

Urban biogeography of fungal endophytes across San Francisco

Emma Gibson¹ and Naupaka Zimmerman¹

¹Department of Biology, University of San Francisco

Corresponding author:

Naupaka Zimmerman¹

Email address: nzimmerman@usfca.edu

ABSTRACT

In natural and agricultural systems, the plant microbiome—the microbial organisms associated with plant tissues and rhizosphere soils—has been shown to have important effects on host physiology and ecology, yet we know little about how these plant-microbe relationships play out in urban environments. Here we characterize the composition of fungal communities associated with living leaves of one of the most common sidewalk trees in the city of San Francisco, California. We focus our efforts on endophytic fungi (asymptomatic microfungi that live inside healthy leaves), which have been shown in other systems to have large ecological effects on the health of their plant hosts. Specifically, we characterized the foliar fungal microbiome of *Metrosideros excelsa* (Myrtaceae) trees growing in a variety of urban environmental conditions. We used high-throughput culturing, PCR, and Sanger sequencing of the internal transcribed spacer nuclear ribosomal DNA (ITS nrDNA) region to quantify the composition and structure of fungal communities growing within healthy leaves of 30 *M. excelsa* trees from 6 distinct sites, which were selected to capture the range of environmental conditions found within city limits. Sequencing resulted in 854 high-quality ITS sequences. These sequences clustered into 85 Operational Taxonomic Units (97% OTUs). We found that these communities encompass relatively high alpha (within) and beta (between-site) diversity. Because the communities are all from the same host tree species, and located in relatively close geographical proximity to one another, these analyses suggest that urban environmental factors such as heat islands or differences in vegetation or traffic density (and associated air quality) may potentially be influencing the composition of these fungal communities. These biogeographic patterns provide evidence that plant microbiomes in urban environments can be as dynamic and complex as their natural counterparts. As human populations continue to transition out of rural areas and into cities, understanding the factors that shape environmental microbial communities in urban ecosystems stands to become increasingly important.

INTRODUCTION

Although major cities only cover a small portion of the Earth's total geographic area, more than 50% of the human population lives in these urban centers, and these cities have major impacts on the biogeochemistry, hydrology, and climate of both their immediate surroundings and of the biosphere as a whole (Schneider, Friedl & Potere, 2009). With increasing urbanization, understanding the ecology of cities and urban settings has become critical to human health and well being. Despite their comparatively small geographic size, the high density of human populations in these environments makes them distinct ecosystems with their own unique dynamics (Sukopp, 1998). Urban environments represent the convergence of humans from around the world, any plant or animal species those humans might have brought with them, and infrastructures such as roads, sewers, and tall buildings. Despite the complexity that these ecosystems present, they are often overlooked by ecologists because more traditional ecology has focused on 'natural' systems. In urban systems, however, human influence is a primary ecological factor (McDonnell & Niemelä, 2011). In recent years, ecologists have begun studying the urban environment just as they would a natural environment, in order to understand the novel environmental conditions this setting presents to the organisms that live there (Wu, 2014).

In this study, we focus on the urban ecology of trees in San Francisco, California. Urban trees can play a major role in shaping the ecosystem of a city. Just as the trees in a forest have a considerable

48 impact on its climate and ecology, trees in cities can have notable effects on a city's environment. For
49 example, plant life in cities can impact temperature, air quality, and other aspects of human health (Willis
50 & Petrokofsky, 2017). The urban heat island effect, which occurs when 'islands' of heat form as heat
51 gets trapped between tall buildings, is one of the most well-documented unique urban anthropogenic
52 environmental conditions (Oke, 1973). Trees in urban environments have been shown to interact with
53 these city-specific environmental factors (Kong et al., 2014). For example, trees in cities have been
54 shown to improve urban air quality by taking up significant amounts of carbon dioxide from city air
55 (Nowak et al., 2014). As pollution generated in urban centers is one of the major contributing factors to
56 worldwide pollution, trees in urban environments may have a role in managing the environmental impacts
57 of urbanization (Alberti et al., 2003).

58 Due to their importance, there is growing interest in understanding and maintaining sustainable urban
59 ecologies. One potentially major factor influencing health of plants both in nature and cities is the plant
60 microbiome. As the widespread availability of DNA sequencing has made it possible to characterize
61 microbial communities more easily and comprehensively, the microbiome has become a major area of
62 interest in numerous organisms (Kyrpides, Elloe-Fadrosch & Ivanova, 2016). Just as the emerging field of
63 human microbiome study has revealed that symbiotic, non-pathogenic microbes can have major impacts
64 on human health (David et al., 2014), plants also host numerous symbiotic microbes. Similar to the
65 human microbiome, the plant microbiome contains great diversity and is comprised of multiple distinct
66 communities in various tissue systems such as the roots (rhizosphere), leaf surface (phyllosphere), and
67 leaf interior (endosphere) (Turner, James & Poole, 2013). All of these microbiomes can have an impact on
68 their host's physiology. For instance, bacterial root microbes have been shown to play a role in the growth
69 of various plant species (Gaiero et al., 2013). They can also play a variety of roles in host physiology,
70 depending on their location within the plant (Schlaeppli & Bulgarelli, 2015). These communities can
71 be quite dynamic, and vary with factors such as plant age (Cavaglieri, Orlando & Etcheverry, 2009).
72 While each of these sets of interactions is important, here we focus on the microbial ecology of fungal
73 microorganisms living in the leaf endosphere. The endosphere is an ideal system for biogeographic
74 studies, both because it is highly diverse, well replicated, and because it is a major interface between the
75 host plant and its environment (Meyer & Leveau, 2012).

76 Endophytic microbes are naturally found in the interior of leaves, whether introduced by natural
77 wounds or openings in the leaf surface, or through penetrating the plant surface with hydrolytic enzymes
78 (Hallmann et al., 1997). Although some of these microbes may be latent pathogens or decomposers
79 waiting for the leaf to die, others are mutualists that may confer a benefit to their host (Carroll, 1988).
80 Endophytes are commonly divided into classes, based on where they are found and what roles they
81 are known to play in host tissues (Rodriguez et al., 2009). In wild grasses, Class 1 endophytic fungi
82 (Clavicipitaceous endophytes) have been shown to protect their hosts by discouraging herbivory, and
83 can even affect host reproductive viability in those same systems (Clay, 1988). In controlled settings,
84 inoculation experiments with Class 3 endophytes (horizontally transmitted and localized to shoots) have
85 shown that these specific taxa of endophytes can also have an impact on their host's overall health,
86 including factors such as resistance and susceptibility to disease (Busby, Ridout & Newcombe, 2016). In
87 nature, Class 3 fungal endophytes can have impacts on their host's physiology, such as limiting pathogen
88 damage (Arnold et al., 2003). In this study, we expected to primarily find Class 3 endophytes, which are
89 known for having high diversity both within populations of plants and within plant tissues (Rodriguez et
90 al., 2009).

91 In the wild, endophytic communities display species diversity comparable to that of any macroscopic
92 community, even among individual trees from the same species (Gazis, Rehner & Chaverri, 2011).
93 However, determining the factors that influence this diversity across disparate environments remains an
94 area of active research. The biodiversity of these communities can be quite high, especially in areas like
95 tropical forests (Arnold & Lutzoni, 2007). In such natural settings, plant-associated microbial community
96 compositions can show clear biogeographic structure (Andrews & Harris, 2000). The urban setting,
97 however, may be distinct, because factors such as rainfall and elevation will likely be less apparent over a
98 smaller geographic area, even as new factors such as proximity to roads and tall buildings may introduce
99 effects of their own. Studies of suburban forests in Japan have indicated that an urban setting has a
100 notable impact on endophytic diversity (Matsumura & Fukuda, 2013). However, the full impact of urban
101 environmental factors on endophytic communities has yet to be completely understood.

