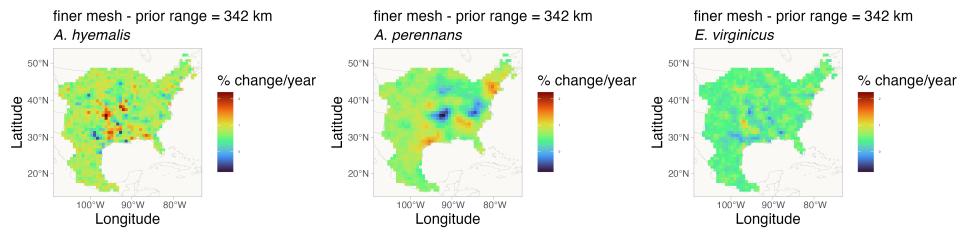
**Figure A16: Overall trend in endophyte prevalence evaluated for endophyte prevalence models with different range priors on spatially structured random effects, and for two different triangulation meshes.** Data used in model fitting is the same across all panels and as in the main text. Note that these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

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

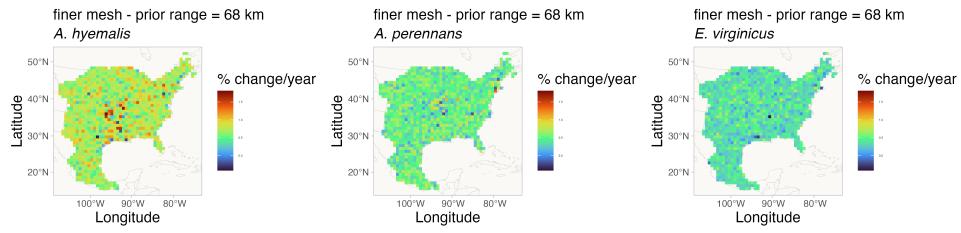


**Figure A17: Spatially-varying trends in endophyte prevalence evaluated for the endophyte prevalence model with different range priors on spatially structured random effects, and for the "standard" mesh.** Data used in model fitting is the same across all panels and as in the main text. Shading indicates 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.

A



B



C

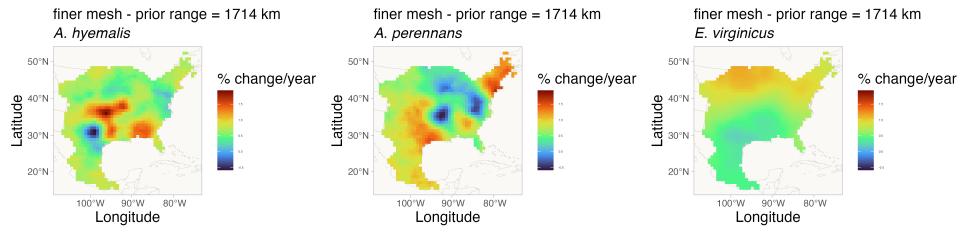


Figure A18: **Spatially-varying trends in endophyte prevalence evaluated for the endophyte prevalence model with different range priors on spatially structured random effects, and for the "finer" mesh.** Data used in model fitting is the same across all panels and as in the main text. Shading indicates 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.

