

¹ Increasing prevalence of plant-fungal symbiosis across two
² centuries of environmental change

³ Joshua C. Fowler^{1,2*}

Jacob Moutouama¹

Tom E. X. Miller¹

⁴ 1. Rice University, Department of BioSciences, Houston, Texas 77006; 2. University of Miami,
⁵ Department of Biology, Miami, Florida;

⁶ * Corresponding author; e-mail: jcf221@miami.edu.

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

⁹ *Keywords:* climate change, *Epichloë*, herbarium, INLA, museum specimen, plant-microbe sym-
¹⁰ biosis, *Poaceae*, spatially-varying coefficients model

¹¹ *Manuscript type:* Research Article.

Abstract

13 Species' distributions and abundances are shifting in response to ongoing global climate change.

14 Mutualistic microbial symbionts can provide hosts with protection from environmental stress that

15 may promote resilience under environmental change, however this change may also disrupt species

16 interactions and lead to declines in hosts and/or symbionts. Symbionts preserved within natural

17 history specimens offer a unique opportunity to quantify changes in microbial symbiosis across broad

18 temporal and spatial scales. We asked how the prevalence of seed-transmitted fungal symbionts

19 of grasses (*Epichloë* endophytes) has changed over time in response to climate change, and how

20 these changes vary across host species' distributions. Specifically, we examined 2,346 herbarium

21 specimens of three grass host species (*Agrostis hyemalis*, *Agrostis perennans*, *Elymus virginicus*)

22 collected over the past two centuries (1824 – 2019) for the presence or absence of *Epichloë* symbiosis.

23 Analysis of an approximate Bayesian spatially-varying coefficients model revealed that endophytes

24 increased in prevalence over the last two centuries from ca. 25% to ca. 75% prevalence, on average,

25 across three host species. Changes in seasonal climate drivers were associated with increasing

26 endophyte prevalence. Notably, increasing precipitation during the peak growing season for *Agrostis*

27 species and decreasing precipitation for *E. virginicus* were associated with increasing endophyte

28 prevalence. Changes in the variability of precipitation and temperature during off-peak seasons were

29 also important predictors of increasing endophyte prevalence. Our analysis performed favorably in

30 an out-of-sample predictive test with contemporary survey data from across 63 populations, a rare

31 extra step in collections-based research. We identified greater local-scale variability in endophyte

32 prevalence in contemporary data compared to model predictions based on historic data, suggesting

33 new directions that could improve predictive accuracy. Our results provide novel evidence for a

34 cryptic biological response to climate change that may contribute to the resilience of host-microbe

35 symbiosis through fitness benefits to symbiotic hosts.

36 Abstract : 300 words

Introduction

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

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

60 Microbial symbionts are common to all macroscopic organisms and can have important effects
61 on their hosts' survival, growth and reproduction (McFall-Ngai et al., 2013; Rodriguez et al., 2009).
62 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

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

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

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

132 Methods

133 Focal species

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

¹⁴² *perennans* is more sparsely distributed, tending to be found in shadier and moister habitats, while
¹⁴³ *A. hyemalis* is commonly found in open and recently disturbed habitats. Both *Agrostis* species are
¹⁴⁴ recorded from throughout the Eastern US, but *A. perennans* has a slightly more northern distri-
¹⁴⁵ bution, whereas *A. hyemalis* is found rarely as far north as Canada and is listed as a rare plant
¹⁴⁶ in Minnesota. *E. virginicus* is a larger and relatively longer-lived species that is more broadly dis-
¹⁴⁷ tributed than the *Agrostis* species. It begins flowering as early as March or April but continues
¹⁴⁸ throughout the summer (most common collection month: July).

¹⁴⁹ Both *Agrostis* species host *Epichloë amarillans* (Craven et al., 2001; Leuchtmann et al., 2014),
¹⁵⁰ while *Elymus virginicus* typically hosts *Epichloë elymi* (Clay and Schardl, 2002). The fungal sym-
¹⁵¹ bionts primarily reproduce asexually and are passed from mother to offspring by vertical transmis-
¹⁵² sion through seeds. These traits contribute to highly specialized interactions between symbiont and
¹⁵³ host. Some host species have been shown to partner with multiple symbiont species or strains, and in
¹⁵⁴ some cases multiple symbiont lineages can co-exist within a host population (Mc Cargo et al., 2014).
¹⁵⁵ However, surveys have typically found limited *Epichloë* genotypic diversity within host populations
¹⁵⁶ (Treindl et al., 2023). Across host populations, concentrations of biologically-active alkaloids and
¹⁵⁷ the genes associated with their production vary substantially (Schardl et al., 2012). In this analysis,
¹⁵⁸ we focus on the presence/absence of *Epichloë* symbionts, and we discuss potential implications of
¹⁵⁹ symbiont genotypic diversity in the Discussion.

¹⁶⁰ *Herbarium surveys*

¹⁶¹ We visited nine herbaria between 2019 and 2022 (see Table A1 for a summary of specimens included
¹⁶² from each collection). With permission from herbarium staff, we acquired seed samples from 1135
¹⁶³ *A. hyemalis* specimens collected between 1824 and 2019, 357 *A. perennans* specimens collected
¹⁶⁴ between 1863 and 2017, and 854 *E. virginicus* specimens collected between 1839 and 2019 (Fig. 1,
¹⁶⁵ Fig. 2A, Fig. A1). We chose our focal species in part because they are commonly represented in
¹⁶⁶ herbarium collections and produce many seeds, meaning that small samples would not diminish the
¹⁶⁷ value of the specimens for future studies. We collected 5-10 seeds per specimen after examining the

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

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

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

191 After collecting seed samples, we quantified the presence or absence of *Epichloë* fungal hyphae
192 in each specimen using microscopy. We first softened seeds with a 10% NaOH solution, then stained
193 the seeds with aniline blue-lactic acid stain and squashed them under a microscope cover slip. We
194 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

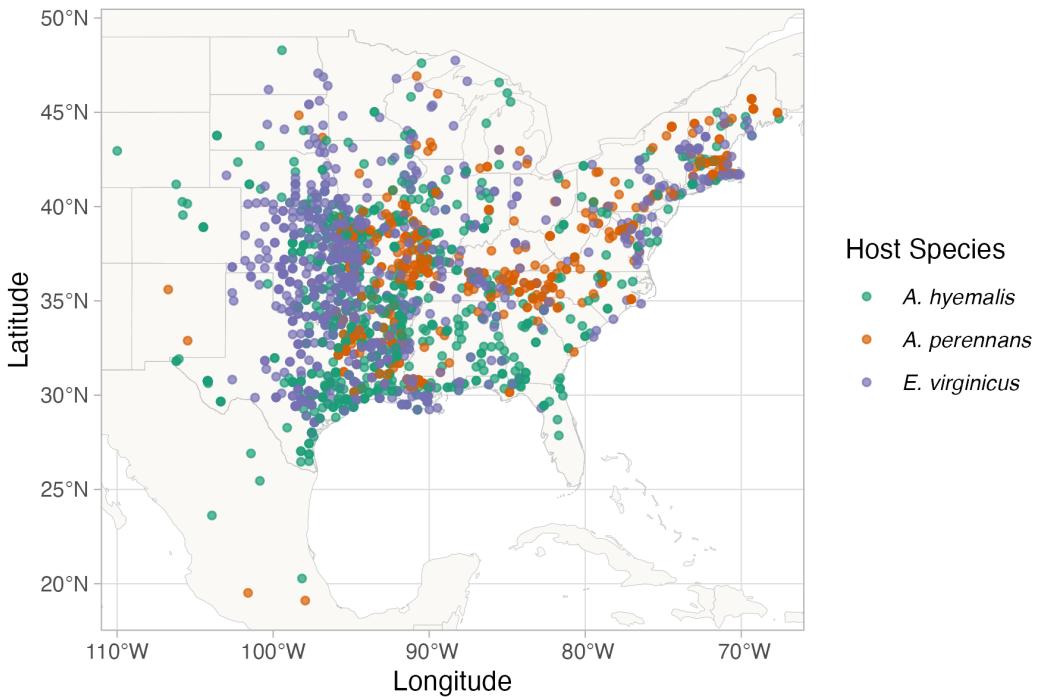


Figure 1: Collection locations of herbarium specimens sampled for *Epichloë* endophytes.

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

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

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

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

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

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

264 *Modeling distributions of host species*

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

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

285 *Assessing the role of climate drivers*

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

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

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

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

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

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

343

Results

344 *How has endophyte prevalence changed over time?*

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

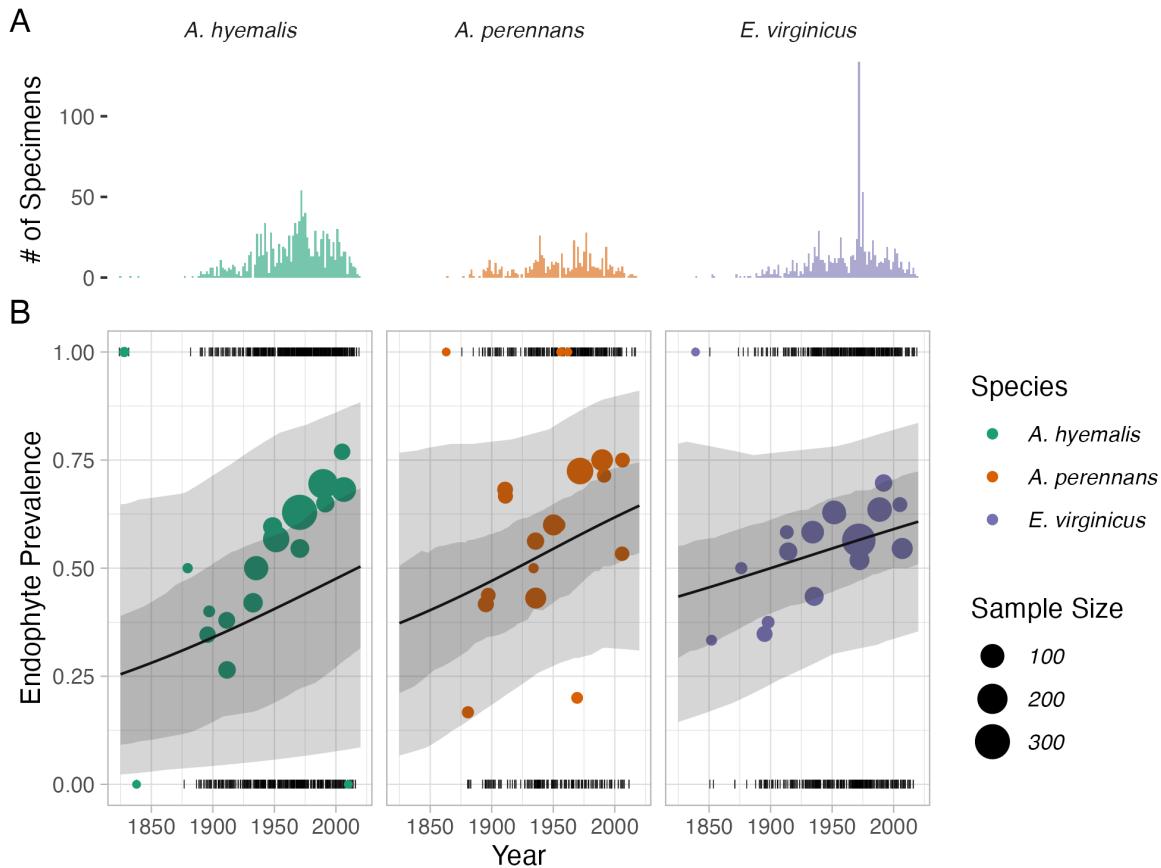


Figure 2: **Temporal trends in endophyte prevalence.** (A) Histograms show the frequency of scored specimens through time for each host species. (B) Lines show mean endophyte prevalence predicted by the endophyte prevalence model over the study period along with the 50% and 95% CI bands incorporating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

