Intro:

Plants face difficult times. With climate change, intensities of many types of plant stressors have increased, creating unpredictable new disturbance regimes, and many sources of plant stress are predicted to increase further[1]. Plant genomes contain many answers to these challenges, especially to environmental challenges which have acted as long-term selective pressures on plants[2]. However, plant genomes alone may not contain sufficient existing genetic variation or plasticity to ensure the success of ecologically or socially important plants in the face of decreased climatic stability and large-scale habitat changes[3], even with rigorous breeding and selection programs.

Other solutions may exist: members of the plant microbiome, with its collective abundance of microbial genetic resources, may protect their hosts through competition for resources[4,5] or active antagonism to potential pathogens[6–8]. In general, evidence has mounted rapidly that microbiomes in general are important determinants in plant health[9–13]. Together, it is proposed that the microbiome metagenome and the plant genome sum to make up the plant "hologenome", which is more extensive and more dynamic than the plant genome alone[14,15]. Microbiomes may become a point of leverage in the near future, wherein humans could increase resilience in plants of interest to environmental and biotic stressors by experimental manipulations of plant-associated microbial communities[16–19]. However, microbiome engineering has yet to achieve widespread success as an agricultural method, most probably due to an underestimation of the ecological complexity that maintains key members of microbiomes[20]. For microbiome engineering to succeed, the intricate dynamics of numerous host-pathogen-microbiome systems have to be examined with detailed ecological methods, to contribute to a more sophisticated understanding of the role of the microbiome in disease.

Here we propose to examine the *Pinus sylvestris / Sphaeropsis sapinea* host-pathogen interaction, with a focus on the augmentation of *Pinus* disease resistance by the natural products of its microbiome. *Pinus sylvestris* is an economically important and ecologically dominant forest tree throughout many regions of northern Eurasia with a long history of management and study[21]. *Sphaeropsis sapinea* (syn. *Diplodia sapinea*) is the causative agent of *Sphaeropsis* tip blight, but is also an extremely common asymptomatic wood endophyte and saprotroph in *P. sylvestris* and numerous other conifer species, with its pathogenic state appearing mostly in otherwise stressed hosts trees.[22]. It is thought to be an emerging pathogen in European forests, possibly due to climate-change induced stress on its host trees[23,24]. In two studies, the presence of endophytes have been suggested to mitigate symptoms of *Sphaeropsis* tip blight, either through antagonism or competition[5,25]. Important work has already been conducted by host researcher Dr. Terhonen and colleagues at University of Göttingen at the proposed site that greatly enhances the feasibility of this project proposal, including:

- extensive fungal endophyte culture surveys, completely archived available for use with the work proposed below[26]

- ITS2 metabarcode surveys of the mycobiome of the host tree in various stages of disease severity[22]
- RNAseq survey of response by host seedlings to inoculation with *S. sapinea*[27].

Questions, Hypothesis, and significance

Questions

Is a "mature" microbiome important to disease resistance in host plants?
If so, what are the mechanisms by disease resistance imbued by the microbiome to its host?

Hypotheses and predictions:

H1 - Plant genome function is augmented very early in life by recruitment of a microbial community.

H2 - This augmentation in plant function is imbued at least partly by chemical compounds originating from the secondary metabolome of the plant microbiome.

P1 - Seedlings with a reduced, less-diverse microbiome will show higher *Sphaeropsis* blight disease severity relative to enhanced-microbiome seedlings.

P2 - The hologenomes of enhanced-microbiome seedlings will contain numerous biosynthetic gene clusters (BSGs) that are greatly reduced or absent in the hologenome of reduced-microbiome seedlings.

P3- After inoculation with the pathogen *S. sapinea,* a subset of transcripts originating from the non-pathogen BSGs will be up-regulated in microbiome-enhanced seedlings.

P4 - Microbiome-enhanced seedlings will show differential phenotypic response to the pathogen: disease severity will covary with the presence of important biosynthetic gene clusters within their microbiome.