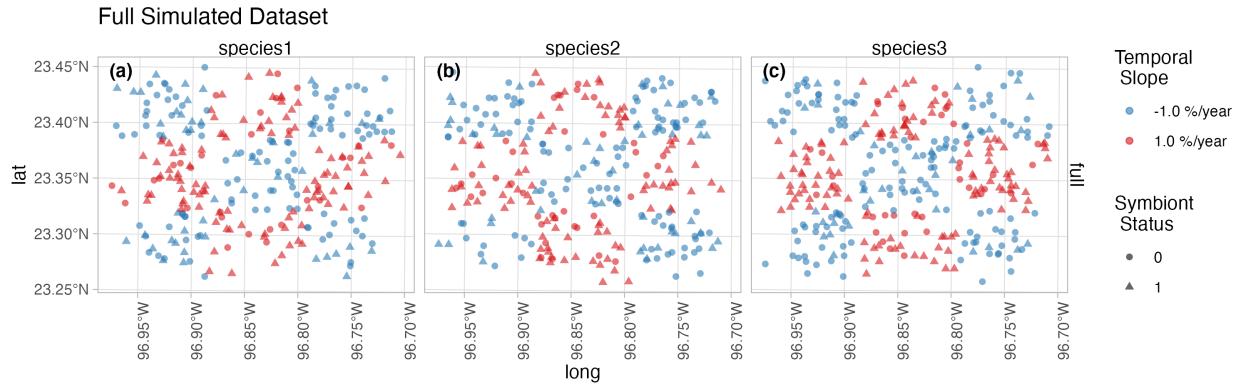
1000

## *Spatially-biased Sample Size Simulation Analysis*

1001 To examine how data that is unevenly distributed across host distributions may influence interpreta-  
1002 tion of spatially-varying coefficients, we performed a simulation analysis. Our focal species, *Agrostis*  
1003 *hyemalis*, *Agrostis perennans*, and *Elymus virginicus*, are widely distributed grasses across the east-  
1004 ern United States that host *Epichloë* fungal endophytes. For logistical reasons, our sampling visits  
1005 to herbaria focused on herbaria in the central southern U.S., which resulted in unevenly distributed  
1006 data across each host species' range. This is particularly noteable for *Agrostis perennans* which has  
1007 the most northern distribution and relatively fewer total collected specimens compared to the other  
1008 focal species. Thus, a significant portion in the northeast of this species' range is relatively sparsely  
1009 sampled. Our analysis presented in the main text identified this region as having strong increase in  
1010 endophyte prevalence. Future visits to herbaria with regional focuses in the Northeastern US would  
1011 certainly garner new specimens that could provide valuable insights into shifting host and symbiont  
1012 distributions.

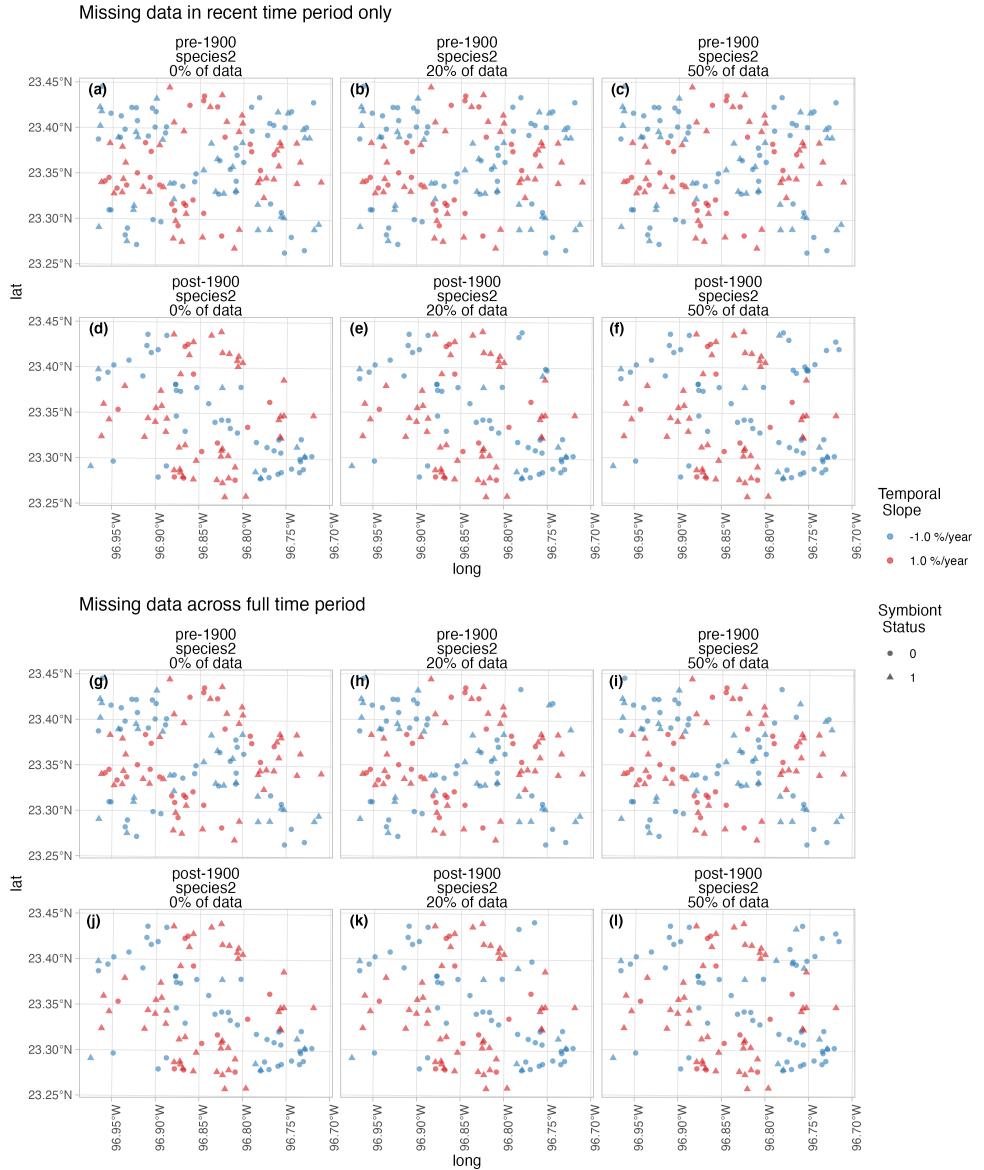
1013 *Simulation of spatially-biased symbiont occurrence data*