102 In this study, we used culturing and barcode gene sequencing to identify the species makeup of

103 endophytic communities in the New Zealand Christmas Tree (Māori: pōhutukawa), *Metrosideros excelsa*
104 (Myrtaceae), throughout San Francisco, CA, to quantify the diversity and biogeographic patterns of foliar
105 fungal communities in this urban environment. Based on studies in natural systems, we expected to find
106 high community diversity both within and between sites, as well as a degree of biogeographic structure to
107 these community compositions. We also anticipated that urban environmental factors might play a role in
108 shaping the biogeography in these endophytic communities.

109 **METHODS**

110 **Host Selection**

111 We selected *Metrosideros excelsa* as a focal host tree. *Metrosideros excelsa* is widely planted throughout
112 San Francisco and as such we were able to obtain samples from a large number of trees in a variety of
113 locations throughout the city.

114 **Site Selection**

115 We selected six sites across the city for leaf sample collection (Fig. 1). When selecting sites, we
116 took factors such as traffic levels, estimated temperature, proximity to tall buildings (greater than the
117 40-foot building-height limit imposed on most residential buildings in San Francisco), elevation, and
118 proximity to the ocean and San Francisco Bay into account. In selecting the sites, we wanted to
119 capture the wide range of potential urban climates. We used street tree location data available from
120 the City of San Francisco (permalink: <https://data.sfgov.org/City-Infrastructure/Street-Tree-List/tkzw-k3nq>), which documents the location and species of every tree in San
122 Francisco, to choose unique locations around the city with sufficient densities of *Metrosideros excelsa*
123 individuals (Fig. 1).

124 **Sample Collection**

125 We collected small branches from 5 trees in each of these sites using a clipper pole, collecting at least 3
126 sun-exposed outer branches from each tree. Because *M. excelsa* is an evergreen tree and the newer leaves
127 are likely to contain fewer fungi, we controlled for leaf age by only collecting branches which contained
128 dark green leaves that appeared to be at least one year old. We collected all samples on the same day,
129 August 26, 2017, to ensure that daily weather patterns and seasonal effects would not have an impact on
130 the microbial community composition. Once collected, leaves were stored in labeled plastic bags and
131 stored at 4°C until culturing. All leaves were processed within 48 hours of collection. Permission to
132 sample leaves was granted by the San Francisco Bureau of Urban Forestry.

133 **Culturing Methods**

134 From each sample, we selected a subset of dark green asymptomatic leaves for culturing. We surface-
135 sterilized the leaves by first rinsing them with distilled water, then immersing them in a sterile petri
136 dish with 95% ethanol for 10 seconds, 0.5% NaOCl for 2 minutes, then 70% ethanol for another 2
137 minutes following Arnold et al. (2007). We emptied the dish between rinses and left it closed inside a
138 sterile biosafety cabinet after the last rinse. We then cut the leaves into small (2mm x 2mm) pieces with
139 flame-sterilized scissors and placed them into 1.5 mL microcentrifuge slant tubes partially filled with 2%
140 VWR-brand malt extract agar (MEA). For each tree, we prepared 6 leaves and made 100 tubes, except for
141 the trees from the downtown site. For the trees at this site, we prepared 140 tubes per tree because we
142 found that they had low isolation frequencies in a preliminary sampling bout. All leaves were prepared
143 this way within 48 hours of the initial leaf sampling, to prevent death of the leaf tissue from altering the
144 fungal community composition.

145 After two weeks, we evaluated the tubes for fungal growth and subcultured the emergent fungi from
146 tubes with growth onto 35mm plates with 2% MEA in order to better evaluate their morphotypes and
147 accumulate sufficient tissue for future barcode gene sequencing and water voucher preparation. We
148 re-evaluated and subcultured these tubes in the following months to capture any late-growing fungi. Each
149 pure culture was vouchered in sterile dI water and stored in a living culture collection in the Zimmerman
150 Lab at the University of San Francisco.

151 **Molecular Methods**

152 We extracted DNA from fungal mycelium using the Sigma RED Extract 'n Amp DNA extraction kit and
153 following previously published protocols (U'Ren et al., 2012). First, we added fungal tissue to sterile

154 tubes filled with 1 mm zirconium oxide beads, then added 100 μ L of Extract 'n' Amp DNA extraction
155 solution. Next, we put the tubes in a bead-beater (Mini-Beadbeater-96, Bio Spec Products Inc.) for one
156 minute. The samples were then placed on a heat block at 95°C for 10 minutes. After the heating step, we
157 added a dilution buffer to each tube and stored them at 4°C until PCR.

158 We performed PCR using fungal-specific primers for the Internal Transcribed Spacer region, a
159 commonly accepted fungal barcode locus (Schoch et al., 2012), using the ITS1F forward primer (5'-CTT
160 GGT CAT TTA GAG GAA GTA A-3') and ITS4 reverse primer (5'-TCC TCC GCT TAT TGA TAT
161 GC-3'). For each PCR reaction, we used 1 μ L of template DNA, 10 μ L Extract 'n' Amp Taq polymerase,
162 6.4 μ L PCR-grade water, 1 μ L bovine serum albumin, 0.8 μ L ITS1F forward primer, and 0.8 μ L ITS4
163 reverse primer (U'Ren et al., 2012). For the PCR reaction, we used a BioRAD T100 thermal cycler with
164 the following cycles: 95°C for 3 minutes; 35 cycles of 95°C for 30 seconds, 54°C for 30 seconds, 72°C
165 for 30 seconds; then 72°C for 10 minutes. To ensure that the fungal DNA successfully amplified, and that
166 the master mix was not contaminated, we ran 5 μ L of each sample and a negative PCR control on a 1%
167 agarose gel with 1X Tris-acetate-EDTA (TAE) buffer and SYBR Safe. We ran the gel at 120 volts for 20
168 minutes and visualized bands using UV transillumination. Successful PCR samples with clean negative
169 controls were kept at 4°C until sequencing preparation.

170 To prepare successfully amplified samples for Sanger sequencing, we first cleaned each sample with 1
171 μ L Thermo Fisher Shrimp Alkaline Phosphatase Exonuclease (ExoSAP-IT). To clean the samples, we
172 used the following cycle on a BioRAD T100 thermal cycler: 37°C for 15 minutes, 80°C for 15 minutes,
173 then hold at 4°C. After cleaning, samples were kept at 4°C until they were ready to be sent for sequencing.
174 Directly before being sent for sequencing, cleaned samples that showed bright bands on their gels were
175 diluted with an additional 15 μ L PCR water, although a small number of samples that had faint bands on
176 the gels were not diluted. Cleaned samples were sent to MCLabs (South San Francisco, CA) for Sanger
177 sequencing.

178 **Computational and Statistical Methods**

179 We used Geneious 11.1 to manually clean and trim the Sanger sequencing data, and to identify and remove
180 failed and low-quality sequences (Kearse et al., 2012). Sequences that appeared to have multiple strong
181 signals at a given position were discarded because they may have originated from from a mixed culture.
182 Usable sequences were cleaned by trimming the ends of low-quality nucleotides, clarifying any ambiguity
183 codes, and resolving incorrect base assignment due to dye blobs. Cleaned sequences are available on
184 GenBank (finalized genbank IDs to go here following acceptance).

