Applicant: Dan Thomas
Proposed advisors: Melissa McCormick (SERC) and Edward Allen Herre (STRI)

Delivering the goods: an environmental survey of secondary metabolites and the microbiome.

Introduction

The intensity and types of environmental challenges facing plants vary at all spatial scales, from landscape-level scales to individual plant bodies. Temporal changes in environmental challenges to plants also range from hourly, to decadal fluctuations in climate, pathogen loads, nutrient availability, and exposure. Plant genomes contain many solutions to these challenges, especially to environmental challenges which have acted as long-term selective pressures on plants. In these times of rapid change, however, the abundance and rhythm of all types of plant stressors has increased, and plant genomes alone may not evolve quickly enough to ensure the success of ecologically or socially important plants in the face of increased pathogen loads, decreased climatic stability, and decreased habitat availability.

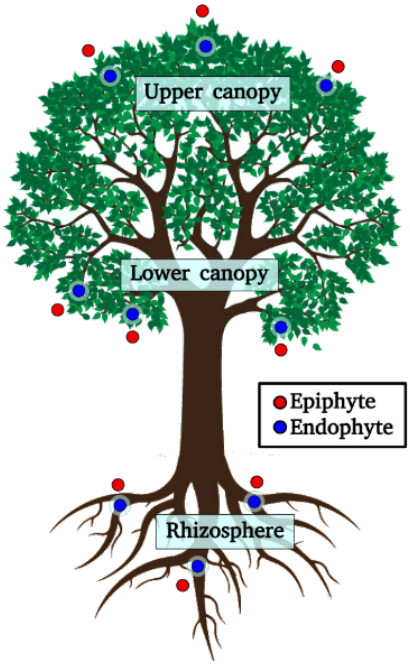
Some have proposed that the plant microbiome, with its collective abundance of microbial genetic resources, may represent an important and adaptive addition to the host plant genome. Together, it is proposed, the microbiome and the plant sum to make up the plant "hologenome", which is more extensive and more dynamic than the plant genome alone[1,2]. The correctness and utility of the hologenome model of the microbiome is actively debated[3–5], but evidence has mounted rapidly that microbiomes in general are important determinants in plant distributions and ecosystem health[6–10]. If microbiomes are to harnessed in any way to help plant health in times of rapid change, their spatial and temporal patterns must be better understood. To this end, biosynthetic gene clusters (BCG) are particularly important to study, as their self-contained nature allows simplified detection of entire pathways, and relatively easier transformation into orthologous systems: in general, BCGs represent a low-hanging fruit in the search for helpful natural metabolic solutions for enhancement of plants and drug synthesis[11–17]. Additionally, characterization of the functional microbiomes of trees is an important

next step in understanding the Janzen-Connell hypothesis of conspecific negative density-dependence in forest trees (see conclusion).

Here we propose to undertake sampling for the functional microbiome of a host tree species, *Eugenia nesiotica*, across the extent of the BCI ForestGEO plot. Endophytes from both leaf and root tissue of *Eugenia* will be surveyed for the presence of biosynthetic gene clusters, through metabarcoding methods. We will survey two of the most important and variable families of microbial natural product pathways, using barcodes targeted to the defining enzymes of these pathways, polyketide synthase (PKS) and nonribosomal ribosomal peptide synthetase (NRPS)[15,18,19].

Hypothesis: We hypothesize that plants recruit microbial symbionts that supplement or compliment plant genome function to enhance tolerance to stressors.

Prediction: Stressors faced by the host plant will predict the relative abundance of "useful" secondary metabolite pathways present in the microbiome. For example, increased pathogen load from nearby congeners of the host plant (see conclusion) will lead to increased incidence of anti-mycotic biosynthetic gene clusters present in the microbiome.