1014 We simulated datasets with varying levels of missing-ness to examine how this missing-ness influ-  
1015 enced the estimation of spatially-varying trend estimates. We first generated 300 data points for  
1016 each of three hypothetical species at random positions across an area approximating the scale of  
1017 our focal data. Each data point was randomly assigned a year of collection across 200 years. We  
1018 then simulated data from a Bernoulli process with trends alternating across nine regions (Fig. A19)  
1019 in a 3X3 grid pattern. This grid pattern was intended to create a complex spatial layout of trends,  
1020 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.

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



**Figure A20: Six simulated datasets representing scenarios of spatially-baised 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.

1028 *Statistical analysis*

1029 We analyzed each dataset with a model given by Eqn. A1 similar in construction to that used in  
1030 our central analysis.

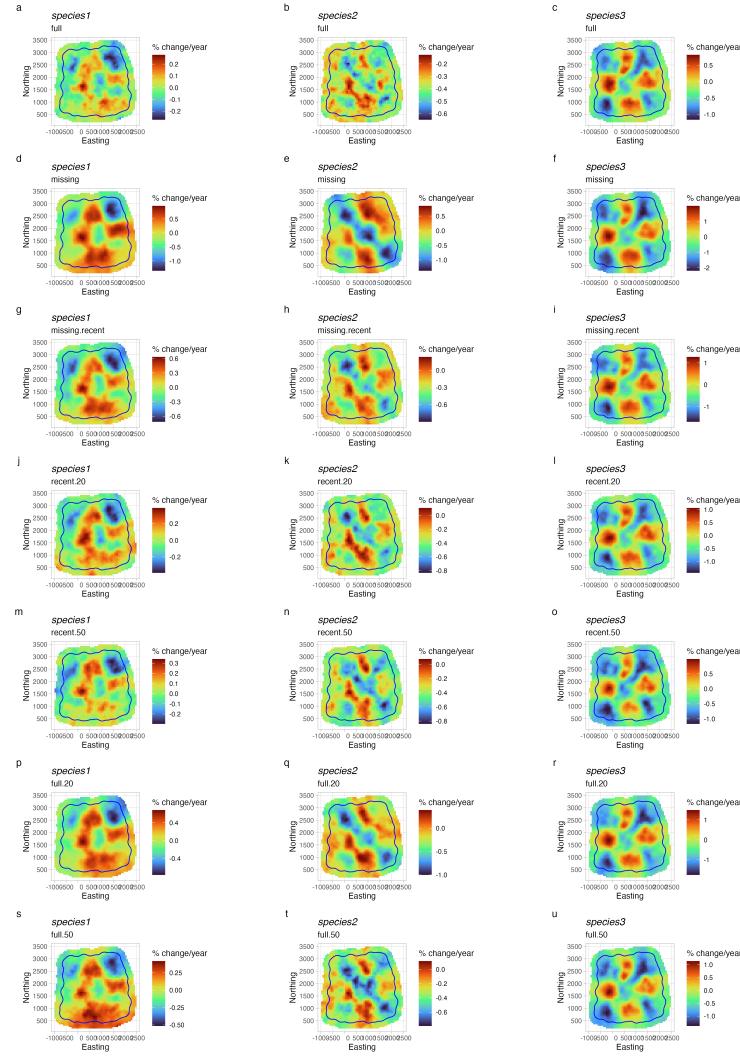
$$\text{logit}(\hat{P}_{h,i}) = A_h + T_h * \text{year}_i + \alpha_{h,l_i} + \tau_{h,l_i} * \text{year}_i + \delta_{l_i} \quad (\text{A1})$$