185 After processing the cleaned sequences to add relevant metadata, we used mothur version 1.39.5 to
186 determine Operational Taxonomic Units (OTUs), based on 97% ITS sequence similarity and abundance-
187 based greedy clustering (Schloss et al., 2009). Next, we used R 4.1.1 (R Core Team, 2021) to analyze the
188 resulting OTU table. We used vegan (Oksanen et al., 2018) to: 1) calculate species accumulation curves
189 to determine species richness from the OTU data, 2) estimate unknown total species richness using the
190 Chao estimator with bias correction for small sample sizes (Chao, 1987; Chiu et al., 2014), 3) calculate
191 Bray-Curtis distances between communities, 4) create a Non-Metric Multidimensional Scaling (NMDS)
192 ordination, and 5) perform a PERMANOVA (Anderson, 2001) to assess differences in communities
193 between sites and between trees of different diameters at breast height (DBH) within sites. We used
194 the Tree-Based Alignment Selector (TBAS) toolkit to assign taxonomic context to each ITS sequence
195 (Carbone et al., 2017). This toolkit matches unknown ITS sequences to the most similar ITS sequences in
196 a large multi-gene phylogeny of confidently assigned taxa. We used the simple features (Pebesma, 2018)
197 and ggmap (Kahle & Wickham, 2013) R packages to produce maps of study sites.

198 Significant differences in fungal leaf isolation frequency and tree diameter were determined using
199 a nonparametric Kruskal-Wallis one-way test of variance and Dunn's Test for multiple comparisons
200 with Holm's correction for multiple hypothesis testing. All analyses, figures, and the final manuscript
201 were generated using a makefile, bash scripts, and R and Rmarkdown. All relevant metadata, as well as
202 analysis and manuscript code, are available on GitHub: [https://github.com/ZimmermanLab/
203 SF-metrosideros-endophytes/](https://github.com/ZimmermanLab/SF-metrosideros-endophytes/) and archived at (Zenodo doi here following acceptance). The
204 complete computational environment for the analyses is documented in a Dockerfile based on rocker
205 containers (Boettiger, 2014) and a renv.lock (Ushey, 2021) file in that same repository.

Methods for Comparing Fungi in San Francisco Trees to New Zealand Trees

In order to compare the foliar microbial populations found in *M. excelsa* throughout San Francisco to those in the tree's native range (New Zealand), we surveyed public GenBank records and performed a literature search. First, we searched Google Scholar on July 7th, 2021, for papers mentioning any of the following sets of search terms: "metrosideros endophytes," "metrosideros excelsa endophytes," "metrosideros excelsa foliage fungus," "metrosideros excelsa endophytes." After this search, we looked through the papers cited by those we originally found in order to find any additional studies.

Then, we searched GenBank for fungal ITS sequences that had *M. excelsa* listed as their host organism. We performed this search using the parameters "host="Metrosideros excelsa", organism=fungi", in order to limit our results to fungal sequences. This search yielded several new sequences that were not included in the previous search because they were unpublished. Using both of these methods, we were able to find about 30 taxa to include as a point of comparison, although many of these were described as pathogenic instead of endophytic.

RESULTS

Overall, we found high diversity in these urban endophytic communities. From 30 different host trees, we successfully cultured and sequenced 854 fungal isolates, which made up 85 distinct 97% ITS OTUs. Some communities encompassed a low number of fungal isolates; species richness analyses showed that in nearly all of the trees analyzed, our sampling did not encompass the complete microbial diversity (Fig. 2). While we found species diversity within some sites that exceeds 20 distinct fungal OTUs, there appears to be greater species diversity in these fungal endophytic communities than this study was able to document. Chao estimated richness overall was 166.6 with a standard error of 35 and is shown per site in Table 1.

Isolation Frequency and Host Tree Diameter at Breast Height

The isolation frequency, or the percentage of leaf pieces that yielded fungal isolates, varied significantly between sites (Kruskal Wallis $p < 0.01$, Fig. 3A). In most sites, the isolation frequency also varied between trees, especially in the Bay site. The diameter at breast height (DBH) of trees we sampled from differed significantly as well (Kruskal Wallis $p < 0.01$; Fig. 3B). Trees within the Bay site also show the greatest variation in this measure. However, there was not an overall significant linear relationship between tree DBH and measured isolation frequency (regression $p = 0.076$). The only site that shows consistently low variation in isolation frequency is the Downtown site, which is also the site with the lowest isolation frequencies overall.

Diversity Patterns

We found 85 total 97% nrITS OTUs among the 30 different trees. Both isolation frequency and the number of fungal species found varied notably between trees. Species accumulation curves showed that the total amount of fungal diversity was not completely sampled in any of the sites (Fig. 2).

The most abundant fungal class that we found in our samples was Dothideomycetes, followed by Sordariomycetes; each had over 300 sequences assigned to them (Table 2). Dothideomycetes was the predominant class in most sites except for the Downtown and Bay sites (Fig. 4). In both of these sites, Sordariomycetes was the most common class. There are several classes that are either absent or present in small numbers in most sites, but more abundant in one or several sites. For example, Eurotiomycetes are more common in the Downtown and Freeway sites, and Leotiomycetes are only abundant in the Mt. Davidson site. Only two sequences were not identified by T-BAS. We queried these two sequences against the NCBI nt database via BLASTn and found them to match with high confidence to the Pezizomycetes, based on similarity to a strain recently identified in Australia (Catcheside, Qaraghuli & Catcheside, 2017).

Biogeographic Patterns

While the NMDS ordination group ellipses (standard error around the centroid for each site) suggest that there might be differences in multivariate dispersion among sites, dispersions were not significantly different from one another ($p = 0.76$; Fig. 5), suggesting that PERMANOVA differences are likely to be caused by compositional differences rather than differences in dispersion. Both site and tree DBH were significant predictors of differences in community composition (Table 3), but their interaction was not significant ($p > 0.05$). Sites that were closer together geographically tended to also be closer to one

257 another in NMDS ordination space. We also tested for the effect of tree DBH on community composition
258 differences within sites, and found that it was significant ($R^2 = 0.11$, PERMANOVA $p = 0.006$).

259 DISCUSSION

260 In this study, we sought to characterize the foliar microbiome of *Metrosideros excelsa* trees across the city
261 of San Francisco. We found that the microbial composition of these urban trees' leaves varied in many
262 aspects, from the isolation frequencies of fungi from their leaves to the taxonomic identities of their fungal
263 isolates. Although *M. excelsa*'s endophytic communities have been characterized in its native range (New
264 Zealand), there have been few studies about these communities outside of its native environment or in an
265 urban setting (McKenzie, Buchanan & Johnston, 1999). In a Hawaiian endemic congener, *Metrosideros*
266 *polymorpha*, the species makeup of foliar fungal endophyte communities has been shown to vary greatly
267 across gradients of environmental factors such as elevation and rainfall (Zimmerman & Vitousek, 2012).
268 While we did not explicitly quantify environmental variation in this study, the community variation we
269 observed could be explained by either environmental or host physiological factors.