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

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

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

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

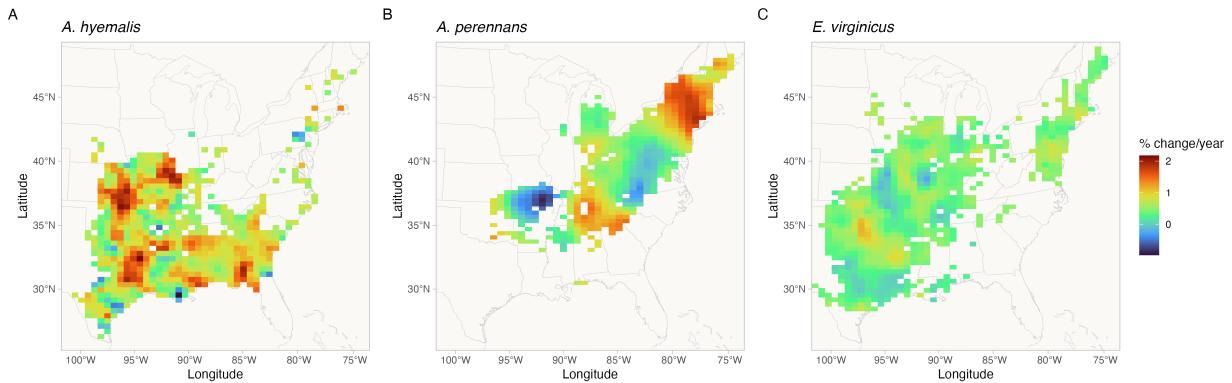


Figure 3: Predicted posterior mean of spatially-varying slopes representing change in endophyte prevalence for each host species (A, *A. hyemalis*; B, *A. perennans*; C, *E. virginicus*). **Spatially-varying trends are estimated from the endophyte prevalence model.** Color indicates the relative change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

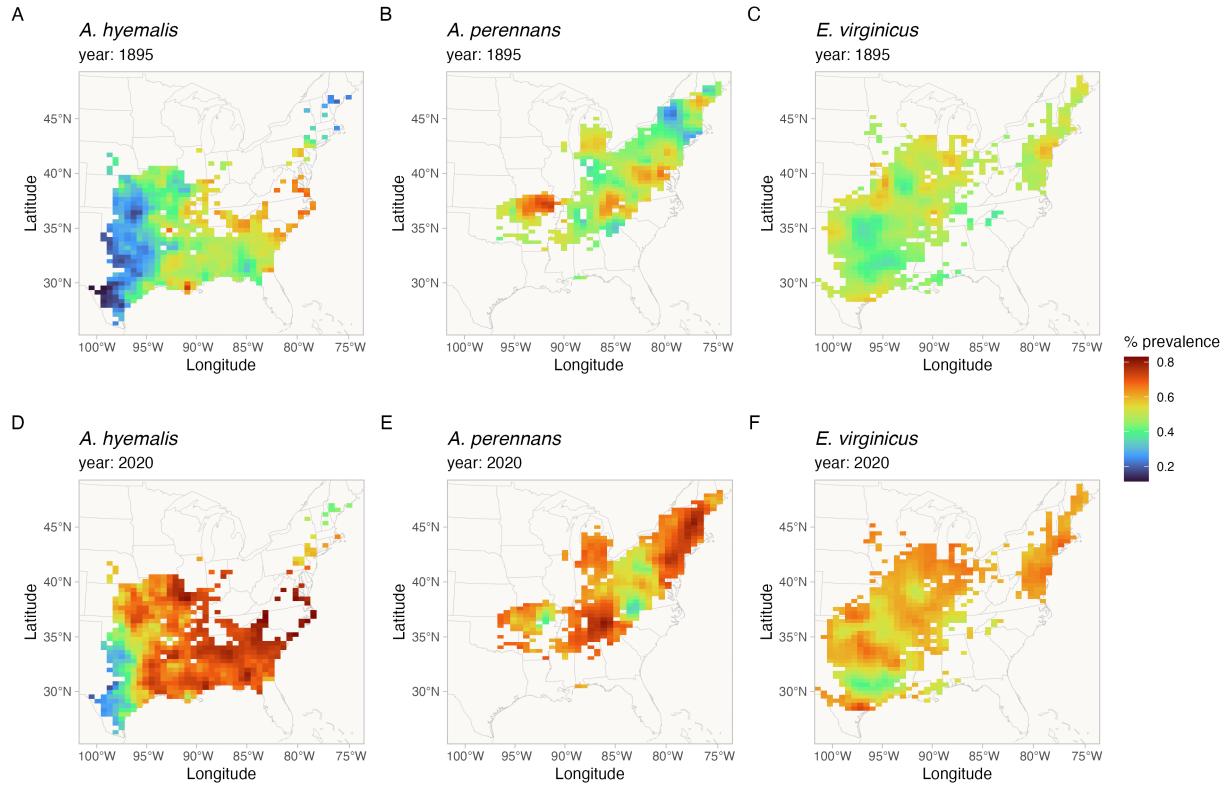


Figure 4: Predicted endophyte prevalence for each host species in 1895 and 2020. Predictions of prevalence come from the endophyte prevalence model. Color indicates the posterior mean endophyte prevalence for *A. hyemalis* (A,D), *A. perennans* (B,E), and *E. virginicus* (C,F). Map lines delineate study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

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

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

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

401 Changes in endophyte prevalence across the ranges of *A. hyemalis* and *E. virginicus* were less
402 strongly driven by changes in climate. Like *A. perennans*, climate during peak growing season
403 (spring for *A. perennans* and summer for *E. virginicus*) emerged most commonly as drivers of
404 changes in endophyte prevalence. Increases in mean spring precipitation were the strongest predictor
405 of increasing trends in endophyte prevalence for *A. hyemalis* (Fig. 5B) (greater than 99% posterior
406 probability of a positive slope). For this species, an increase of 100 mm in spring precipitation was
407 associated with 0.6% per year stronger increases in endophyte prevalence relative to regions with
408 no change in precipitation. The next greatest slopes were those associated with variability in spring
409 precipitation (greater than 96% posterior probability of a negative slope), as well as in the mean
410 and variability of autumn climate (greater than 98% probability of negative and positive slopes,
411 respectively). Changes in endophyte prevalence in *E. virginicus* were not strongly associated with
412 changes in most climate drivers, but regions with reduced variability in autumn precipitation (Fig.
413 5B) and with cooler and more variable summer temperatures (Fig. 5A,C) experienced the largest

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

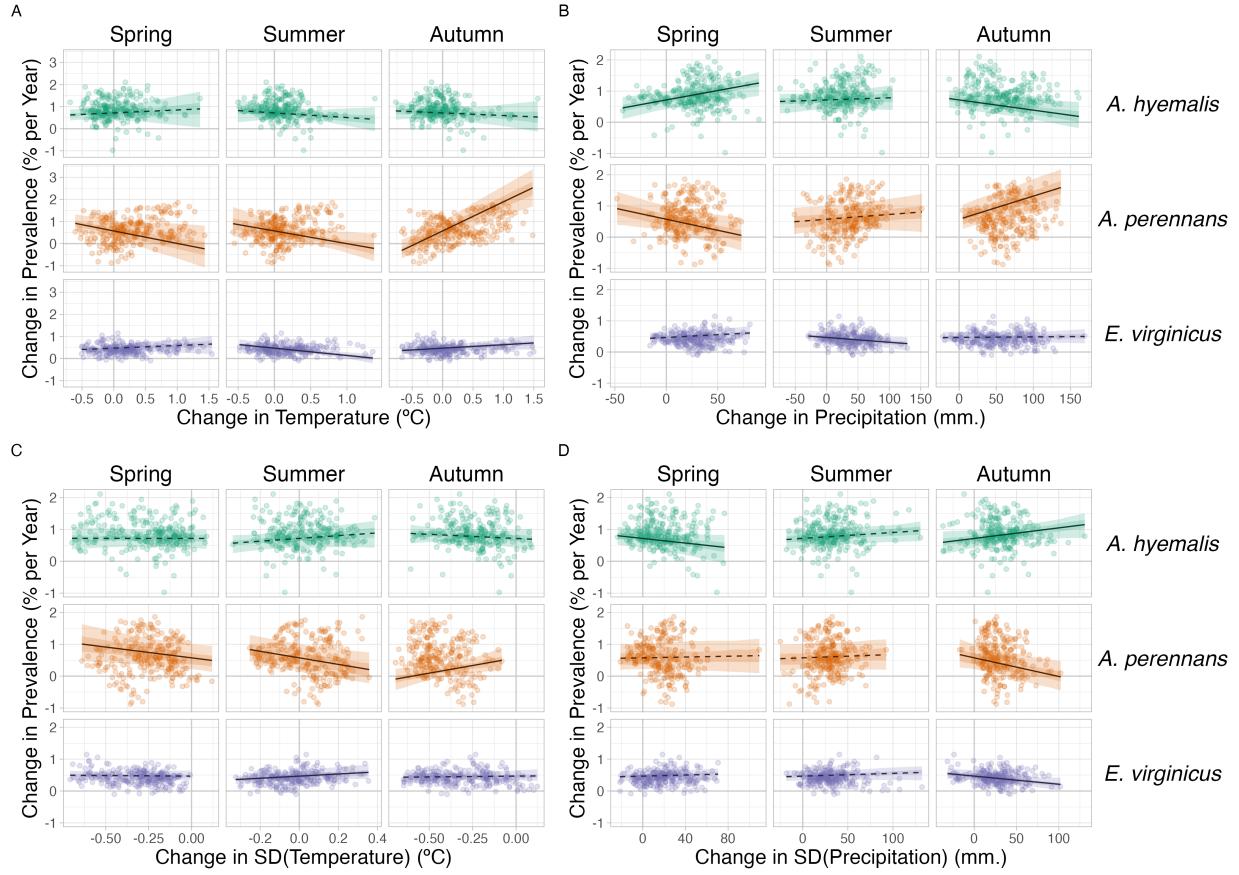


Figure 5: Relationships between predicted trends in endophyte prevalence and changes in seasonal climate drivers. Lines show marginal predicted relationship between spatially-varying trends in endophyte prevalence and changes in mean and variability of climate ((A): mean temperature, (B): cumulative 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.

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

421 Tests of the endophyte prevalence model's predictive performance as quantified by AUC and by
422 visual posterior predictive checks, indicated good predictive ability. Model performance was similar
423 between historic herbarium specimens used as training data and the out-of-sample test data from
424 contemporary surveys (AUC = 0.79 and 0.77 respectively; Fig. A5-A4). The model successfully
425 captured broad regional trends in endophyte prevalence seen in the contemporary survey data,
426 such as decline endophyte prevalence in *A. hyemalis* towards western longitudes (Fig. 6A) and **an**
427 **increase towards** northern latitudes (Fig. 6B). **It is noteable that** model predictions for endophyte
428 prevalence exhibited relatively little local geographic variation, whereas the out-of-sample survey
429 data were highly variable with populations spanning 0% to 100% endophyte-symbiotic plants (Fig.
430 6C), **indicating population-to-population variation not captured in the endophyte prevalence model.**

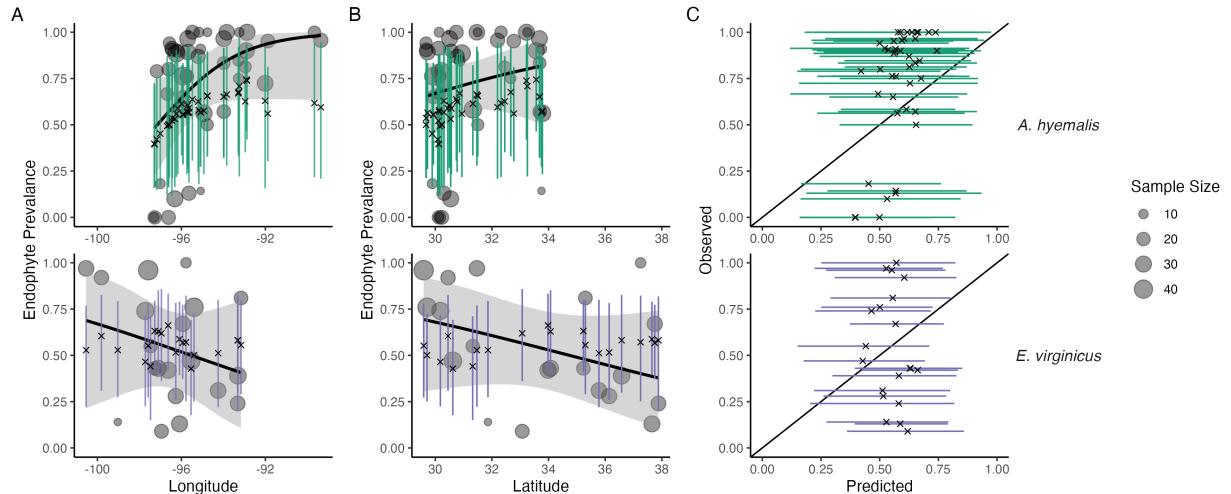


Figure 6: **Predictive performance for contemporary test data.** (A) The endophyte prevalence model, trained on historic herbarium collection data, performed modestly at predicting prevalence in contemporary population surveys. The model captured regional trends across (A) longitude and (B) latitude. 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

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

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

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

451 Based on previous work demonstrating that endophytes can shield their hosts from drought
452 stress (reviewed in Decunta et al. (2021)), we generally predicted that drought conditions would be
453 a driver of increasing endophyte prevalence. In contrast to this expectation, increasing prevalence
454 for *A. perennans* was associated with **both** increasing autumn temperature and precipitation (Fig.
455 5). To our knowledge, the response of the symbiosis in *A. perennans* to drought has not been
456 examined experimentally, but in a greenhouse experiment, endophytes had a positive effect on host
457 reproduction under shaded, low-light conditions (Davitt et al., 2010). Our results also hint that it
458 may be useful to investigate whether lagged climate effects are important predictors of host fitness
459 in this system (Evers et al., 2021). Endophyte prevalence of the autumn-flowering *A. perennans* was
460 strongly linked with decreasing spring precipitation, and that of the spring-flowering *A. hyemalis*
461 was associated with decreasing autumn precipitation (Fig. 5B). For *A. hyemalis*, endophytes could
462 be playing a role helping hosts weather autumn-season droughts, which **is likely also** an important
463 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

⁴⁹¹ of these fitness benefits playing out across the heterogeneous map of a changing climate and other
⁴⁹² anthropogenic drivers.