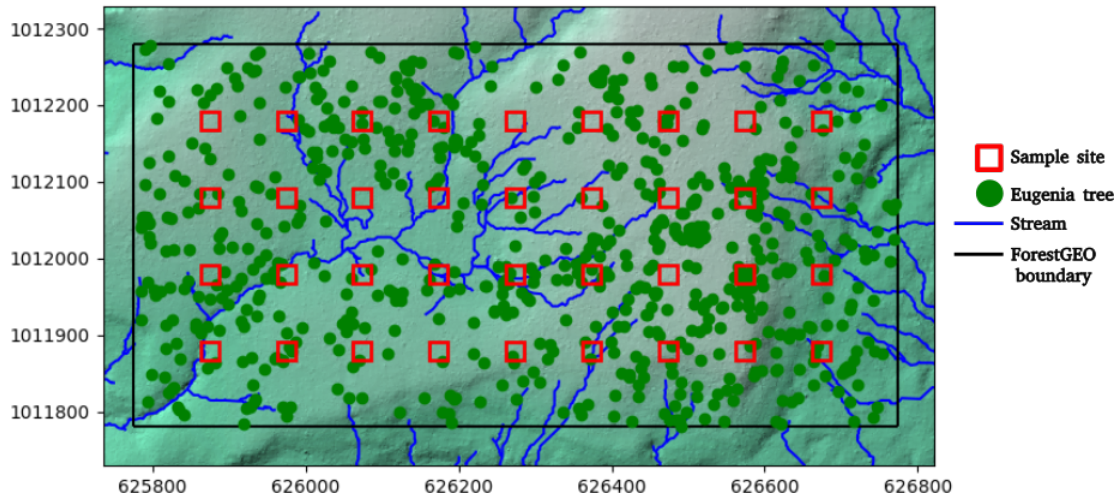
Methods

Host tree

Subject tree species *Eugenia nesiotica* was selected for its medium population density on the BCI forest dynamics plot (see map), for its history of research at BCI[20,21], and for its close relation to economically important species *E. uniflora*, whose transcriptome and genome are subject to ongoing investigation[22,23].

Sampling scheme

Spatial analysis of the presence of “useful” metabolites produced by plant-associated microbes will be coupled with micro-site data, estimates of disease and herbivory load, and the proximity of conspecific trees. Leaf and root tissue will be collected at 36 locations within the plot according to a spatial grid with a grain of 100m. Epiphyte and endophyte microbial populations will be sampled. Aerial plant tissues will be sampled at upper and lower canopy sites, to test micro-site variability in microbiome expression due to position in the canopy. Three paired samples will be collected from each of three zones of an individual tree, then pooled by zone and by community (endophyte or epiphyte) totaling to 6 genomic DNA samples per tree.



Sample preparation

Roots will be gently brushed clean of soil. To sample epiphytes of leaves and roots, leaves and root sections will submerged in 20 ml of 0.1% tween / sterile water mixture, and vortex-ed for 5 minutes. This wash fluid will be centrifuged, and the pellet frozen for later DNA extraction. Root and leaf endophytes sampling will follow Thomas et al.[24], using a bleach-based surface sterilization method, followed by tissue homogenization and DNA extraction. Bulk DNA will be purified from host plant tissue and epiphyte wash-pellets using column kits, and selectively PCR-amplified with above PKS and NRPS primers (with added sample barcodes and Illumina platform adapters). Sequences will be

sequenced using a Illumina MiSeq platform sequencer, with a MiSeq V3 kit (2 x 300 bp). Plant pathogen loads and damage from herbivory will be phenotyped from photos[25] of 10 randomly selected from leaves from upper and lower zones of the tree. Soil from each tree will also be harvested for soil moisture estimates.