Significance:

As discussed above, the microbiome of plants has been shown to be important to fungal health historically, and may represent one of the great hopes for increasing climate resilience in plants in the current era of rapid change. However, the inverse of this hope must also be examined. Within each plant is an ecosystem as complex and diverse as the habitats of the host plant. While many have acknowledged the destabilizing effects of climate change on macro-ecological systems like forests, the effects of climate change on the interior ecosystem that is the microbiome are likely to be just as drastic and unpredictable.

In this way, systems such as the *Pinus sylvestris / Sphaeropsis* host/pathogen-system are particularly important, because they are representative of the new era of disease that will face forests, cropping systems, and other ecosystems: **with climate change, plant microbiomes will behave unpredictably.** Large, disruptive changes in ecological roles of their associated microorganisms are possible. The *Pinus/Sphaeropsis* system may serve as an example: a facultative, weak pathogen that is historically considered a typical member of the host microbiome has rather quickly elevated its status as an emerging threat to the pines of Europe and Asia, likely due to increases in stress that are ultimately attributable to climate change.

In these times of rapid change, therefore, the urgency of detailed research into the dynamics of microbiomes within plants has increased - the leverage points of the microbiome must be explored. To this end, biosynthetic gene clusters are particularly important to study, as their self-contained nature allows simplified detection of entire pathways, and relatively easier transformation into heterologous systems. Because so much of the capacity of microbial secondary metabolome resides in BGCs, they are a natural place to search for the mechanisms by which microbes interact with one another. In general, BCGs represent a low-hanging fruit in the search for helpful natural metabolic solutions for enhancement of plants and drug synthesis[28–34]. Their diversity is astounding, however, and here we propose to filter the search for important gene clusters by the use of transcriptomic and phenotypic approaches. This pipeline for discovery of useful microbial gene clusters may be generalizable to host-pathogen systems elsewhere.

Methods

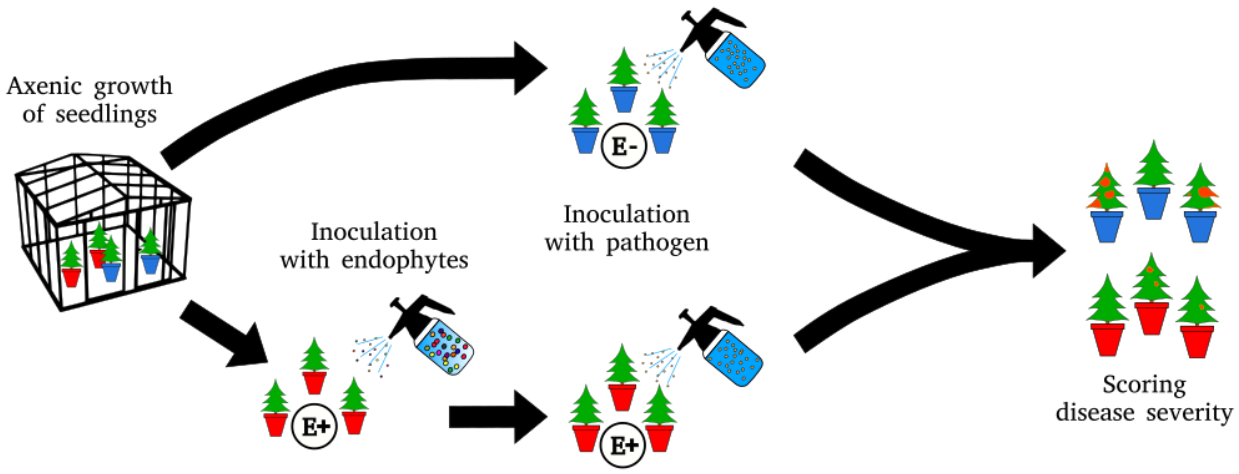
Summary:

We will experimentally reduce the microbiome of *P sylvestris* seedlings using axenic growth methods, and restore mycobiome of some seedlings by inoculating with endophytes associated with a low-disease state. Following this, we will challenge both reduced-microbiome and regular seedlings with *S. sapinea*, and observe changes in the holo-transcriptome. Gene clusters will be sampled to estimate the capacity of secondary metabolome of the microbiome of these seedlings, and the holo-transcriptome, and the holo-transcriptome will be monitored for changes after infection with *S. sapinea*, and correlated with disease severity, to determine what portion of the microbiome meta-genome is activated by the disease.