⁴⁹³ Our results indicate that *Epichloë* symbiosis has likely improved host fitness in stressful envi-
⁴⁹⁴ ronments leading to increasing prevalence. What is less clear is how this will influence future range
⁴⁹⁵ shifts. Based on our analysis, it is likely that the symbiosis will facilitate range shifts for hosts by
⁴⁹⁶ improving population growth at range edges. Previous population surveys (Rudgers and Swafford,
⁴⁹⁷ 2009; Semmarin et al., 2015; Sneck et al., 2017) attributed environment-dependent gradients in
⁴⁹⁸ endophyte prevalence to symbiont-derived fitness benefit's allowing hosts to persist in environments
⁴⁹⁹ where they otherwise could not (Afkhami et al., 2014; Kazenel et al., 2015). However, symbiont-
⁵⁰⁰ facilitated range shifts require that endophytes be present in the populations to be able to support
⁵⁰¹ population growth. The arid western range edge of *A. hyemalis* has had historically low endophyte
⁵⁰² prevalence (Fig. 4), and while prevalence has increased most quickly in the regions with histor-
⁵⁰³ ically low endophyte prevalence (Fig. A11), the complete absence of endophytes at range edges
⁵⁰⁴ in this species would make it impossible for prevalence to increase without dispersal of symbiotic
⁵⁰⁵ seeds (Fowler et al., 2023). These factors potentially contribute to the ability of the host species
⁵⁰⁶ to track its environmental niche. Another interesting question is the degree to which symbiotic
⁵⁰⁷ and non-symbiotic hosts, which occupy overlapping but distinct niches, are likely to experience
⁵⁰⁸ distribution shifts in tandem or at different rates in future. More extreme climate stresses, which
⁵⁰⁹ are expected more frequently in the future (Seneviratne et al., 2021), have the potential to alter
⁵¹⁰ the costs and benefits of the interaction. The past indicates a resilient symbiosis, but it will be
⁵¹¹ crucial to understand whether more extreme future climate conditions could tip this interaction to
⁵¹² deteriorate.

⁵¹³ *Steps towards forecasts of host-microbe symbioses*

⁵¹⁴ The combination of a spatially-explicit model and historic herbarium specimens allowed us to
⁵¹⁵ identify regions of both increasing and decreasing endophyte prevalence, however we see several
⁵¹⁶ next steps toward the goal of predicting host and symbiont niche-shifts in response to future cli-

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

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

544 *Herbaria for global change research*

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

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

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

596 *Conclusion*

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

614 **Acknowledgments**

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

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

623 **Statement of Authorship**

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

627 **Data and Code Availability**

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

632

Literature Cited

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

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

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

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

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

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

- 769 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
- 771 J J Heberling and D J Burke. Utilizing herbarium specimens to quantify historical mycorrhizal
772 communities. *Applications in plant sciences*, 7(4):e01223, 2019.
- 773 R J Hijmans, S Phillips, J Leathwick, J Elith, and Robert J Hijmans. Package ‘dismo’. *Circles*, 9
774 (1):1–68, 2017.
- 775 J HilleRisLambers, M A Harsch, A K Ettinger, K R Ford, and E J Theobald. How will biotic
776 interactions influence climate change-induced range shifts? *Annals of the New York Academy of
777 Sciences*, 1297(1):112–125, 2013.
- 778 IPCC. Climate change 2021: The physical science basis, 2021. URL <https://www.ipcc.ch/report/ar6/wg1/>.
- 780 N J B Isaac, M A Jarzyna, P Keil, L I Dambly, P H Boersch-Supan, E Browning, S N Freeman,
781 N Golding, G Guillera-Arroita, P A Henrys, et al. Data integration for large-scale models of
782 species distributions. *Trends in Ecology & Evolution*, 35(1):56–67, 2020.
- 783 A Jiménez-Valverde. Insights into the area under the receiver operating characteristic curve (auc)
784 as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*,
785 21(4):498–507, 2012.
- 786 D Kahle and H Wickham. Package ‘ggmap’. *Retrieved September*, 5:2021, 2019.
- 787 M R Kazenel, C L Debban, L Ranelli, W Q Hendricks, Y A Chung, T H Pendergast IV, N D
788 Charlton, C A Young, and J A Rudgers. A mutualistic endophyte alters the niche dimensions of
789 its host plant. *AoB plants*, 7:plv005, 2015.
- 790 R A Knapp, G M Fellers, P M Kleeman, D AW Miller, V T Vredenburg, E B Rosenblum, and C J
791 Briggs. Large-scale recovery of an endangered amphibian despite ongoing exposure to multiple
792 stressors. *Proceedings of the National Academy of Sciences*, 113(42):11889–11894, 2016.

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

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

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

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

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

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

934

Appendix A

935

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

938 **Authors:**

939 Joshua C. Fowler^{1,2*}

940 Jacob Moutouama¹

941 Tom E. X. Miller¹

942

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

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

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

946 **Contents:**

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

948

Supplemental Figures

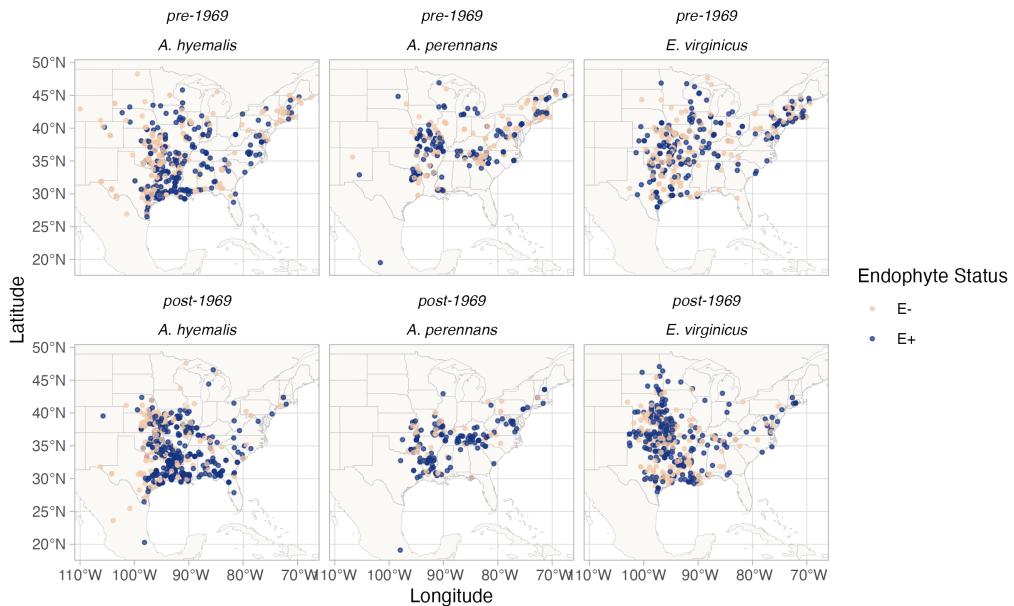


Figure A1: Endophyte presence/absence in specimens of each host species. Points show collection locations colored according to whether the specimen contained endophytes (E+; blue points) or did not contain endophytes (E-, tan points). To visualize temporal change, the data are faceted before and after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

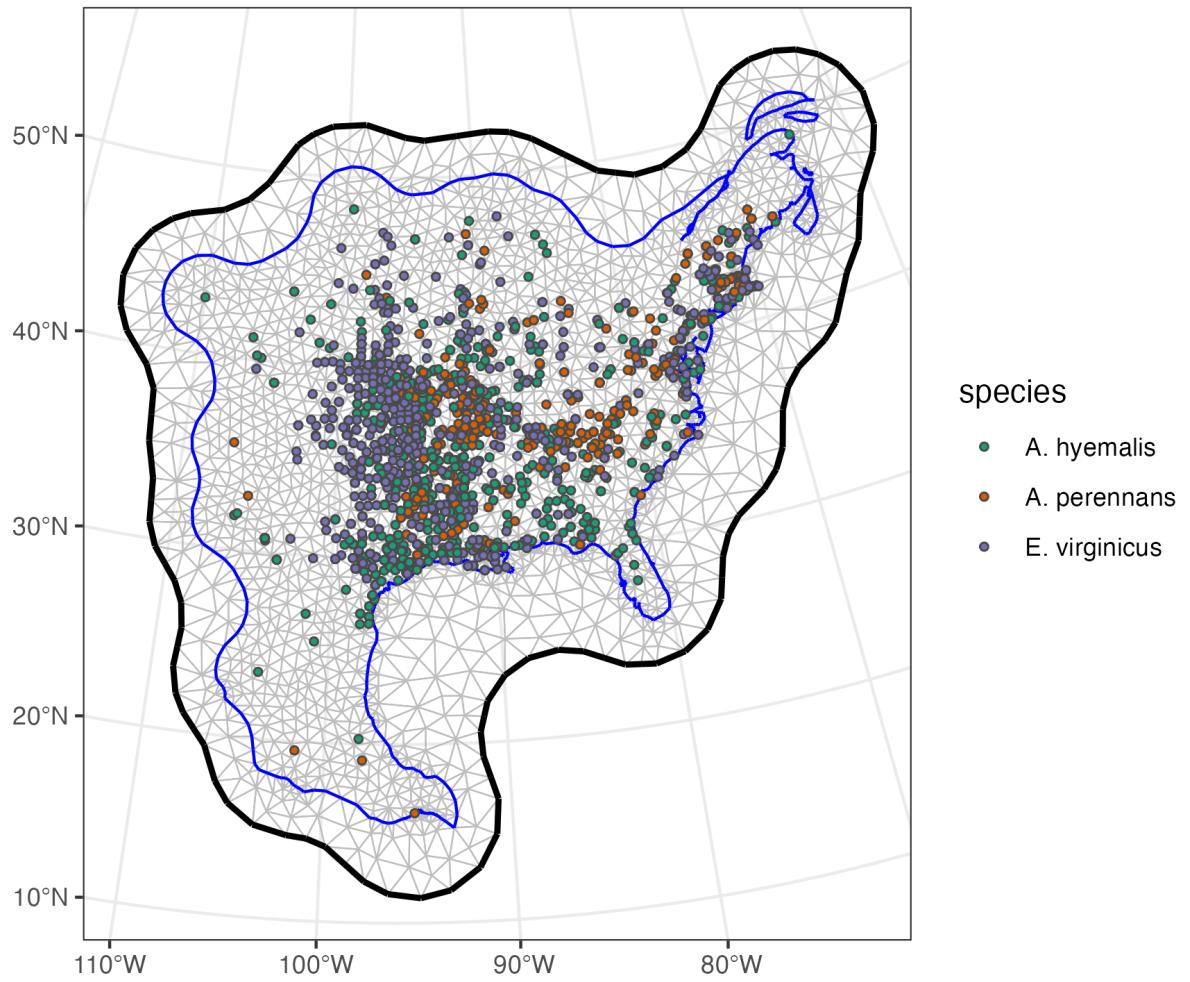


Figure A2: Triangulation mesh used to estimate spatial dependence between data points.

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.