270 Isolation Frequency and Tree size

271 Both isolation frequency and DBH varied among sites. The trees in the Bay site show the greatest range
272 in both isolation frequency and DBH (Fig. 3). The Mt. Davidson trees have a larger median and range
273 of DBH than trees at the Balboa site (Fig. 3B), and also have a significantly higher median isolation
274 frequency than the Downtown site. Just as it shows the greatest range of isolation frequencies and tree
275 sites, the Bay site also has some of the least similarity between its fungal communities on the NMDS
276 ordination (Fig. 5). The trees in the Mt. Davidson site, which have some of the most similar communities
277 (Fig. 5), also have fairly large trees with similar DBH. Although there appears to be a general pattern
278 with larger trees hosting a greater number of fungal endophytes, this relationship was not significant; thus
279 DBH cannot explain all of the variation in isolation frequencies.

280 Diversity Patterns

281 Endophytic microbiomes in tree leaves are known to be highly diverse in natural systems, and that also
282 appears to be the case in these urban trees. The species accumulation curves indicate that even in the
283 most well-represented sites, there are still many potentially undiscovered OTUs within these endophytic
284 communities (Fig. 2). Additionally, there is a considerable amount of taxonomic variation between
285 certain sites, in addition to the generally high diversity of each site (Fig. 4). In such cases, it is likely that
286 environmental factors play a role in shaping the endophytic communities of these trees. While the impact
287 of environmental factors may be less evident when focusing on fungal endophyte community richness, it
288 becomes more apparent when looking at the identities of these endophytes. The composition of these
289 communities can vary greatly among trees with similar isolation frequencies, as demonstrated by the
290 taxonomic composition of the Mt. Davidson and Bay sites (Fig. 4). It is possible that prevailing onshore
291 weather dynamics, which generally lead the western part of the city to be colder and foggier than the
292 eastern portion, explain some of the patterns we observed. In general, there was higher compositional
293 similarity between trees from the same location than between trees of a similar size (Fig. 5 and Table 3).

294 Biogeographic Patterns

295 There appears to be a degree of biogeographic structure to the endophytic communities within these urban
296 trees. Communities of trees from the same site generally cluster closer together than trees from different
297 sites (Fig. 5 and Table 3). Although the Balboa site appears to have a reasonably high isolation frequency
298 (Fig. 3A), many of the slant tubes that showed fungal growth failed to grow into a larger culture, leading
299 to a lower number of usable sequences (Fig. 2B). This low number of sequences per tree may indicate
300 that the divergence in community composition that appears to be present might be due to the low sample
301 size, rather than actual difference between the microbial communities between the Balboa trees (Fig. 5).

302 However, even sites with low within-site diversity appear to be more similar to other trees from
303 geographically close sites. For example, the Downtown trees do not cluster with the Mt. Davidson,
304 Balboa, or Ocean trees in the ordination, indicating that they have fairly different fungal communities
305 from those sites (Fig. 5). Generally, the westernmost sites (Ocean and Balboa) cluster to one side of the
306 ordination, while the Easternmost sites (Bay and Downtown) cluster to the other, and the sites closer to
307 the middle of the city (Mt. Davidson and Freeway) cluster between them (Fig. 5). Notably, the Downtown

308 and Freeway sites cluster fairly close together (Fig. 5) despite being fairly distant geographically (Fig.
309 1). In this case, it is possible that there are environmental factors that cause the communities to be more
310 compositionally similar. One of the most likely environmental similarities between these two sites is that
311 they are both exposed to high traffic levels, with one site located in a bustling downtown and the other
312 located near the intersection of two large freeways (280 and 101). Carbon dioxide emissions, particulates,
313 or both of these pollutants could possibly be having an effect on both these trees and their microbiomes.
314 Endophytic bacteria have been shown to have a phytoremediating effect (Afzal, Khan & Sessitsch, 2014),
315 so it is possible that endophytic fungi may be playing a similar role in helping these plants persist in
316 high-traffic and thus lower air quality conditions.

317 **Community composition compared to related trees in other locations**

318 In widely distributed plant species, fungal endophyte composition can differ quite considerably in a
319 species from its native home to the places it has been introduced (Taylor, Hyde & Jones, 1999). *M.*
320 *excelsa* is native to New Zealand, but it and its close relatives in the *Myrtaceae* are widespread across the
321 world. Across all locations, the most abundant classes found within these *M. excelsa* fungal communities
322 were Dothideomycetes and Sordariomycetes by a notable margin (Table 2). At broad taxonomic levels
323 such as Class, the endophytes found in this study appear to be similar to those found in other *Myrtaceae*
324 trees elsewhere. For instance, a study of *Myrtaceae* trees in South America, the first and second most
325 common fungal isolates were also identified as Dothideomycetes followed by Sordariomycetes (Vaz et
326 al., 2014). Dothideomycetes and Sordariomycetes are also the first and second most abundant classes in
327 the endophytic communities of *M. excelsa*'s close relative in Hawai'i, *Metrosideros polymorpha*, as well
328 (Zimmerman & Vitousek, 2012). This could indicate that these fungal classes are well-adapted to living
329 as endophytes within these trees, but it is also possible that these communities are actually different when
330 considered at a finer taxonomic resolution.

331 In order to determine if the fungi we identified were simply members of ubiquitous taxa or if they
332 were potentially more specific to *M. excelsa*, we compared our results to a study of endophytes in
333 two species of trees in the nearby Northern California county of Mendocino. This study looked at
334 two species of host plants: an angiosperm, *Vaccinium ovatum*, and a gymnosperm, *Pinus muricata*.
335 Overall, urban *M. excelsa* shared few of its prominent endophytic classes with *P. muricata*, but it shared
336 Dothideomycetes as a common class with *V. ovatum* (Oono et al., 2020). The most common class in *P.*
337 *muricata* was Atractiellomycetes (Oono et al., 2020), which did not appear prominently in any of the
338 trees we studied. Based on this admittedly limited comparison, it appears that angiosperm species may
339 have more endophytic similarities to each other than to gymnosperm species in the same geographic
340 area. However, because we found that the second most common fungal class in our *M. excelsa* leaves,
341 Sordariomycetes, was not present in large quantities in any of the trees in the study by Oono and
342 colleagues, finer-scale host species filtering is also likely to be playing a role in shaping the communities
343 that establish in *M. excelsa*. This conclusion has been suggested in other more comprehensive studies of
344 fungal endophyte host plants as well (U'Ren et al., 2019; Darcy et al., 2020).

345 **Comparison of trees in San Francisco and New Zealand**

346 When comparing *M. excelsa*'s endophytic taxa found in the literature to those that we collected, we found
347 that the taxa between the regions were somewhat similar on an order level, and even more so on a class
348 level. Only one of the orders that was extremely common in New Zealand (Mycosphaerellales) was
349 absent in San Francisco, and only two of the most common orders in San Francisco (Capnodiales and
350 Pezizales) were absent in New Zealand. However, most of the families that are found in New Zealand are
351 absent in San Francisco, and vice versa. We only found four families in common: Mycosphaerellaceae,
352 Glomerellaceae, Nectriaceae, and Helotiaceae. Of these, Mycosphaerellaceae is the only family that is
353 present in New Zealand and was prevalent (over 50 isolates) in San Francisco. The other families that
354 they had in common were only present in very small numbers in San Francisco, and we were unable to
355 find documentation of three most prevalent families in San Francisco (Xylariaceae, Botryosphaeriaceae,
356 and Cladosporiaceae) being isolated from leaves in New Zealand. Although this could be due to the small
357 sample size, it suggests that *M. excelsa*'s endophytic populations may be influenced primarily by their
358 environment. However, additional research into both the endophytic populations of other trees in San
359 Francisco and *M. excelsa* in its native range would be required to draw further conclusions.