PKS and NRPS metabarcoding

This metabarcode study will broadly target polyketide synthases (PKS) and nonribosomal peptide synthases (NRPS) with degenerate primers. These data will paint a general picture of some broad classes of secondary metabolites that tree microbiomes are predicted to be producing differentially in response to environmental challenges. These enzymes encode backbone molecules for two of the broadest families of microbial secondary metabolites (polyketides and nonribosomal peptides)[18,26], and exist in diverse homologs that can be used to infer medium-to-fine grain portrait of the secondary metabolome of the plant microbiome. Amplicon library preparation will follow Thomas et al.[24], for sequencing on Illumina MiSeq platform. Primer pairs for this survey will be based on two studies conducted by Charlop-Powers et al,[15,19] with primers that have a long history of use for environmental sampling of secondary metabolome from microbiomes. In soil, these primers tended to produce amplicon sizes of 200-350 bp, suitable for MiSeq platform sequencing. These primers on public databases show very few matches to plant sequences, with the vast majority of observations of NRPS and type I PKSes in public databases originate from microbial sequences. However, the possibility exists that that plants are greatly under-sampled for type I PKS gene clusters[27]. Since our hypothesis focuses on the complementary role of the microbiome, we will attempt enhance primers with plant-blocking oligonucleotides[28]. The sampling depth necessary to capture the full diversity of secondary metabolites is unknown, but estimated numbers of unique PKS and NRPS types in soil samples in both studies by Charlop-Powers et al. range from 2000 to less-than-10,000 enzyme “OTUs”. This is well within the sampling depth proposed here (~$23x10^6$ reads per MiSeq run /216 samples ~ 100,000 reads),

even given the extreme skewness of OTU abundance distributions from MiSeq runs. If time and budgetary constraints allow, our metabarcode survey will be augmented with more targeted PCR surveys for the presence of biosynthetic pathways of select microbial products. We will emphasize the detection of key metabolic steps in the microbial (non-plant) biosynthesis of anti-herbivory and pathogen response: ergot alkaloids[29–31] (a terpene but with NRPS-catalyzed pathway steps[30]) and other alkaloids, fungal anti-mycotic compounds such as echinocandins (NRPS)[26,32], polyene anti-mycotic including Nystatin and amphotericin (PKS) [33–35], and strobilurins (PKS)[36,37].

Bioinformatic and Statistical Analyses

Metabarcoding results will be subject to standard bioinformatic pipelines to reduce sequencing artifacts and group highly similar reads[38]. Following this, amplicon reads will be evaluated and grouped using closest-known biosynthetic gene clusters using the software package eSNaPD[15,39] and public databases. The resulting gene-cluster-by-sample-site matrix will be examined standard numerical ecology-based methods for modeling large matrices as dependent objects of environmental data ("regression" of metabarcoding results against environmental data and spatial patterns using redundancy analyses, including MEM and variation-partitioning methods)[40–43]. Read libraries will also be mined for specific biosynthetic pathways of interest (alkaloids, antibiotics, antioxidants, etc), or unknown gene clusters that may appear be acting to control disease, to the level of specificity allowed by the homology of the polyketide and peptide backbones. and these will be treated as point patterns[44], along with the results of any targeted PCR tests that may also be performed. Point patterns of secondary products of interest will analyzed for co-occurrence with environmental challenges, herbivory, pathogen loads, and proximity and size of conspecifics and cluster size of nearest conspecifics.

Conclusion: pertinence to research on-going at Smithsonian Institute (SI) and Smithsonian Tropical Research Institute (STRI):

One of the greatest contributions to tropical ecology in recent memory by SI and STRI is the empirical testing of the Janzen-Connell hypothesis of conspecific negative density dependence (CNDD), and exploration of the mechanisms behind CNDD[45–48]. Dr. Allen Herre and colleagues at STRI found that soil-born enemies may be main culprits in the negative density dependence patterns of rare trees in the tropics, but noted that unlike rare trees, highly abundant trees do not show patterns of CNDD[20]. Insufficient gene flow and genetic variation in the genomes of rare plants trees has been proposed as one mechanism for CNDD[21].