Seedling preparation, endophyte inoculation, and pathogen inoculation

In the first year, microbiome-reduced seedlings will be grown in mostly-sterile conditions[35] . Seedlings will be grown from seed from the research site of Blumenstein et al[22].,a stand of Scots pine (52.327653 °N 11.189848 °E) close to Behnsdorf in Saxony-Anhalt, Germany. This also the site of origin for the endophyte cultures to be used for E+ seedlings[26]. The most commonly recovered fungal endophytes observed in *P. sylvestris* from the previous study site will be incorporated into the endophyte inoculation mixtures[36]. Endophyte inoculation slurries will

be prepared using scraped conidia from agar endophyte cultures or from using lightly-homogenized mycelium from liquid cultures where necessary. Half of seedlings of *P. sylvestris* will be misted with the endophyte inoculation slurry, and half with sterile malt controls. Successful endophyte infection will be confirmed through re-isolation of endophytes from inoculated seedling material 1-2 weeks following inoculations. Pathogen inoculation will take place 1-several weeks after this, following Vornam et al.[27], using a strain of *S. sapinea* isolated from the study site.



Metabarcode-based exploration of fungal biosynthetic gene clusters:

Epiphytic and endophytic microbial communities will be sampled separately. To sample epiphytes of leaves will be submerged in tween / sterile water mixture, and vortex-ed to dislodge epiphyte biomass. This wash fluid will be centrifuged, and the pellet frozen for later DNA extraction. Leaf endophytes sampling will follow Thomas et al.[37], using a bleach-based surface sterilization method, followed by tissue homogenization and DNA extraction. Bulk DNA will be purified using column kits, and selectively PCR-amplified with PKS and NRPS primers (see below). Amplicon libraries will be prepped according to Illumina recommendations[38] and will be sequenced using a Illumina MiSeq platform sequencer, with a MiSeq V3 kit (2 x 300 bp). Metabarcoding results will be subject to standard bioinformatic pipelines to reduce sequencing artifacts and group highly similar reads[39]. Following this, amplicon reads will be evaluated and grouped using closest-known biosynthetic gene clusters using the software package eSNaPD[32,40] and public databases.

The biosynthetic capabilities of the microbiome will be modeled by sampling for the presence of two families of biosynthetic gene clusters. This will be accomplished using two metabarcode surveys of type I polyketide synthases (PKS) and nonribosomal peptide synthases (NRPS). These data will paint a general picture of some broad classes of secondary metabolites that tree microbiomes are predicted to produce 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)[41,42] and exist in diverse homologs that can be used to infer medium-to-fine grain portrait of the secondary metabolome of the plant microbiome. Primer pairs for this survey will be based on two studies conducted by Charlop-Powers et al.[32,43] with degenerate primers that target the ketosynthase (KS) region of type I polyketide synthases (PKS) and adenylation (AD) domains of nonribosomal peptide synthases (NRPS). Limited, initial exploration of these primers *in-silico* of KS-targeting primer sites has shown that they are present in both fungal and bacterial PKSes (see jupyter notebook), and AD primers are expected to perform as well in fungal NRPS searches. In soil, these primers tended to produce amplicon sizes of 350-500 bp, suitable for paired-end 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 originating from microbial sequences (see jupyter notebook). However, the possibility exists that plants are greatly under-sampled for type I PKS gene clusters[44]. Since our hypothesis focuses on the complementary role of the microbiome, we will attempt to enhance primers with plant-blocking oligonucleotides[45]. 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”. As both enzyme families (PKS and NRPS) are thought to occur only rarely in higher plants, and as plant-blocking nucleotides will be sought for primers, it should be possible to sample endophytic PKS and NRPS gene clusters as well as epiphytic.