360 It is important to note the difference in methods between how we collected our samples and how
361 others studying *M. excelsa* in New Zealand collected theirs. Whereas we took care to only collect samples

362 from unfallen, asymptomatic leaves, most of the papers referenced in this comparison use either leaf litter,
363 dead leaves, or obviously symptomatic leaves. Therefore, it is reasonable to infer that the data we are
364 comparing our results to is biased towards pathogenic and decomposing fungi, and underrepresents the
365 non-harmful endophytic populations in *M. excelsa*'s native range. Additionally, the number of fungal
366 isolates we were able to find using literature and GenBank searches was much lower than the number of
367 samples we collected ourselves, making quantitative comparison unfeasible. However, we think that such
368 comparisons are still worthwhile as a preliminary look into the extent that *M. excelsa* in San Francisco
369 may maintain its native endophytic microbiome.

370 CONCLUSIONS

371 These findings indicate that the endophytic microbiomes of urban trees are complex and diverse, and
372 may show a degree of biogeographic structure that reflects their natural counterparts. Additionally, it is
373 likely that urban environmental factors play a considerable role in shaping the endophytic communities of
374 these trees. This study has demonstrated that the urban endophytic microbiome is highly diverse and has
375 the potential to be structured by urban environmental factors. Our species accumulation curves indicate
376 that the full diversity of these endophytic communities has yet to be sampled (Fig. 2). Even in the small
377 geographic area of San Francisco, we found notable trends in microbiome composition that may vary with
378 geographic factors, host physical factors such as DBH and age, and uniquely urban environmental factors,
379 such as traffic. A combination of environmental factors and perhaps host physiology therefore appear to
380 be the driving force behind the diversity of these microbiomes. While it is difficult to determine, given the
381 data we have, the exact mechanisms that influence the composition of these communities, the amount
382 of species diversity and biogeographic structure we present here indicate that the foliar microbiomes of
383 urban trees may be just as complex and dynamic as those of trees in nature, and are worth of further
384 concerted study.

385 ACKNOWLEDGEMENTS

386 This study was conducted as an undergraduate Biology honors thesis project by EG under the guidance
387 of NZ. Joshua Copeland and Julian Murdzek assisted with culture processing, vouchering, and DNA
388 extractions. Jeff Oda contributed supplies and equipment expertise.

389 REFERENCES

- 390 Afzal M, Khan QM, Sessitsch A. 2014. Endophytic bacteria: Prospects and applications for the phytore-
391 mediation of organic pollutants. *Chemosphere* 117:232–242.
- 392 Alberti M, Marzluff JM, Shulenberg E, Bradley G, Ryan C, Zumbrunnen C. 2003. Integrating humans
393 into ecology: Opportunities and challenges for studying urban ecosystems. *AIBS Bulletin* 53:1169–
394 1179.
- 395 Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology*
396 26:32–46.
- 397 Andrews JH, Harris RF. 2000. The ecology and biogeography of microorganisms on plant surfaces.
398 *Annual review of phytopathology* 38:145–180.
- 399 Arnold AE, Henk DA, Eells RL, Lutzoni F, Vilgalys R. 2007. Diversity and phylogenetic affinities of
400 foliar fungal endophytes in loblolly pine inferred by culturing and environmental PCR. *Mycologia*
401 99:185–206.
- 402 Arnold AE, Lutzoni F. 2007. Diversity and host range of foliar fungal endophytes: Are tropical leaves
403 biodiversity hotspots? *Ecology* 88:541–549.
- 404 Arnold AE, Mejia LC, Kylo D, Rojas EI, Maynard Z, Robbins N, Herre EA. 2003. Fungal endophytes
405 limit pathogen damage in a tropical tree. *Proceedings of the National Academy of Sciences* 100:15649–
406 15654.
- 407 Boettiger C. 2014. An introduction to docker for reproducible research, with examples from the r
408 environment. *arXiv(2015) ACM SIGOPS Operating Systems Review, Special Issue on Repeatability*
409 *and Sharing of Experimental Artifacts*. 49(1), 71-79:1410.0846v1.
- 410 Busby PE, Ridout M, Newcombe G. 2016. Fungal endophytes: Modifiers of plant disease. *Plant*
411 *Molecular Biology* 90:645–655. DOI: 10.1007/s11103-015-0412-0.

- Carbone I, White JB, Miadlikowska J, Arnold AE, Miller MA, Kauff F, U'Ren JM, May G, Lutzoni F. 2017. T-BAS: Tree-based alignment selector toolkit for phylogenetic-based placement, alignment downloads and metadata visualization: An example with the pezizomycotina tree of life. *Bioinformatics* 33:1160–1168. DOI: 10.1093/bioinformatics/btw808.
- Carroll G. 1988. Fungal endophytes in stems and leaves: From latent pathogen to mutualistic symbiont. *Ecology* 69:2–9.
- Catcheside PS, Qaraghuli S, Catcheside DE. 2017. A new species of small black disc fungi, smardaea australis (pezizales, pyronemataceae), is described from australia. *Swainsona* 31:17–101.
- Cavaglieri L, Orlando J, Etcheverry M. 2009. Rhizosphere microbial community structure at different maize plant growth stages and root locations. *Microbiological research* 164:391–399.
- Chao A. 1987. Estimating the population-size for capture recapture data with unequal catchability. *Biometrics* 43:783–791. DOI: 10.2307/2531532.
- Chiu C, Wang Y, Walther B, Chao A. 2014. An improved nonparametric lower bound of species richness via a modified good-turing frequency formula. *Biometrics* 70:671–682. DOI: 10.1111/biom.12200.
- Clay K. 1988. Fungal endophytes of grasses: A defensive mutualism between plants and fungi. *Ecology* 69:10–16.
- Darcy J, Swift S, Cobian G, Zahn G, Perry B, Amend A. 2020. Fungal communities living within leaves of native hawaiian dicots are structured by landscape-scale variables as well as by host plants. *Mol Ecol* 29:3103–3116. DOI: 10.1111/mec.15544.
- David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA, others. 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505:559.
- Gaiero JR, McCall CA, Thompson KA, Day NJ, Best AS, Dunfield KE. 2013. Inside the root microbiome: Bacterial root endophytes and plant growth promotion. *American journal of botany* 100:1738–1750.
- Gazis RO, Rehner S, Chaverri P. 2011. Species delimitation in fungal endophyte diversity studies and its implications in ecological and biogeographic inferences. *Molecular Ecology* 20:3001–3013. DOI: 10.1111/j.1365-294X.2011.05110.x.
- Hallmann J, Quadt-Hallmann A, Mahaffee W, Kloepper J. 1997. Bacterial endophytes in agricultural crops. *Canadian Journal of Microbiology* 43:895–914.
- Kahle D, Wickham H. 2013. Ggmap: Spatial visualization with ggplot2. *The R Journal* 5:144–161.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. DOI: 10.1093/bioinformatics/bts199.
- Kong F, Yin H, James P, Hutrya L, He H. 2014. Effects of spatial pattern of greenspace on urban cooling in a large metropolitan area of eastern china. *Landscape and Urban Planning* 128:35–47. DOI: 10.1016/j.landurbplan.2014.04.018.
- Kyrpides NC, Elloe-Fadrosh EA, Ivanova NN. 2016. Microbiome data science: Understanding our microbial planet. *Trends in microbiology* 24:425–427.
- Matsumura E, Fukuda K. 2013. A comparison of fungal endophytic community diversity in tree leaves of rural and urban temperate forests of kanto district, eastern japan. *Fungal Biology* 117:191–201. DOI: <https://doi.org/10.1016/j.funbio.2013.01.007>.
- McDonnell MJ, Niemelä J. 2011. The history of urban ecology. *Urban Ecology*:9.
- McKenzie EHC, Buchanan PK, Johnston PR. 1999. Fungi on pohutukawa and other metrosideros species in new zealand. *New Zealand Journal of Botany* 37:335–354. DOI: 10.1080/0028825X.1999.9512637.
- Meyer KM, Leveau JH. 2012. Microbiology of the phyllosphere: A playground for testing ecological concepts. *Oecologia* 168:621–629.
- Nowak DJ, Hirabayashi S, Bodine A, Greenfield E. 2014. Tree and forest effects on air quality and human health in the united states. *Environmental Pollution* 193:119–129. DOI: <https://doi.org/10.1016/j.envpol.2014.05.028>.
- Oke TR. 1973. City size and the urban heat island. *Atmospheric Environment (1967)* 7:769–779.
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2018. *Vegan: Community ecology package*.
- Oono R, Black D, Slessarev E, Sickler B, Strom A, Apigo A. 2020. Species diversity of fungal endophytes across a stress gradient for plants. *New Phytologist* 228:210–225. DOI: <https://doi.org/10.1111/nph.16111>.