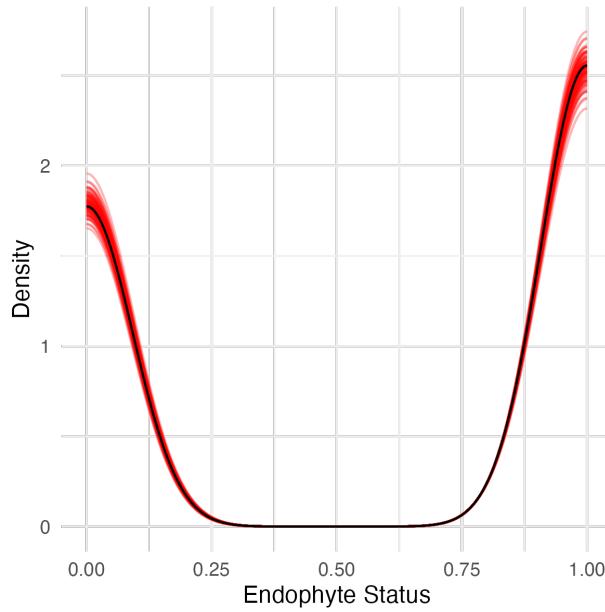


Figure A3: **Graphical posterior predictive check of the endophyte prevalence model fit.**

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).

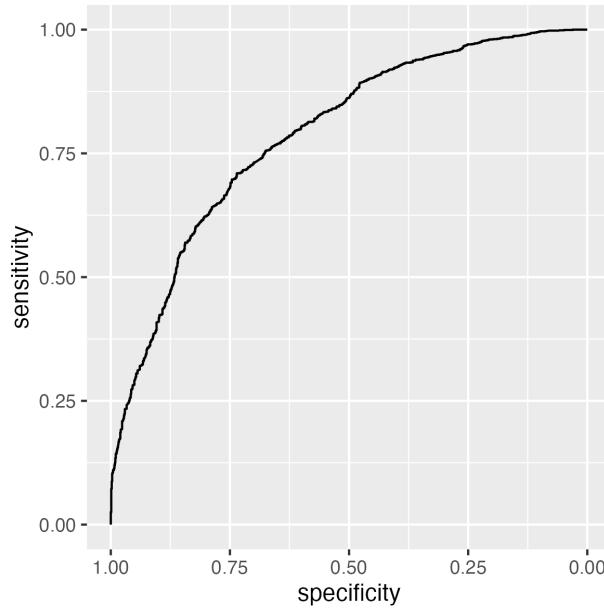


Figure A4: **ROC plot showing performance of the endophyte prevalence model in classifying observations according to endophyte status within the in-sample training data from herbarium collections.** The curves show adequate model performance for observed data. The AUC value is 0.79.

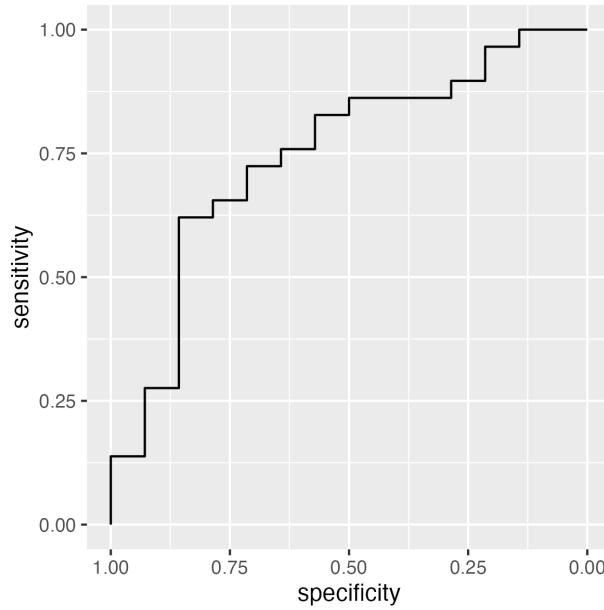


Figure A5: **ROC plot showing performance of the endophyte prevalence model in classifying observations according to endophyte status within the out-of-sample test data from contemporary surveys.** The curves show adequate model performance for test data. The AUC value is 0.77.

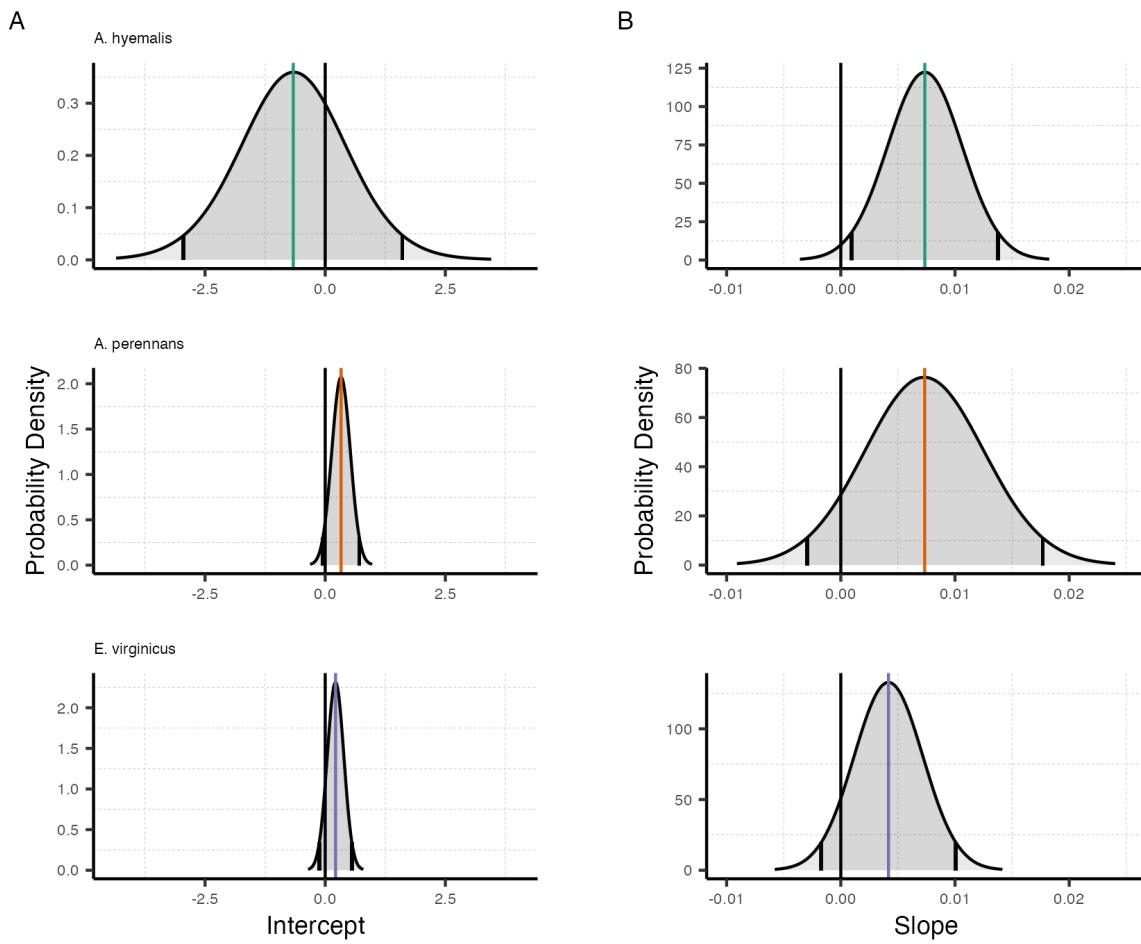


Figure A6: Posterior estimates of parameters describing global intercept and temporal trends from the endophyte prevalence model. Density curves show the probability density along with mean (colored line) and 95% CI (black lines) for the (A) intercept and (B) slope terms, A and T respectively from Eqn.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

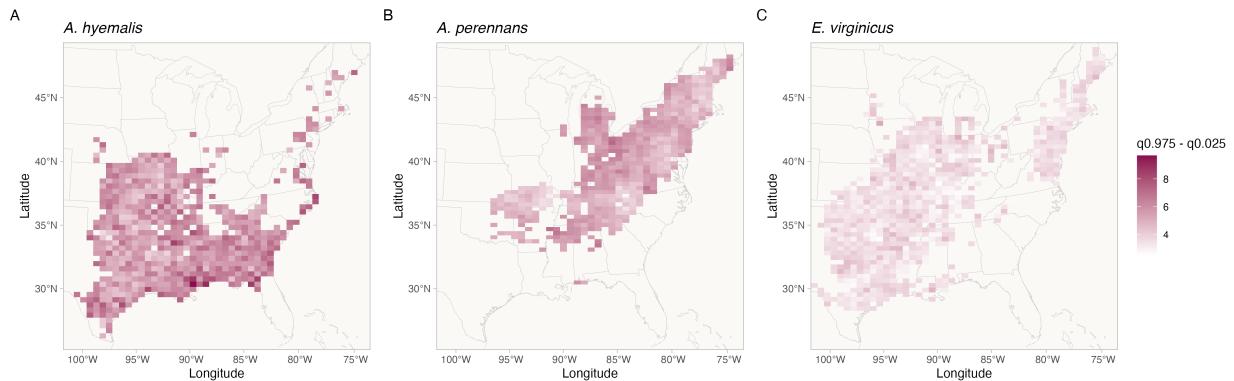


Figure A7: Credible interval width of temporal trends in endophyte prevalence across the distribution of each host species estimated from the endophyte prevalence model.

Shading represents the range of the 95% posterior credible interval given in units of $\%$ change in prevalence/year for spatially varying slopes, τ from Eqn. 1. Map lines delineate study areas and do not necessarily depict accepted national boundaries.

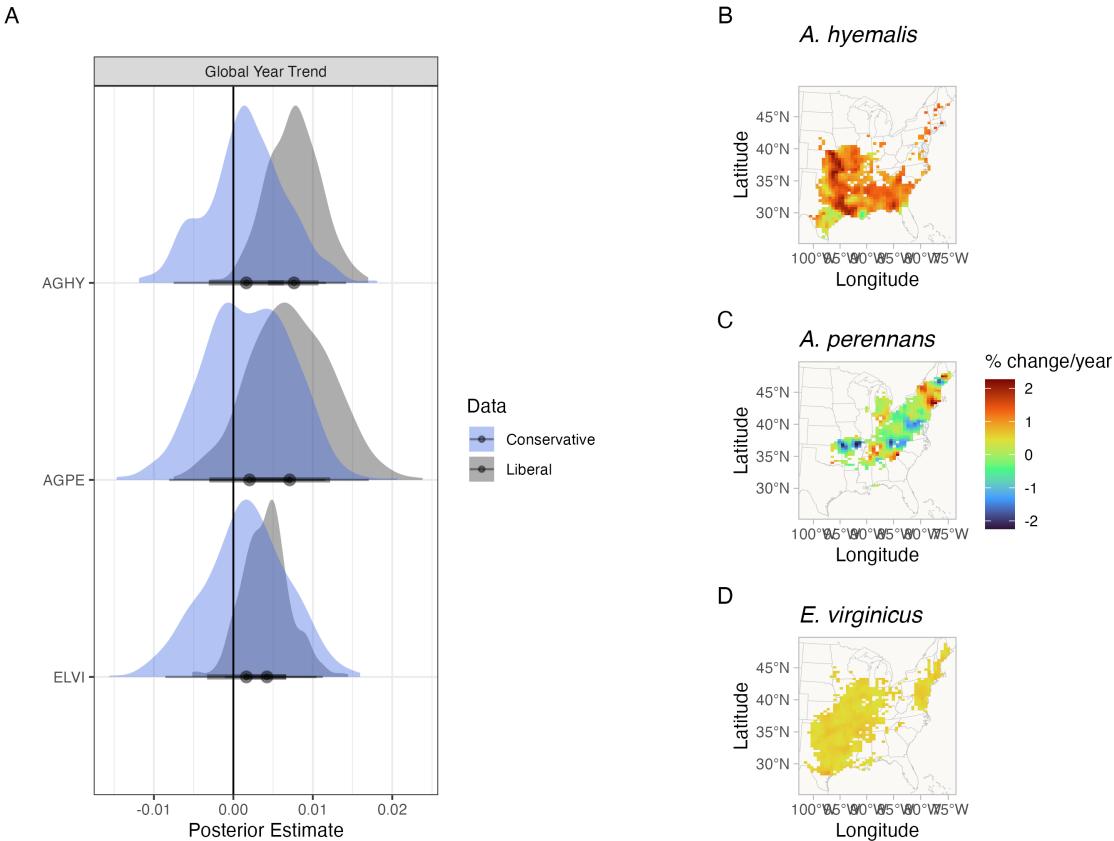


Figure A8: Comparison of endophyte prevalence model estimates fit to data with liberal versus conservative endophyte scores. Liberal and conservative scores document uncertainty in the endophyte identification process. Each specimen was given both a liberal and conservative scores. In cases of uncertain 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 (τ 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.