We will emphasize the detection of enzymes in pathways of microbial secondary metabolites that could be "useful" to plant hosts in response to pathogen stress, including antifungal compounds and antioxidants. Our survey will broadly target all microbial gene clusters secondary metabolites with polyketide or non-ribosomal peptide backbones (and hybrids), which will likely uncover a vast array of pathways, of which only a subset will be applicable to our questions. As such, we will first examine results for:

- Fungal anti-mycotic compounds such as echinocandins (NRPS pathway)[42,46], polyene anti-mycotic including Nystatin and amphotericin (PKS)[47–49], and strobilurins (PKS)[50,51] in response to high pathogen pressure,
- Pigments including anthraquinones[52,53] and melanins[54–57] (PKS pathways) that may potentially act as antioxidants and UV-protectants.

RNAseq-based exploration of the host, pathogen, and microbiome transcriptome:

RNAseq survey methods will be applied using a tripartite approach: reads will be aligned to host using available genomic and transcriptomic resources[21,58], to pathogen using existing genomes[59,60], and by mining the remaining, unaligned reads for reads of fungal origins. Host and pathogen transcriptomes will be sequenced using standard Illumina platform sequencing and bioinformatic pipelines[27,61,62], including important transcriptomic studies already undertaken[27].

The latter method, mining of unaligned reads to characterize the transcriptome of the microbiome, is relatively uncommon and warrants further explanation:

RNAseq studies of large organisms very often leave 15-20% of total reads unaligned to the genome of the targeted organism. These reads are typically discarded, as RNA expression in the large organism is usually the focus of RNAseq studies. These discarded reads can originate from sequencer error and laboratory contamination, but many also originate from the transcriptome of the microbiome. Through careful curation of these "trash" read libraries, without minimum abundance cutoffs, we can draw an outline of metabolic activity in the microbiome. A similar method has recently been used successfully on the holo-transcriptome of sponges[70].