The project here proposed suggests another, complementary mechanism - if functional microbiomes are key to pathogen resistance, **then recruitment of beneficial microorganisms into the microbiome may function as a positive-density-dependent feedback.** Dr. Herre and colleagues have observed that many important "horizontally-acquired" microbes can be transmitted to seedlings through litter of older trees[49]. Abundant trees may therefore be able to transmit their microbiomes to seedlings more successfully than rare trees simply as a function of volume of leaf litter, a positive-density-dependent feedback in seedling success. Clusters of abundant species may represent larger "target" areas not only for pathogens but for environmental recruitment of beneficial microbial partners. Dr. Melissa McCormick at STRI and her colleagues found hypersensitivity by host plants to availability of symbiotic soil fungal partners on the landscape[50]. This indicates that the availability of mutualistic symbioses may be important to overcoming the inhibitory effects of soil-born enemies and the negative density: both plant and fungal mycorrhizal partners presumably benefit by the presence of the other, a positive density feedback. However, McCormick and colleagues have found that taxonomic diversity patterns in fungal microbiome don't necessarily co-vary in predictable ways with host post plant diversity[51]. This and other studies[2] indicate that classical taxonomic approaches to microbiome study need to be augmented by *functional* surveys of the microbiome such as here proposed.

1.....Rosenberg, E. & Zilber-Rosenberg, I. Microbes drive evolution of animals and plants: the hologenome concept. *mBio* **7**, e01395-15 (2016).
2.. .Lemanceau, P., Blouin, M., Muller, D. & Moënne-Loccoz, Y. Let the core microbiota be functional. *Trends in plant science* **22**, 583–595 (2017).
3.....Douglas, A. E. & Werren, J. H. Holes in the Hologenome: Why Host-Microbe Symbioses Are Not Holobionts. *mBio* **7**, (2016).
4.....Moran, N. A. & Sloan, D. B. The Hologenome Concept: Helpful or Hollow? *PLOS Biology* **13**, e1002311 (2015).
5.....Theis, K. R. *et al.* Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes. *mSystems* **1**, (2016).
6..…Laforest-Lapointe, I., Paquette, A., Messier, C. & Kembel, S. W. Leaf bacterial diversity mediates plant diversity and ecosystem function relationships. *Nature* **546**, 145–147 (2017).
7.....Gazis, R. & Chaverri, P. Wild trees in the Amazon basin harbor a great diversity of beneficial endosymbiotic fungi: is this evidence of protective mutualism? *Fungal Ecology* **17**, 18–29 (2015).
8..Mejía, L. C. *et al.* Endophytic fungi as biocontrol agents of *Theobroma cacao* pathogens. *Biological Control* **46**, 4–14 (2008).
9.....Rodriguez, R. J. *et al.* Stress tolerance in plants via habitat-adapted symbiosis. *ISME J* **2**, 404–416 (2008).
10..…Müller, D. B., Vogel, C., Bai, Y. & Vorholt, J. A. The plant microbiota: systems-level insights and perspectives. *Annual review of genetics* **50**, 211–234 (2016).
11..Theobald, S. *et al.* Uncovering secondary metabolite evolution and biosynthesis using gene cluster networks and genetic dereplication. *Scientific Reports* **8**, 17957 (2018).
12.....Keller, N. P. Translating biosynthetic gene clusters into fungal armor and weaponry. *Nature Chemical Biology* **11**, 671–677 (2015).
13.....Cragg, G. M. & Newman, D. J. Natural products: A continuing source of novel drug leads. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1830**, 3670–3695 (2013).
14.....Slot, J. C. & Gluck-Thaler, E. Metabolic gene clusters, fungal diversity, and the generation of accessory functions. *Current Opinion in Genetics & Development* **58–59**, 17–24 (2019).
15.Charlop-Powers, Z. *et al.* Global biogeographic sampling of bacterial secondary metabolism. *eLife* **4**, e05048 (2015).
16..Nielsen, J. C. *et al.* Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in Penicillium species. *Nature Microbiology* **2**, 1–9 (2017).
17..·Rokas, A., Mead, M. E., Steenwyk, J. L., Raja, H. A. & Oberlies, N. H. Biosynthetic gene clusters and the evolution of fungal chemodiversity. *Nat. Prod. Rep.* **37**, 868–878 (2020).
18. Bills, G. F. & Gloer, J. B. Biologically Active Secondary Metabolites from the Fungi. *Microbiology Spectrum* **4**, (2016).
19.....Charlop-Powers, Z., Owen, J. G., Reddy, B. V. B., Ternei, M. A. & Brady, S. F. Chemical-biogeographic survey of secondary metabolism in soil. *Proc Natl Acad Sci USA* **111**, 3757 (2014).
20.....Mangan, S. A. *et al.* Negative plant–soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* **466**, 752–755 (2010).
21.....Marden, J. H. *et al.* Ecological genomics of tropical trees: how local population size and allelic diversity of resistance genes relate to immune responses, cosusceptibility to pathogens, and negative density dependence. *Molecular Ecology* **26**, 2498–2513 (2017).
22..…Guzman, F., Kulcheski, F. R., Turchetto-Zolet, A. C. & Margis, R. De novo assembly of Eugenia uniflora L. transcriptome and identification of genes from the terpenoid biosynthesis pathway. *Plant Science* **229**, 238–246 (2014).
23.....Eugenia uniflora (ID 504673) - BioProject - NCBI. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA504673.