1111/nph.16709.

Pebesma E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* 10:439–446. DOI: 10.32614/RJ-2018-009.

R Core Team. 2021. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rodriguez RJ, White Jr JF, Arnold AE, Redman RS. 2009. Fungal endophytes: Diversity and functional roles. *New Phytologist* 182:314–330. DOI: <https://doi.org/10.1111/j.1469-8137.2009.02773.x>.

Schlaeppi K, Bulgarelli D. 2015. The plant microbiome at work. *Molecular Plant-Microbe Interactions* 28:212–217.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, others. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology* 75:7537–7541.

Schneider A, Friedl MA, Potere D. 2009. A new map of global urban extent from MODIS satellite data. *Environmental Research Letters* 4:044003.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Bolchacova E, Voigt K, Crous PW, others. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for fungi. *Proceedings of the National Academy of Sciences* 109:6241–6246.

Sukopp H. 1998. Urban ecology—scientific and practical aspects. In: *Urban ecology*. Springer, 3–16.

Taylor JE, Hyde KD, Jones EBG. 1999. Endophytic fungi associated with the temperate palm, trachycarpus fortunei, within and outside its natural geographic range. *New Phytologist* 142:335–346.

Turner TR, James EK, Poole PS. 2013. The plant microbiome. *Genome biology* 14:209.

U'Ren JM, Lutzoni F, Miadlikowska J, Laetsch AD, Arnold AE. 2012. Host and geographic structure of endophytic and endolichenic fungi at a continental scale. *American Journal of Botany* 99:898–914.

U'Ren J, Lutzoni F, Miadlikowska J, Zimmerman N, Carbone I, May G, Arnold A. 2019. Host availability drives distributions of fungal endophytes in the imperilled boreal realm. *Nat Ecol Evol* 3:1430–1437. DOI: 10.1038/s41559-019-0975-2.

Ushey K. 2021. *Renv: Project environments*.

Vaz AB, Fontenla S, Rocha FS, Brandao LR, Vieira ML, De Garcia V, Goes-Neto A, Rosa CA. 2014. Fungal endophyte β -diversity associated with myrtaceae species in an andean patagonian forest (argentina) and an atlantic forest (brazil). *Fungal Ecology* 8:28–36.

Willis KJ, Petrokofsky G. 2017. The natural capital of city trees. *Science* 356:374–376. DOI: 10.1126/science.aam9724.

Wu J. 2014. Urban ecology and sustainability: The state-of-the-science and future directions. *Landscape and Urban Planning* 125:209–221.

Zimmerman NB, Vitousek PM. 2012. Fungal endophyte communities reflect environmental structuring across a hawaiian landscape. *Proceedings of the National Academy of Sciences* 109:13022–13027. DOI: 10.1073/pnas.1209872109.

Table 1. Estimated richness of OTU counts across sites using the Chao estimator suggest that OTU richness is higher than observed and is characterized by large uncertainty at several sites.

Site	OTU Count	Chao Estimated Richness	Chao Richness SE
Balboa	15	19.9	5.0
Bay	35	105.5	50.9
Downtown	27	44.1	11.8
Freeway	29	144.6	128.9
Mt. Davidson	37	75.7	24.7
Ocean	24	184.0	175.7

Table 2. Overall sequenced fungal isolate counts by Class and Order. All fungi successfully sequenced in this study were placed within the Ascomycota.

Class	Order	Number of sequences
Dothideomycetes	Capnodiales	266
Dothideomycetes	Botryosphaerales	127
Dothideomycetes	Pleosporales	52
Dothideomycetes	unclassified,Capnodiales	3
Dothideomycetes	Trypetheliales	1
Dothideomycetes	unclassified	1
Eurotiomycetes	Eurotiales	57
Leotiomycetes	Helotiales	19
Pezizomycetes	Pezizales	54
Sordariomycetes	Xylariales	237
Sordariomycetes	Diaporthales	49
Sordariomycetes	Coniochaetales	38
Sordariomycetes	Sordariales	9
Sordariomycetes	Hypocreales	7
Sordariomycetes	Glomerellales	1

Table 3. PERMANOVA model output based on Bray-Curtis distance. Foliar fungal community composition in leaves of *Metrosideros excelsa* across San Francisco is significantly related to both site of sampling and tree diameter at breast height (DBH), but the interaction between these parameters is not significant.

Term	df	Sum of Squares	Mean of Squares	F	R^2	p-value
Site ID	5	3.642	0.728	3.015	0.373	0.001
DBH (cm)	1	0.556	0.556	2.302	0.057	0.007
Site x DBH	5	1.208	0.242	1.000	0.124	0.484
Residuals	18	4.349	0.242	NA	0.446	NA
Total	29	9.755	NA	NA	1.000	NA