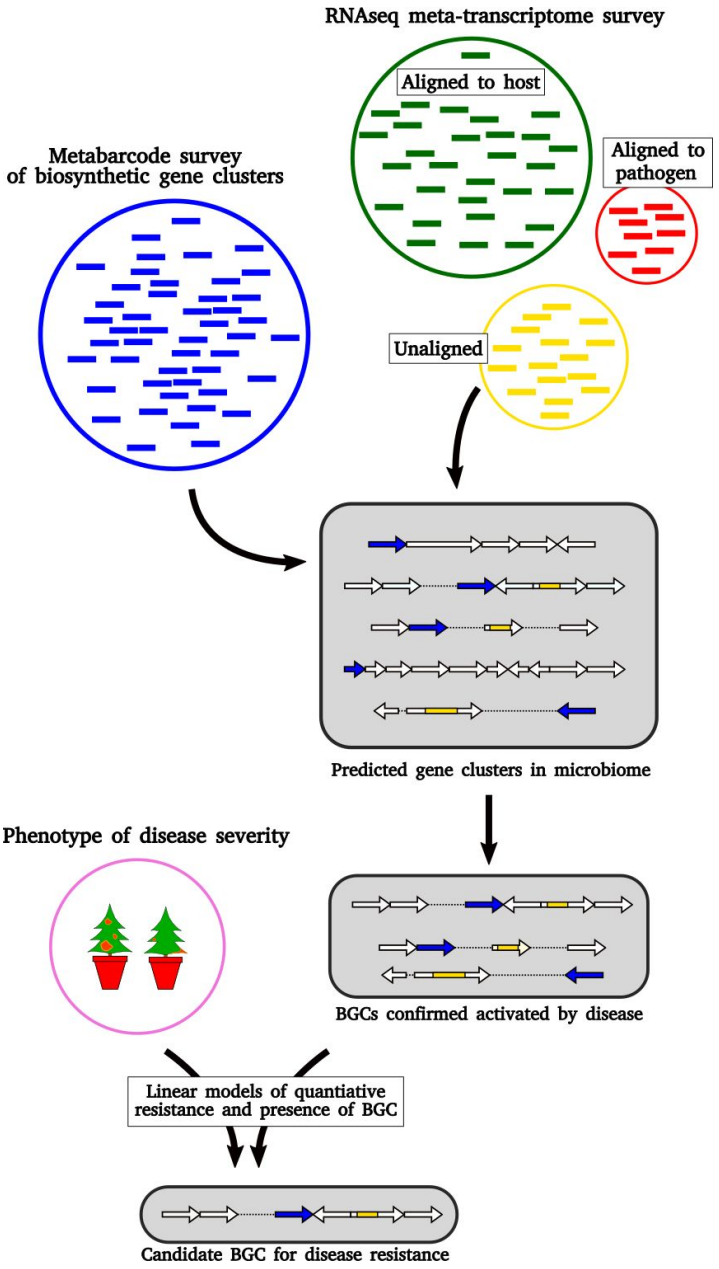
Preliminary test-of-concept:
In a first exploration of this method, we aligned 11 transcriptome samples from a publicly available dataset[63] to host *Pseudotsuga menziesii* genome[64]. For each of these samples, approximately 15% of reads are unalignable to the host. We blasted the reads not aligned to host genome against a database of ~10,000 fungal genes randomly selected from NCBI genbank. The average proportion of total reads (including those aligned to host) with high confidence matches to fungal genes in this database was 0.4%. This amounted to average of ~50,000 fungal reads per sample, with each sample containing an average of ~11,700,000 reads before alignment to host . If this rate applies to other sequencing events, in a single lane of illumina HiSeq 3000/4000 with 50bp per read, we may expect to see as many as 2,700,000 fungal reads. Some amount of these reads will be due to laboratory contamination or undigested ribosomal RNA, but presumably a large portion are reads originating from the microbiome. Prokaryotic reads were not examined but are of equal interest, and could be estimated similarly, provided RNAseq libraries were not created via poly-A enrichment methods. It is worth noting that this amount of read coverage is comparable to results from Roche 454 Pyrosequencing platforms, which have been successfully used for RNAseq studies[65,66]. Simply expanding the reference database of fungal genes to include the large number of published fungal genomes will likely increase this number of recovered microbiome reads.

Phenotyping of plant disease

All seedlings will be photo documented immediately prior to inoculation by *Sphaeropsis* and at weekly intervals following for several weeks until mortality of all diseased seedlings and/or stabilization of disease symptoms in all living samples occurs. Camera setup will standardized and pots will be marked to allow repeated standardized photography from several angles of all seedlings. Disease severity will be classified using machine-learning algorithms specialized for plant-phenotyping computer vision problems, using the PlantCV package in Python[67]. If additional training data are required, twigs of varying disease severity will be collected from the field site, where *Sphaeropsis* tip blight is common.

Synthesis: analysis of gene cluster metabarcoding, RNAseq surveys, and disease severity phenotyping:

The two proposed sequence-based surveys will allow us to make high-confidence calls about the presence and activity of important gene clusters, wherever the two datasets coincide. There is a substantial body of literature, software and public databases cataloguing and predicting biosynthetic gene clusters from synthase-barcodes[40,43,68,69]. Dr. Terhonen and colleagues retain a substantial library of endophyte cultures isolated from *P. sylvestris* which can be examined for BGCs alongside the ecological samples to provide further information on the microbiome BGCs that are likely present, specific to the current *P. sylvestris* system. The Information from these sources will help discovery of un-aligned fungal transcripts that originate from "nearby" genes within predicted gene clusters, even if the exact PKS or NRPS match is not detected in the transcriptome. This will increase the radius of detection of active PKS and NRPS gene clusters in the RNAseq library from samples in the disease state. Taken together, these two methods will give us an estimate of the biosynthetic potential of plant microbiomes, and an estimate of the portion of this potential that is involved in pathogen response. Linear models modeling the relationship between phenotype (disease severity) and the presence/absence and activity of gene clusters can then help to filter to gene clusters that may be important to host disease resistance.