1031 Where symbiont presence/absence of the  $i^{th}$  specimen ( $P_{h,i}$ ) was modeled as a Bernoulli re-  
1032 sponse variable with expected probability of symbiont occurrence  $\hat{P}_{h,i}$  for each host species  $h$ . We  
1033 modeled  $\hat{P}_{h,i}$  as a linear function of intercept  $A_h$  and slope  $T_h$  defining the global trend in endophyte  
1034 prevalence specific to each host species as well as with spatially-varying intercepts  $\alpha_{h,l_i}$  and slopes  
1035  $\tau_{h,l_i}$  associated with location ( $l_i$ , the unique latitude-longitude combination of the  $i$ th observation).  
1036 Similar to the SVC model of our central analysis (Eqn. 1), we estimated a shared variance term  
1037 with the spatially-dependent random effect  $\delta_{l_i}$ , intended to account for residual spatial variation.  
1038 However in this analysis we omit i.i.d.-random effects terms associated with collector and scorer  
1039 identity ( $\chi_{c_i}$  and  $\omega_{s_i}$  in Eqn. 1) for the sake of simplicity.

1040 *Influence of spatially-biased sampling on model interpretation*

1041 Our analysis of the full simulated data shows that our model is suitably flexible to capture complex  
1042 spatial patterns in temporal trends (Fig. A21 a-c). Beyond this, the model also qualitatively  
1043 captures the spatial patterns in temporal trends even with large amounts of data missingness (i.e  
1044 missing up to 80% of the datapoints (Fig. A21 p-r)).

1045 While this analysis is not an exhaustive examination of the influence of sampling bias on our  
1046 results for several reasons (including not examining how different strengths in temporal trends,  
1047 different spatial arrangements of missing-ness influence model estimates, or different sample sizes  
1048 influence results), it demonstrates that the spatially-varying modelling framework implemented in  
1049 INLA we employ can suitably recover regional trends even with significant spatially-bias within  
1050 data collection, and further the analysis is likely robust to temporally-structured bias (missing data



**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.

1051 within recent collection period). Future work could more fully explore the scenarios that cause  
1052 this ability to break down. We expect this simulation reflects what may be a common scenario for  
1053 research investigating global change using natural history specimens. Collection effort by trained  
1054 taxonomists and professional collectors peaked in the past, and collections contain relatively fewer  
1055 modern specimens in many regions. Additionally, most global change research necessarily involves  
1056 accessing many specimens across collections. Research efforts such as ours will be unable to access  
1057 every specimen from all possible collections. Ongoing digitization efforts will make it possible to  
1058 more clearly assess how much data is missing from a particular study compared to the actual  
1059 holdings of natural history collections, but ultimately, the decision of what data and collections to  
1060 include is a question of sample size and study design.