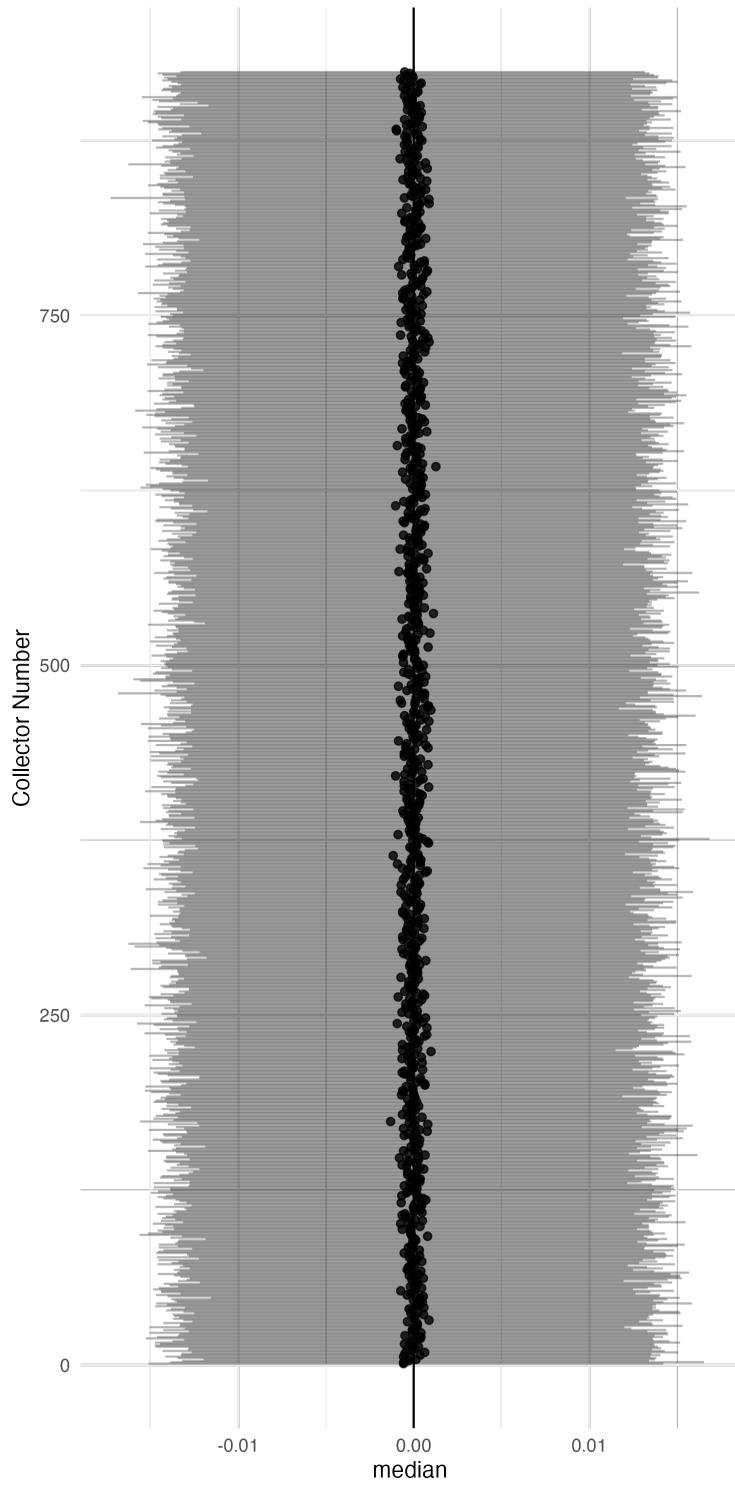


Figure A9: **Posterior estimates of collector random effects from endophyte prevalence model.** Collector random effects are denoted χ in Eqn. 1 and represent variance associated with researchers who collected historic herbarium specimens. Points show posterior median along with 95% CI for each of 924 individual collectors.

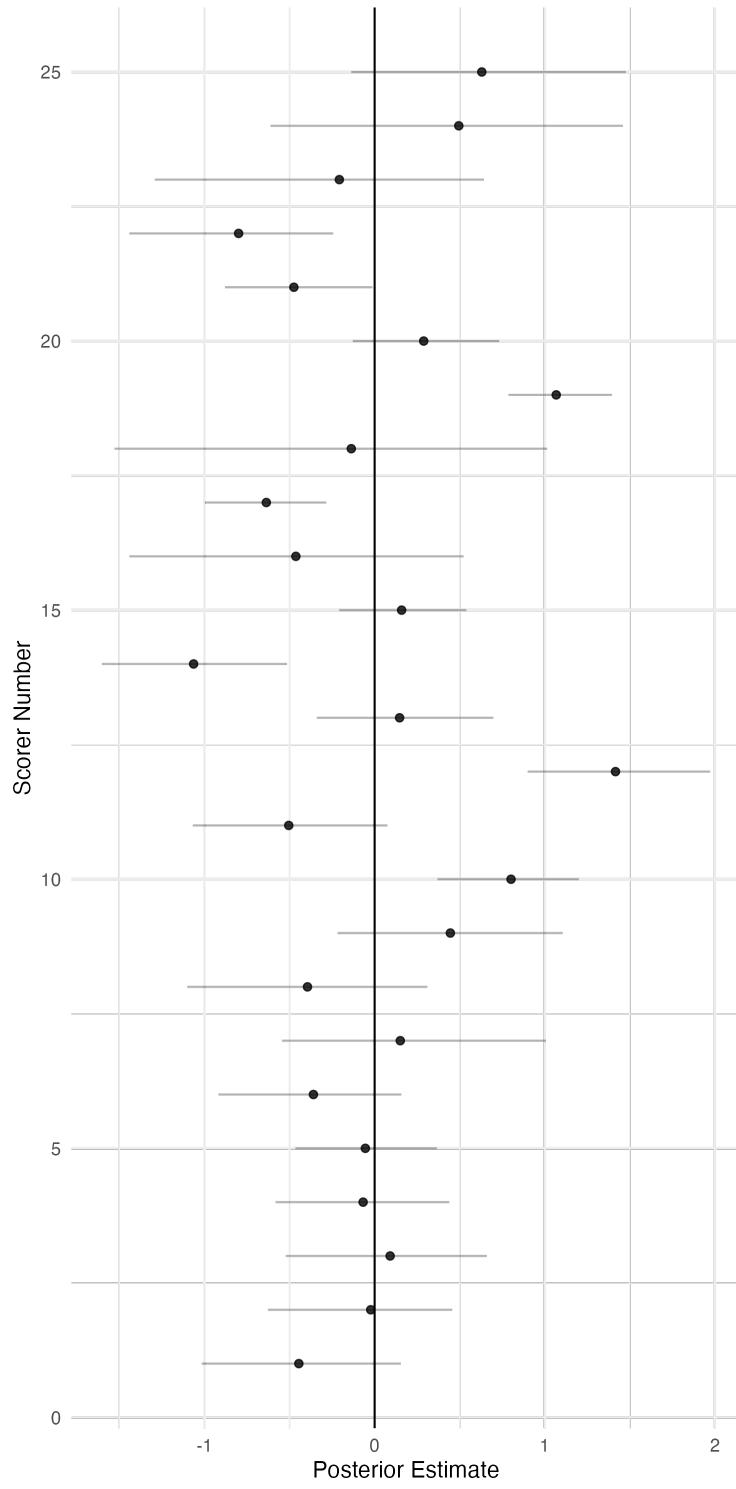


Figure A10: **Posterior estimates of scorer random effects from endophyte prevalence model.** Scorer random effects are denoted ω in Eqn. 1 and represent variance associated with researchers who identified *Epichloë* endophytes within herbarium specimen tissue samples. Points show posterior median along with 95% CI for each of 25 individual scorers.

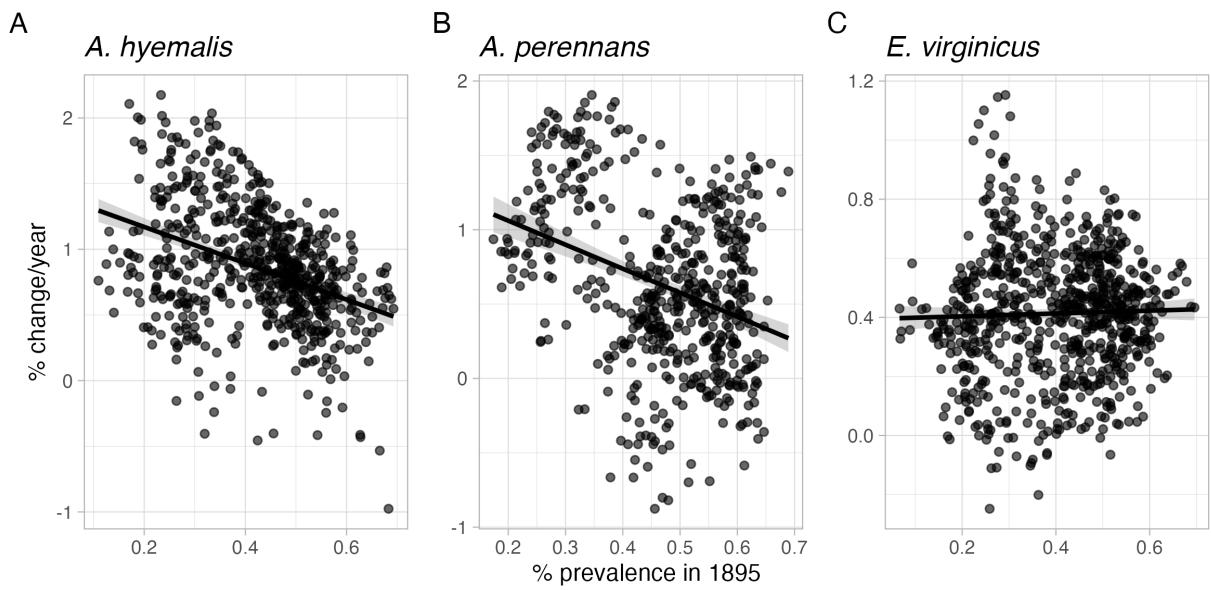


Figure A11: Relationship between initial prevalence and temporal trends in prevalence estimated from the endophyte prevalence model. Points show predicted posterior mean temporal trend for each species at pixels across each host distribution ((A) *A. hyemalis*, (B) *A. perennans*, and (C) *E. virginicus*).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

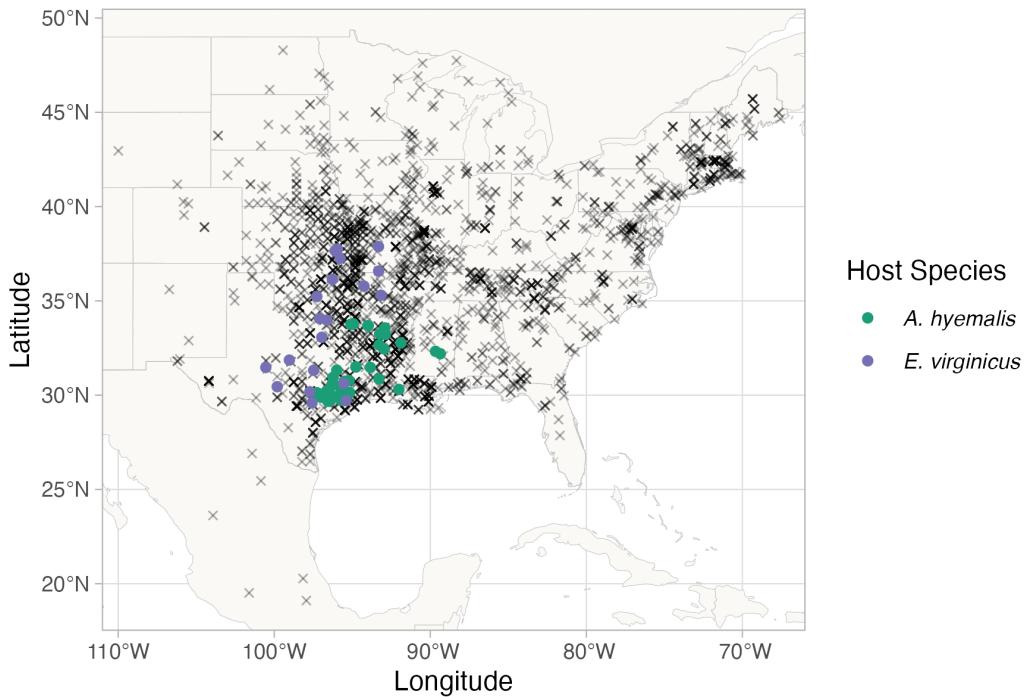


Figure A12: **Locations of contemporary surveys of endophytes used as "test" data to evaluate predictive ability of the endophyte prevalence model.** Points are locations of host populations surveyed between 2013 and 2019 for endophytes, colored by species (*A. hyemalis*: green, *E. virginicus*: purple).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