24.....Thomas, D., Vandegrift, R., Roy, B. A., Hsieh, H.-M. & Ju, Y.-M. Spatial patterns of fungal endophytes in a subtropical montane rainforest of northern Taiwan. *Fungal Ecology* **39**, 316–327 (2019).
25....Ubbens, J. R. & Stavness, I. Deep Plant Phenomics: A Deep Learning Platform for Complex Plant Phenotyping Tasks. *Front. Plant Sci.* **8**, (2017).
26.....Bills, G. *et al.* New insights into the echinocandins and other fungal non-ribosomal peptides and peptaibiotics. *Nat. Prod. Rep.* **31**, 1348–1375 (2014).
27. Schläpfer, P. *et al.* Genome-Wide Prediction of Metabolic Enzymes, Pathways, and Gene Clusters in Plants. *Plant Physiology* **173**, 2041–2059 (2017).
28.. .Vestheim, H., Deagle, B. E. & Jarman, S. N. Application of Blocking Oligonucleotides to Improve Signal-to-Noise Ratio in a PCR. in *PCR Protocols* (ed. Park, D. J.) 265–274 (Humana Press, 2011). doi:10.1007/978-1-60761-944-4_19.
29.. .Haarmann, T. *et al.* The ergot alkaloid gene cluster in Claviceps purpurea: Extension of the cluster sequence and intra species evolution. *Phytochemistry* **66**, 1312–1320 (2005).
30.....Lorenz, N., Wilson, E. V., Machado, C., Schardl, C. L. & Tudzynski, P. Comparison of Ergot Alkaloid Biosynthesis Gene Clusters in Claviceps Species Indicates Loss of Late Pathway Steps in Evolution of C. fusiformis. *Appl. Environ. Microbiol.* **73**, 7185–7191 (2007).
31..Lorenz, N., Wilson, E. V., Machado, C., Schardl, C. L. & Tudzynski, P. Claviceps fusiformis strain PRL 1980 ergot alkaloid synthesis (EAS) gene cluster, complete sequence. *NCBI* http://www.ncbi.nlm.nih.gov/nuccore/EU006773.1 (2016).
32.....Hüttel, W., Youssar, L., Grüning, B. A., Günther, S. & Hugentobler, K. G. Echinocandin B biosynthesis: a biosynthetic cluster from Aspergillus nidulans NRRL 8112 and reassembly of the subclusters Ecd and Hty from Aspergillus pachycristatus NRRL 11440 reveals a single coherent gene cluster. *BMC Genomics* **17**, 570 (2016).
33.…Yao, T. *et al.* Characterization of the biosynthetic gene cluster of the polyene macrolide antibiotic reedsmycins from a marine-derived Streptomyces strain. *Microbial Cell Factories* **17**, 98 (2018).
34.....Morishita, Y., Zhang, H., Taniguchi, T., Mori, K. & Asai, T. The Discovery of Fungal Polyene Macrolides via a Postgenomic Approach Reveals a Polyketide Macrocyclization by trans-Acting Thioesterase in Fungi. *Org. Lett.* **21**, 4788–4792 (2019).
35.....Ge, B., Liu, B. & Zhang, K. Streptomyces albulus nystatin gene cluster, complete sequence; and hypothetical protein genes, complete cds. http://www.ncbi.nlm.nih.gov/nuccore/MG742725.1 (2018).
36.. Nofiani, R. *et al.* Strobilurin biosynthesis in Basidiomycete fungi. *Nature Communications* **9**, 3940 (2018).
37. Nofiani, R., Willis, C. L., Simpson, T. J. & Cox, R. J. Strobilurus tenacellus strobilurin gene cluster, partial sequence. *NCBI* http://www.ncbi.nlm.nih.gov/nuccore/KY070339.1 (2020).
38.....Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996–998 (2013).
39. Reddy, B. V. B., Milshteyn, A., Charlop-Powers, Z. & Brady, S. F. eSNaPD: A Versatile, Web-Based Bioinformatics Platform for Surveying and Mining Natural Product Biosynthetic Diversity from Metagenomes. *Chemistry & Biology* **21**, 1023–1033 (2014).
40. Dray, S. *et al.* Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs* **82**, 257–275 (2012).
41.....Bauman, D., Drouet, T., Dray, S. & Vleminckx, J. Disentangling good from bad practices in the selection of spatial or phylogenetic eigenvectors. *Ecography* **41**, 1638–1649 (2018).
42.Borcard, D., Gillet, F. & Legendre, P. Spatial analysis of ecological data. in *Numerical ecology with R* 227–292 (Springer, 2011).
43.....Legendre, P. & Legendre, L. *Numerical Ecology*. vol. 24 (Elsevier, 2012).
44.. Baddeley, A., Rubak, E. & Turner, R. *Spatial point patterns: methodology and applications with R*. (Chapman and Hall/CRC, 2015).