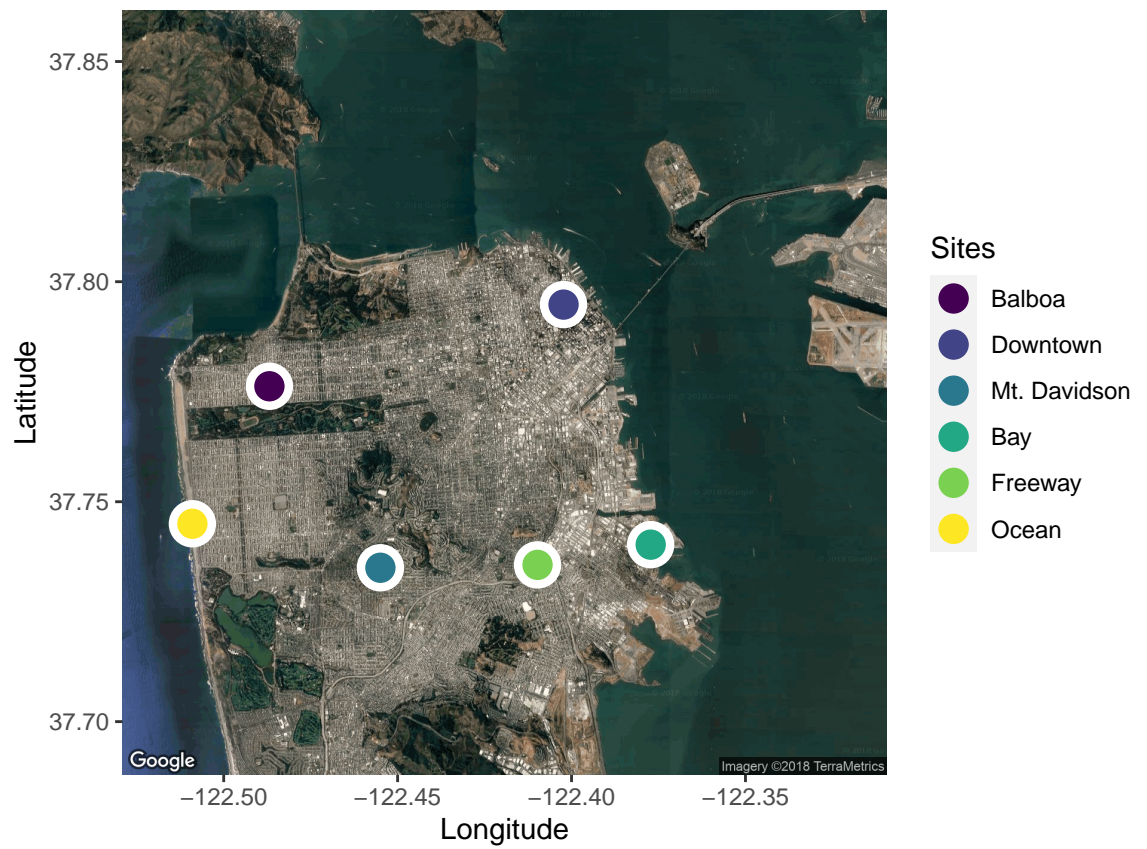


Figure 1. A map of the locations sampled across the city of San Francisco, California, USA. Five trees from each site were sampled. Sites were selected to span the range of environmental conditions across the city. Map data: Google, ©2018 TerraMetrics.

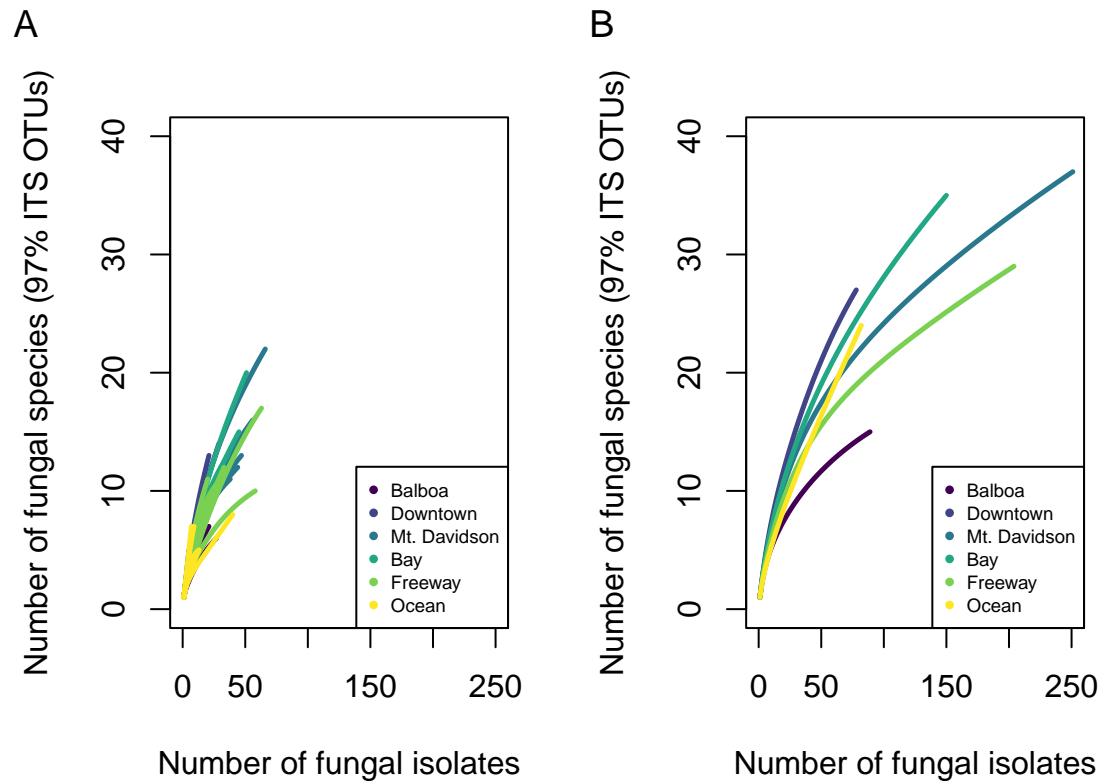


Figure 2. Species accumulation curves showing fungal community OTU richness. Each line in A represents the OTU richness for a single host tree. Each line in B represents the combined OTU richness of all trees (n = 5) in one site.