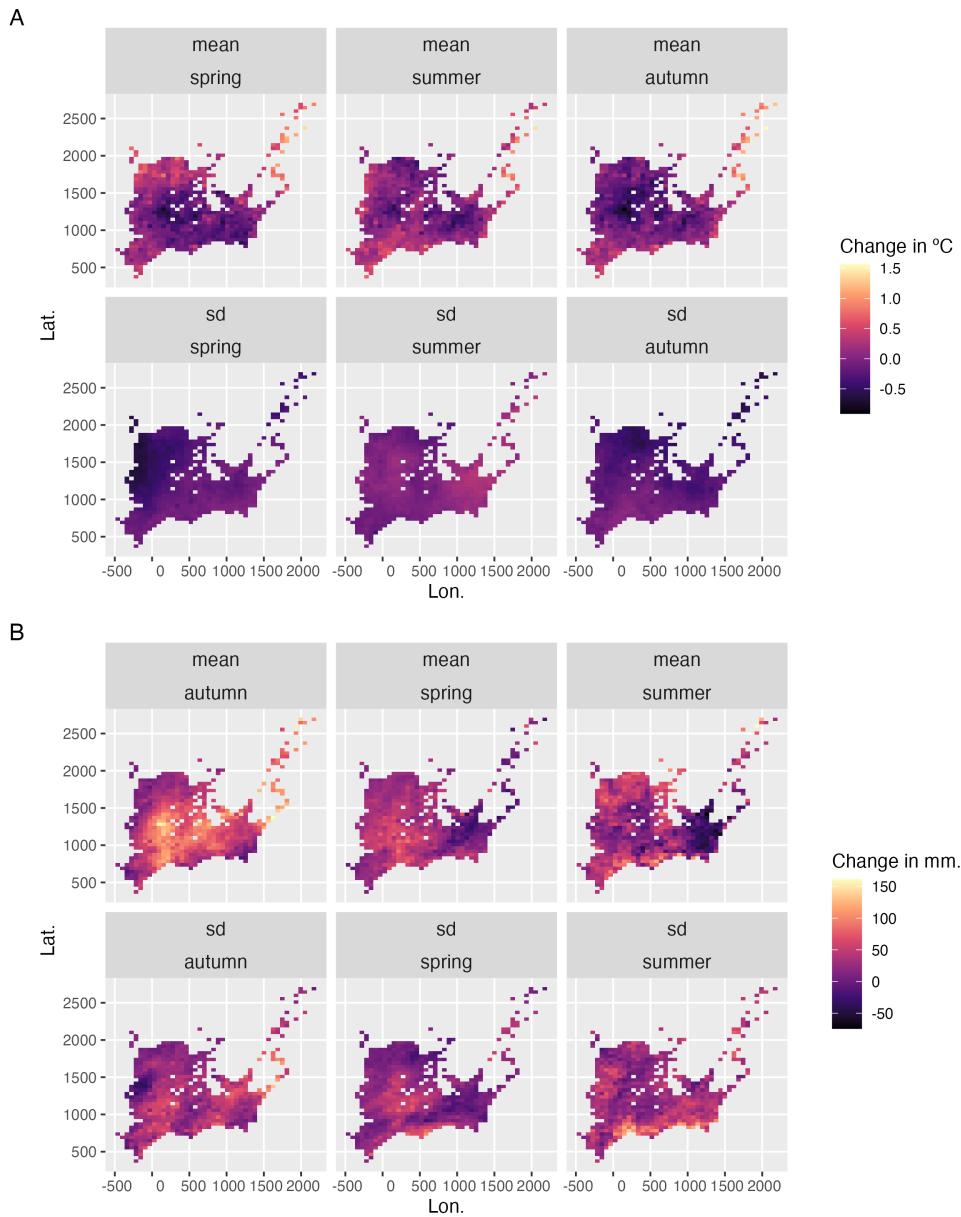


Figure A13: **Change in seasonal climate variables between the periods 1895-1925 and 1990-2020 across the distribution of *A. hyemalis*.** Color represents change in (A) seasonal temperature ($^{\circ}\text{C}$) and (B) seasonal precipitation (mm.). Maps show pixels covering the modeled distribution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

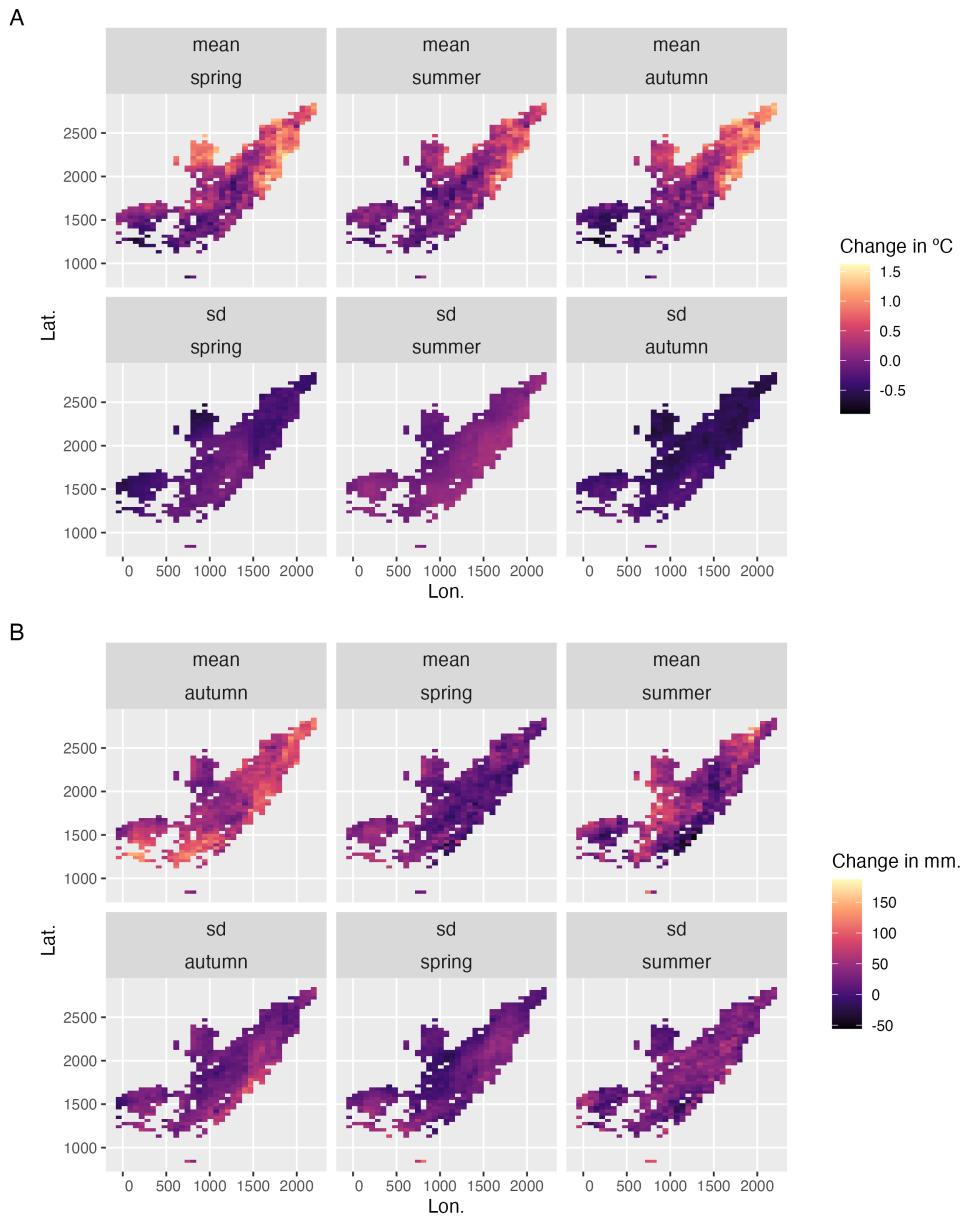


Figure A14: **Change in seasonal climate variables between the periods 1895-1925 and 1990-2020 across the distribution of *A. perennans*.** Color represents change in (A) seasonal temperature ($^{\circ}\text{C}$) and (B) seasonal precipitation (mm.). Maps show pixels covering the modeled distribution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

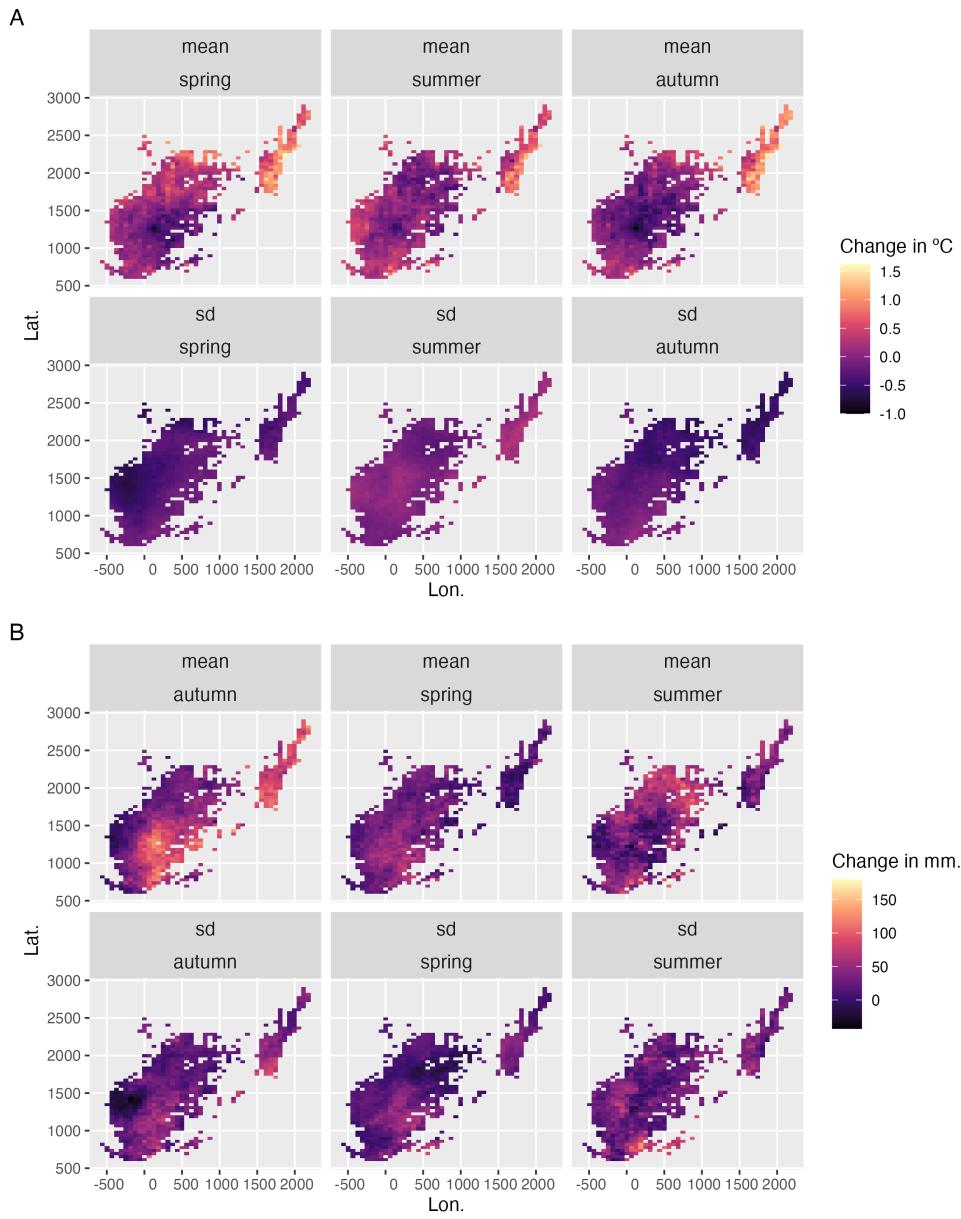


Figure A15: **Change in seasonal climate variables between the periods 1895-1925 and 1990-2020 across the distribution of *E. virginicus*.** Color represents change in (A) seasonal temperature ($^{\circ}\text{C}$) and (B) seasonal precipitation (mm.). Maps show pixels covering the modeled distribution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

950

Supporting Methods

951

ODMAP Protocol

952 [Overview](#)

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

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

955 **Study area:** Eastern North America

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

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

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

959 **Boundary:** Natural.