45.....Janzen, D. H. Herbivores and the number of tree species in tropical forests. *The American Naturalist* **104**, 501–528 (1970).
46..Connell, J. H., Den Boer, P. J. & Gradwell, G. R. Dynamics of populations. *Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands. chapter On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees* 298–313 (1971).
47...Johnson, D. J., Beaulieu, W. T., Bever, J. D. & Clay, K. Conspecific Negative Density Dependence and Forest Diversity. *Science* **336**, 904–907 (2012).
48.....Harms, K. E., Wright, S. J., Calderón, O., Hernández, A. & Herre, E. A. Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature* **404**, 493–495 (2000).
49. Christian, N., Herre, E. A., Mejia, L. C. & Clay, K. Exposure to the leaf litter microbiome of healthy adults protects seedlings from pathogen damage. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20170641 (2017).
50.....Rock-Blake, R., McCormick, M. K., Brooks, H. E. A., Jones, C. S. & Whigham, D. F. Symbiont abundance can affect host plant population dynamics. *American Journal of Botany* **104**, 72–82 (2017).
51..…Griffin, E. A., Harrison, J. G., McCormick, M. K., Burghardt, K. T. & Parker, J. D. Tree Diversity Reduces Fungal Endophyte Richness and Diversity in a Large-Scale Temperate Forest Experiment. *Diversity* **11**, 234 (2019).