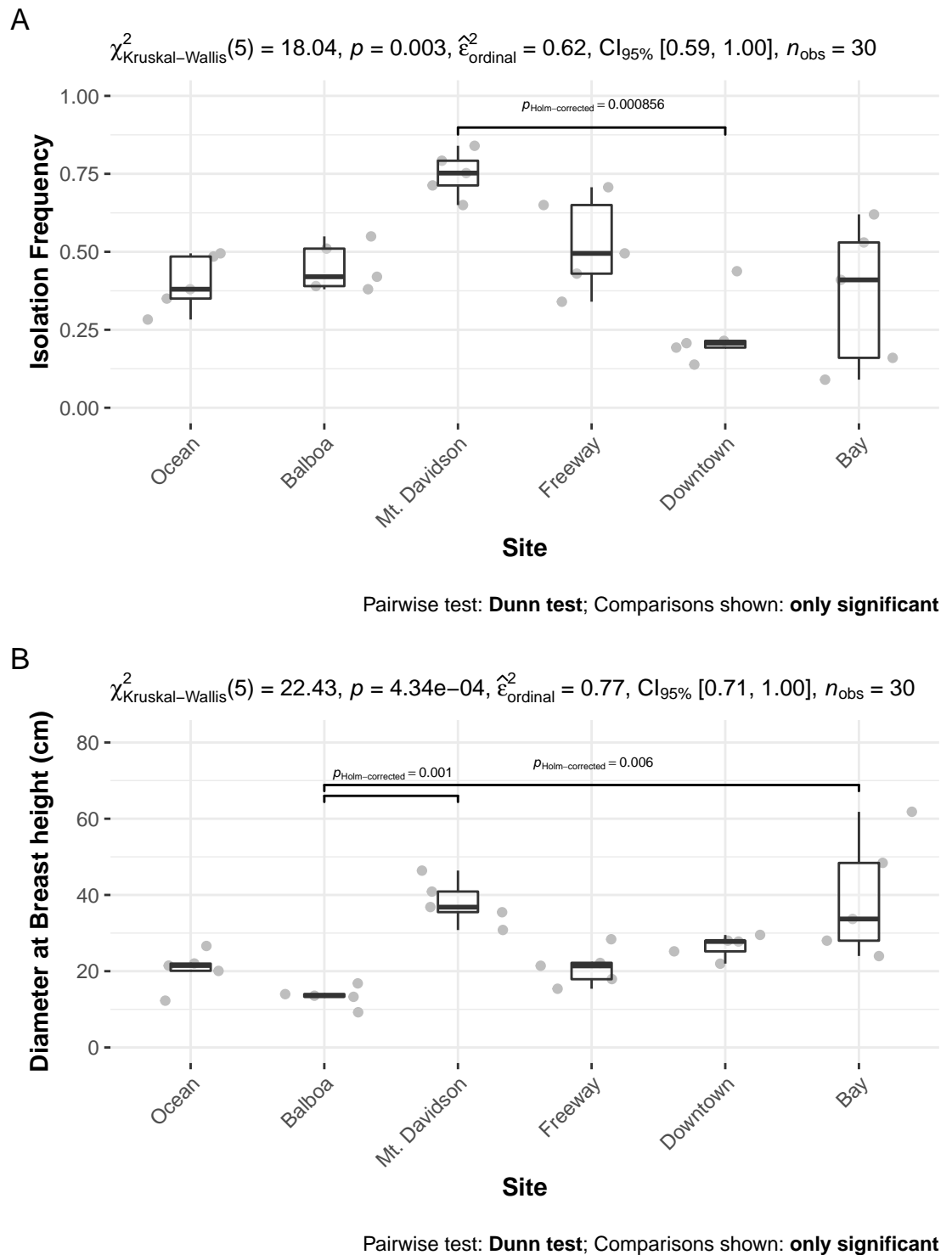


Figure 3. Isolation frequencies (A) and tree diameters (B) at each site. Isolation frequency (A) is a measure of how many slant tubes showed signs of fungal growth, out of how many total slant tubes were made. Approximately 100 slant tubes were made for each tree except for the trees in the downtown site, which had 140 slant tubes per tree because they had low isolation frequencies during the initial sampling. Sites along x axis are arranged left to right in order of their geographic location from west (left) to east (right).

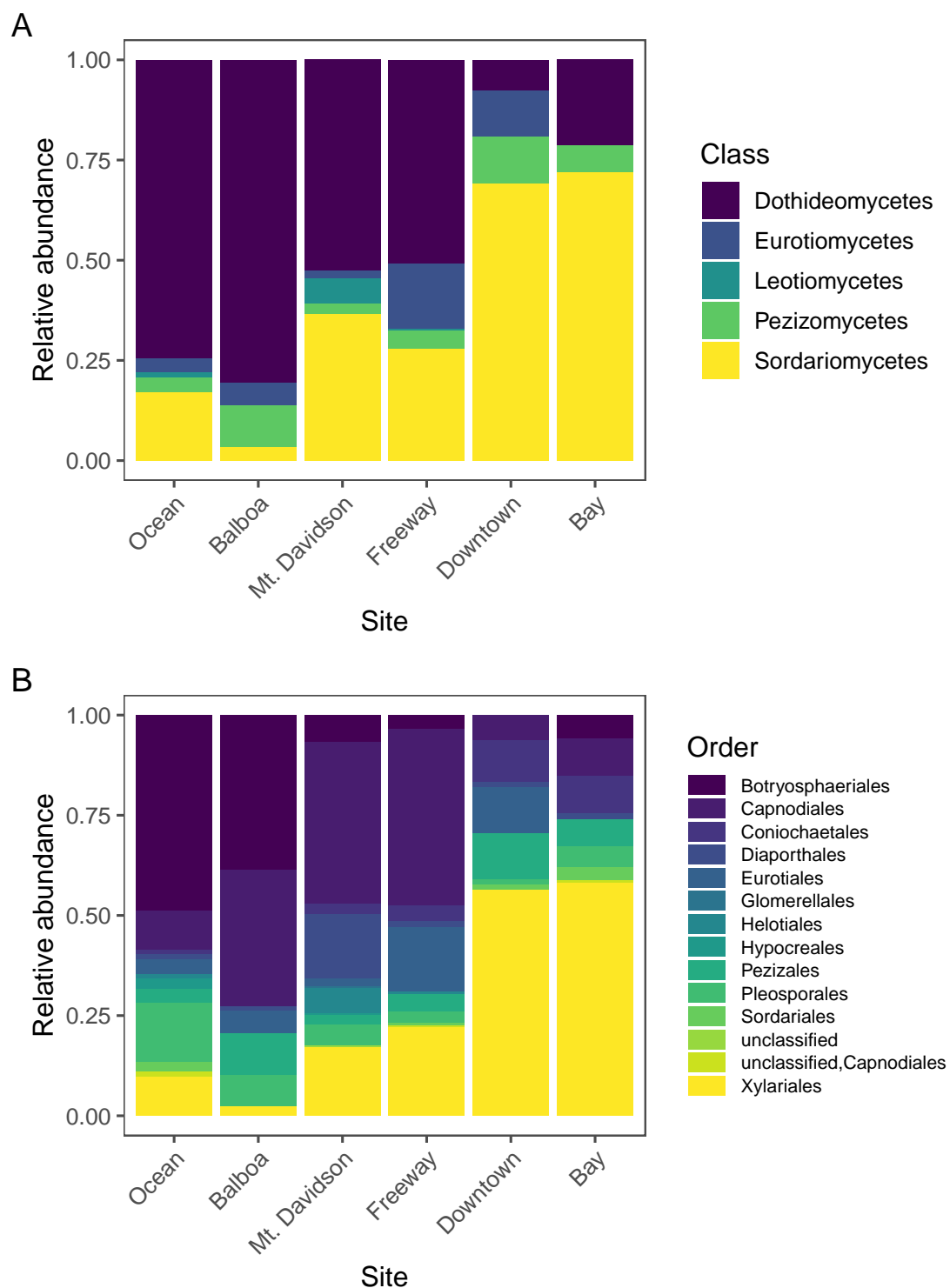


Figure 4. Normalized relative abundances of taxonomic groupings (Classes in A and Orders in B) in each site. Sites along x axis are arranged left to right in order of their geographic location from west (left) to east (right).

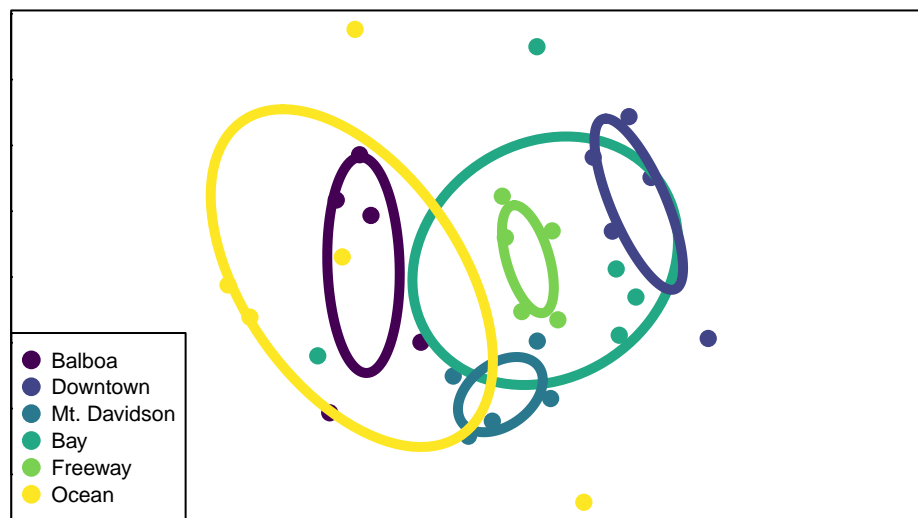


Figure 5. NMDS ordination of community compositions. Each point represents the endophytic community of one tree, and the size of the point corresponds to that tree's DBH, while the color of said point corresponds to the site that tree is from. Points that are closer together indicate that the trees they represent have similar community compositions. The ellipses show the standard error around the centroid of all points within a site, and are also color-coded according to which site they represent.