960 [Data](#)

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

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

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

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

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

967 [Model](#)

968 **Model assumption:** We assumed that the target species are at equilibrium with their environment.
969

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

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

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

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

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

977 [Assessment](#)

978 We used AUC to test model performance.

979 [Prediction](#)

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

981 *Mesh and Prior Sensitivity Analysis*

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

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

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

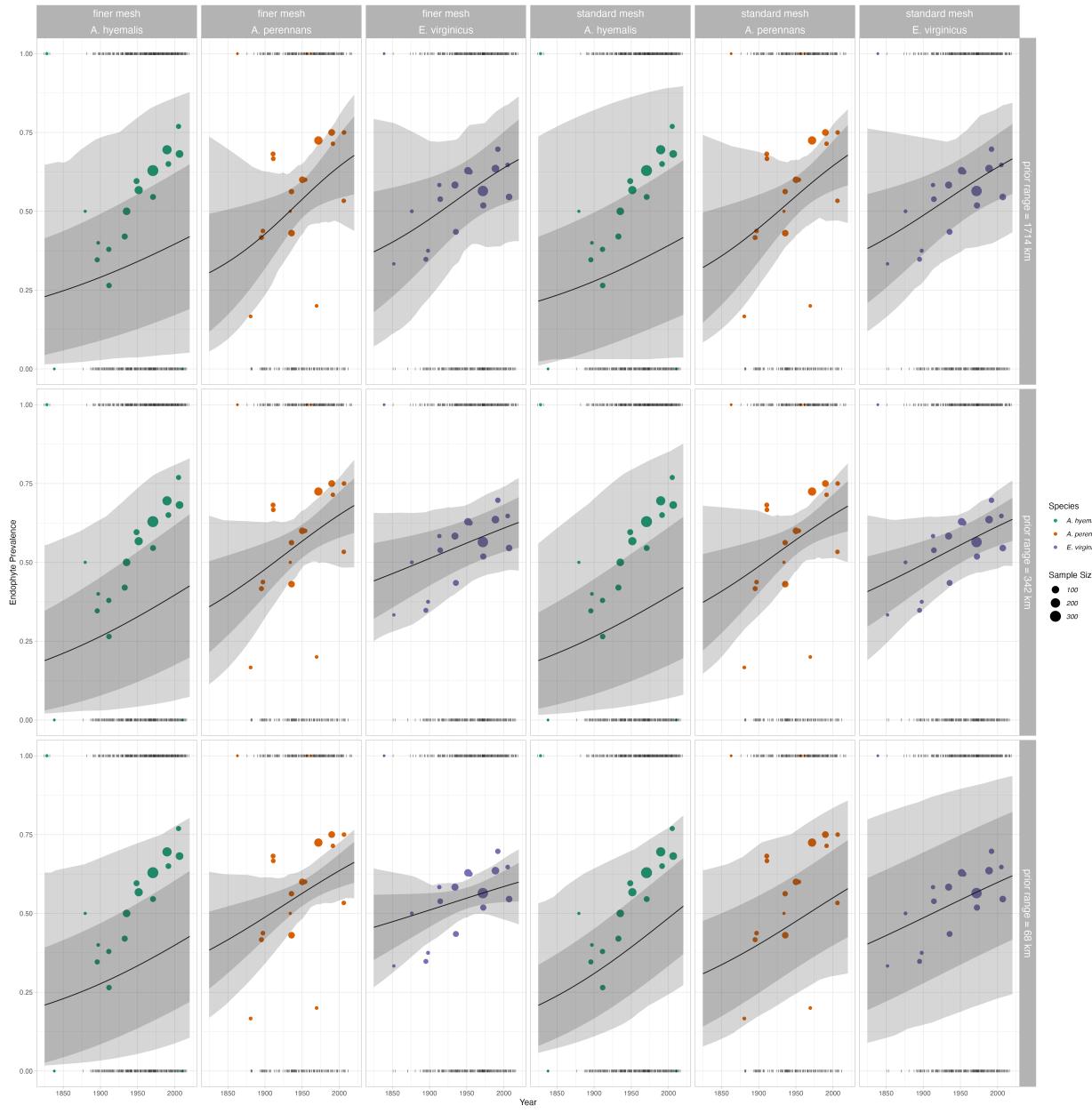


Figure A16: Overall trend in endophyte prevalence evaluated for endophyte prevalence models with different range priors on spatially structured random effects, and for two different triangulation meshes. Data used in model fitting is the same across all panels and as in the main text. Note that these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

992 For spatially-varying temporal trends, we found that models with different priors predicted
993 consistent spatial patterns in temporal trends, although the range of this prediction varied depending
994 on the prior and mesh (Fig. A17 - A18). One noteworthy result of this analysis is that combinations
995 of prior choice and mesh can introduce instability in model fitting. This is evident in A17 panel B
996 and A18 panel B, where the prior range is smaller than the minimum vertex length of the mesh.
997 Model fitting takes an extended time period and the model struggles to identify variation across
998 space. Results with a set of prior ranges (Fig. A17 - A and C; Fig. A18 - A and C) result in
999 models that estimate trends across space of the same direction and order of magnitude, although
1000 the "smoothness" of these predictions vary.

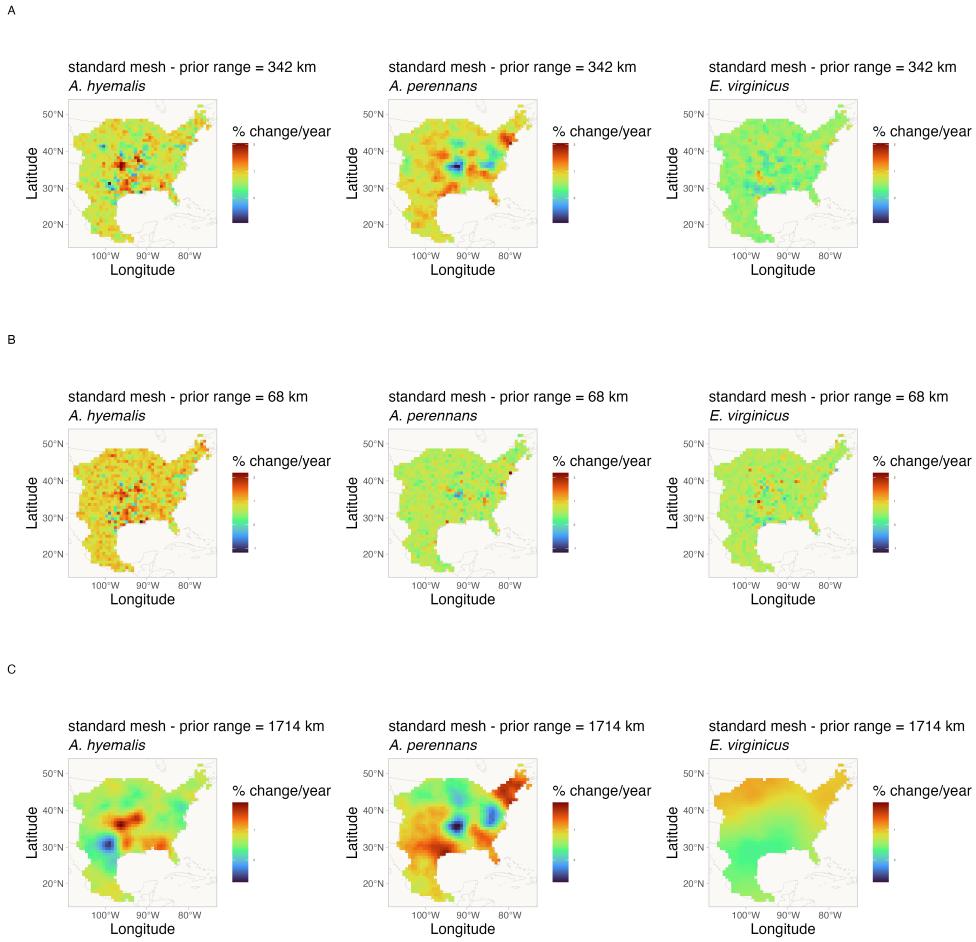


Figure A17: Spatially-varying trends in endophyte prevalence evaluated for **the endophyte prevalence model with different range priors on spatially structured random effects, and for the "standard" mesh**. Data used in model fitting is the same across all panels and as in the main text. Shading indicates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

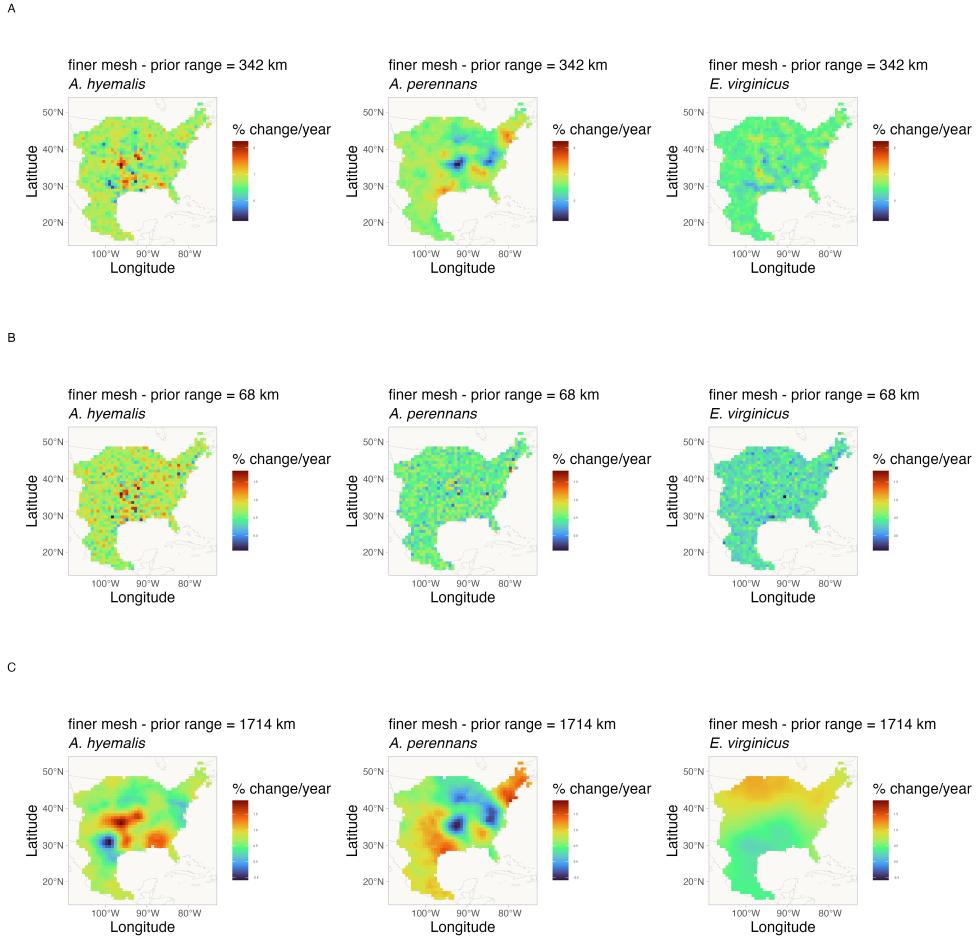


Figure A18: Spatially-varying trends in endophyte prevalence evaluated for the endophyte prevalence model with different range priors on spatially structured random effects, and for the "finer" mesh. Data used in model fitting is the same across all panels and as in the main text. Shading indicates the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

1001

Spatially-biased Sample Size Simulation Analysis

1002 To examine how data that is unevenly distributed across host distributions may influence interpreta-
1003 tion of spatially-varying coefficients, we performed a simulation analysis. Our focal species, *Agrostis*
1004 *hyemalis*, *Agrostis perennans*, and *Elymus virginicus*, are widely distributed grasses across the east-
1005 ern United States that host *Epichloë* fungal endophytes. For logistical reasons, our sampling visits
1006 to herbaria focused on herbaria in the central southern U.S., which resulted in unevenly distributed
1007 data across each host species' range. This is particularly noteable for *Agrostis perennans* which has
1008 the most northern distribution and relatively fewer total collected specimens compared to the other
1009 focal species. Thus, a significant portion in the northeast of this species' range is relatively sparsely
1010 sampled. Our analysis presented in the main text identified this region as having strong increase in
1011 endophyte prevalence. Future visits to herbaria with regional focuses in the Northeastern US would
1012 certainly garner new specimens that could provide valuable insights into shifting host and symbiont
1013 distributions.

1014 *Simulation of spatially-biased symbiont occurrence data*

1015 We simulated datasets with varying levels of missing-ness to examine how this missing-ness influ-
1016 enced the estimation of spatially-varying trend estimates. We first generated 300 data points for
1017 each of three hypothetical species at random positions across an area approximating the scale of
1018 our focal data. Each data point was randomly assigned a year of collection across 200 years. We
1019 then simulated data from a Bernoulli process with trends alternating across nine regions (Fig. A19)
1020 in a 3X3 grid pattern. This grid pattern was intended to create a complex spatial layout of trends,
1021 where trends were either an increase of 1% per year or a decrease of 1% per year.

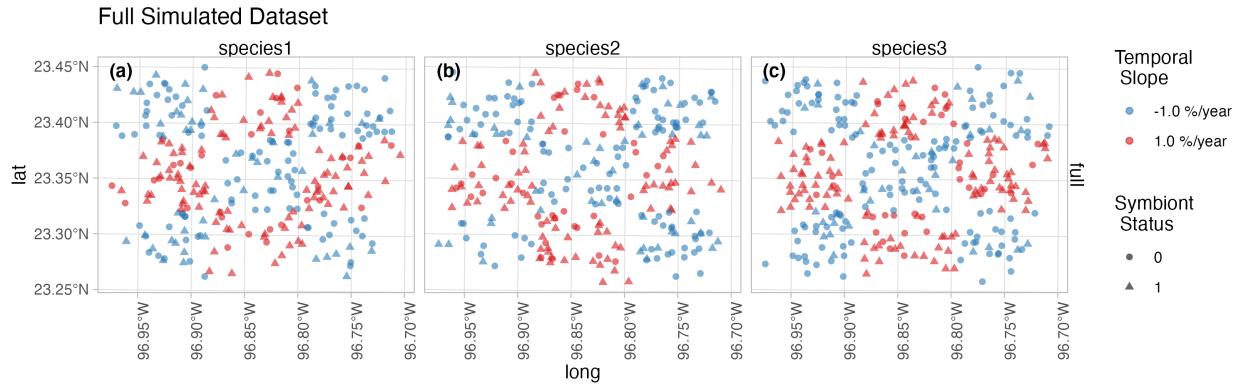


Figure A19: Full simulated dataset of symbiotic association with spatially-varying temporal trends. Color indicates the slope parameter used to simulate trends in endophyte status across nine "regions" for three species. Data are assigned collection years across a period of 200 years. Shape indicates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

1022 From this full data, we generated six additional datasets with missing-ness in the northeast
 1023 region of the simulated data for hypothetical species 2. The data remained the same for Species 1
 1024 and for species 3 across all datasets. For these six datasets, we removed data points at random in
 1025 six ways: 0% of datapoints in northeast region, 0% of recent datapoints, only 20% of datapoints,
 1026 only 20% of recent datapoints, only 50% of datapoints, and only 50% of recent datapoints (Fig.
 1027 A20). We define the datapoints as part of the recent time period if they occur later than the median
 1028 year. The result is 6 scenarios exploring degrees of spatial and temporal bias.

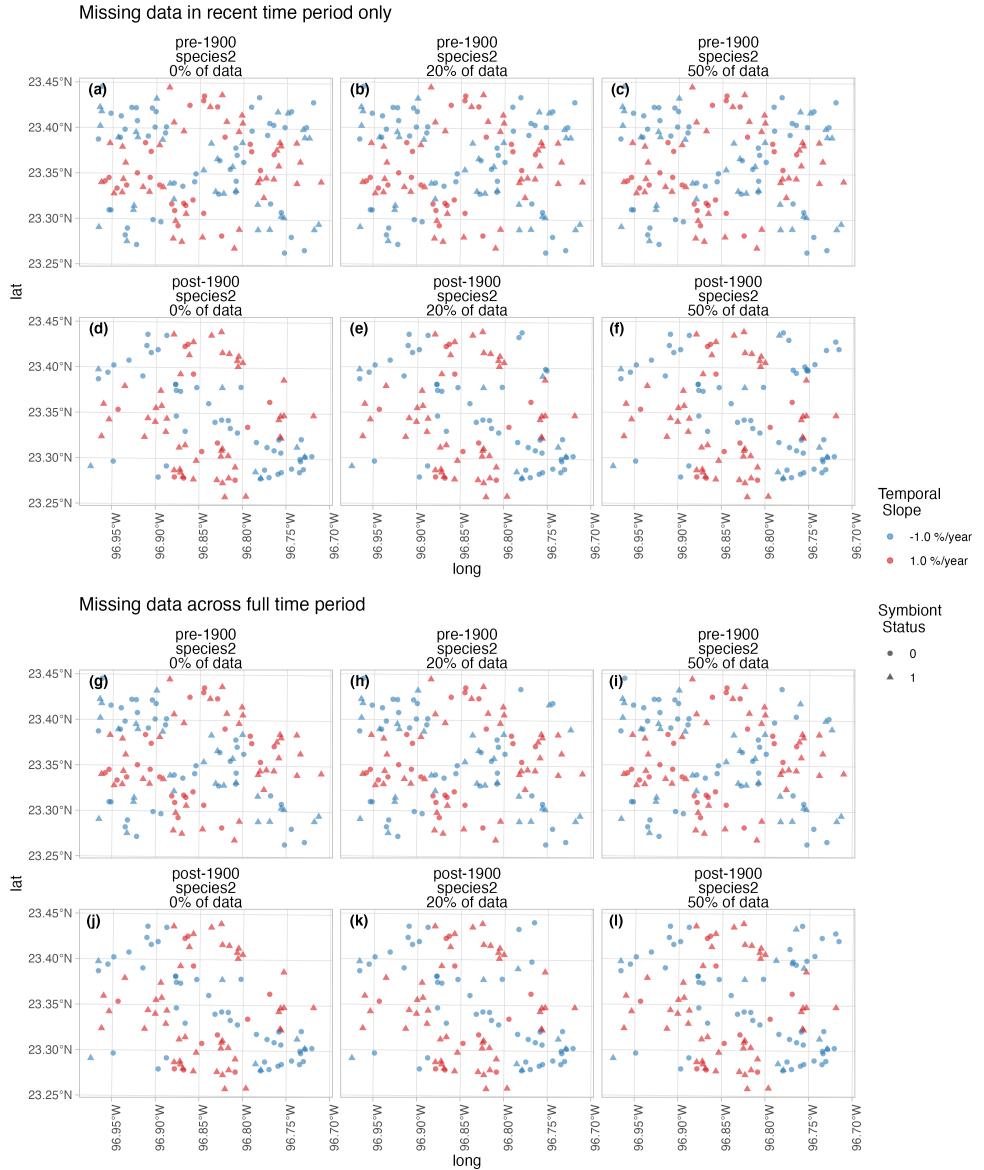


Figure A20: Six simulated datasets representing scenarios of spatially-baised missingness for Species 2. Missingness was imposed in the northeast region for six scenarios: 0% of recent datapoints available (a,d); only 20% of recent datapoints (b,e); only 50% of recent datapoints (c,f); 0% of datapoints 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.

1029 *Statistical analysis*

1030 We analyzed each dataset with a model given by Eqn. A1 similar in construction to that used in
1031 our central analysis.

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

1032 Where symbiont presence/absence of the i^{th} specimen ($P_{h,i}$) was modeled as a Bernoulli re-
1033 sponse variable with expected probability of symbiont occurrence $\hat{P}_{h,i}$ for each host species h . We
1034 modeled $\hat{P}_{h,i}$ as a linear function of intercept A_h and slope T_h defining the global trend in endophyte
1035 prevalence specific to each host species as well as with spatially-varying intercepts α_{h,l_i} and slopes
1036 τ_{h,l_i} associated with location (l_i , the unique latitude-longitude combination of the i th observation).
1037 Similar to the SVC model of our central analysis (Eqn. 1), we estimated a shared variance term
1038 with the spatially-dependent random effect δ_{l_i} , intended to account for residual spatial variation.
1039 However in this analysis we omit i.i.d.-random effects terms associated with collector and scorer
1040 identity (χ_{c_i} and ω_{s_i} in Eqn. 1) for the sake of simplicity.

1041 *Influence of spatially-biased sampling on model interpretation*

1042 Our analysis of the full simulated data shows that our model is suitably flexible to capture complex
1043 spatial patterns in temporal trends (Fig. A21 a-c). Beyond this, the model also qualitatively
1044 captures the spatial patterns in temporal trends even with large amounts of data missingness (i.e
1045 missing up to 80% of the datapoints (Fig. A21 p-r)).

1046 While this analysis is not an exhaustive examination of the influence of sampling bias on our
1047 results for several reasons (including not examining how different strengths in temporal trends,
1048 different spatial arrangements of missing-ness influence model estimates, or different sample sizes
1049 influence results), it demonstrates that the spatially-varying modelling framework implemented in
1050 INLA we employ can suitably recover regional trends even with significant spatially-bias within
1051 data collection, and further the analysis is likely robust to temporally-structured bias (missing data

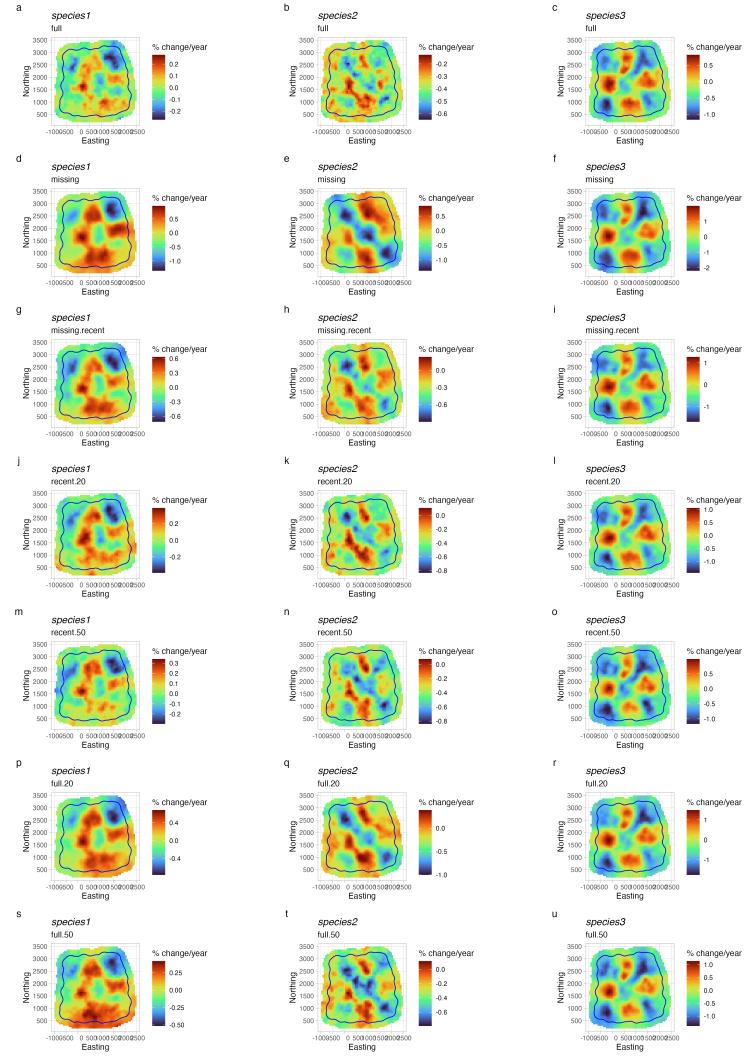


Figure A21: Mean predicted spatially-varying trend in symbiont prevalence across datasets with different levels of missingness. Color indicates the estimated mean temporal trend within each pixel across the simulated data. Panels show estimates for models fit to different levels of missing data for species 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.

1052 within recent collection period). Future work could more fully explore the scenarios that cause
1053 this ability to break down. We expect this simulation reflects what may be a common scenario for
1054 research investigating global change using natural history specimens. Collection effort by trained
1055 taxonomists and professional collectors peaked in the past, and collections contain relatively fewer
1056 modern specimens in many regions. Additionally, most global change research necessarily involves
1057 accessing many specimens across collections. Research efforts such as ours will be unable to access
1058 every specimen from all possible collections. Ongoing digitization efforts will make it possible to
1059 more clearly assess how much data is missing from a particular study compared to the actual
1060 holdings of natural history collections, but ultimately, the decision of what data and collections to
1061 include is a question of sample size and study design.