Timeline:

Year 1: Seedlings will be planted as soon as possible in year one, to allow for use of seedlings at field in year two. While seedlings are growing, primers for secondary metabolite gene clusters will be refined for use on fungal endophytes and tested *in-silico* and *in-vivo* using available culture libraries and *P. sylvestris* tissue. Bioinformatic methods will be further developed using

existing transcriptomes from *P. sylvestris*[21,27] and available fungal genomes from confirmed endophytes of *P. sylvestris*. E+ inoculation will be prepared as seedlings mature.

Year 2: In late spring of year 2, E+ Seedlings will be inoculated with endophyte community inoculum, and the project enters the intensive sampling phase. Beginning 6-8 weeks following inoculations, bench preparation of samples and sequencing will take place. Analysis and writing phase will be in the final six months of the second year.

References

1. Seidl, R. *et al.* Forest disturbances under climate change. *Nature Climate Change* **7**, 395–402 (2017).
2. Carmody, M., Waszczak, C., Idänheimo, N., Saarinen, T. & Kangasjärvi, J. ROS signalling in a destabilised world: A molecular understanding of climate change. *Journal of Plant Physiology* **203**, 69–83 (2016).
3. Anderson, J. T. & Song, B.-H. Plant adaptation to climate change—Where are we? *Journal of Systematics and Evolution* **58**, 533–545 (2020).
4. Blumenstein, K. *et al.* Nutritional niche overlap potentiates the use of endophytes in biocontrol of a tree disease. *BioControl* **60**, 655–667 (2015).
5. Oliva, J., Ridley, M., Redondo, M. A. & Caballol, M. Competitive exclusion amongst endophytes determines shoot blight severity on pine. *Functional Ecology* **35**, 239–254 (2021).
6. Evans, H. C., Holmes, K. A. & Thomas, S. E. Endophytes and mycoparasites associated with an indigenous forest tree, *Theobroma gileri*, in Ecuador and a preliminary assessment of their potential as biocontrol agents of cocoa diseases. *Mycological Progress* **2**, 149–160 (2003).
7. Tanney, J. B., McMullin, D. R. & Miller, J. D. Toxigenic foliar endophytes from the Acadian forest. in *Endophytes of forest trees: biology and applications* (eds. Pirttilä, A. M. & Frank, A. C.) 343–381 (Springer International Publishing, 2018). doi:10.1007/978-3-319-89833-9_15.
8. Tanney, J. B., McMullin, D. R., Green, B. D., Miller, J. D. & Seifert, K. A. Production of antifungal and antiinsectan metabolites by the Picea endophyte Diaporthe maritima sp. nov. *Fungal Biology* **120**, 1448–1457 (2016).
9. 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).
10. 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).
11. Mejía, L. C. *et al.* Endophytic fungi as biocontrol agents of *Theobroma cacao* pathogens. *Biological Control* **46**, 4–14 (2008).
12. Rodriguez, R. J. *et al.* Stress tolerance in plants via habitat-adapted symbiosis. *ISME J* **2**, 404–416 (2008).
13. 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).
14. Rosenberg, E. & Zilber-Rosenberg, I. Microbes drive evolution of animals and plants: the hologenome concept. *mBio* **7**, e01395-15 (2016).
15. Lemanceau, P., Blouin, M., Muller, D. & Moënne-Loccoz, Y. Let the core microbiota be functional. *Trends in plant science* **22**, 583–595 (2017).
16. Tsoi, R., Dai, Z. & You, L. Emerging strategies for engineering microbial communities. *Biotechnology Advances* **37**, 107372 (2019).
17. Foo, J. L., Ling, H., Lee, Y. S. & Chang, M. W. Microbiome engineering: Current applications and its future. *Biotechnology Journal* **12**, 1600099 (2017).
18. Busby, P. E. *et al.* Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLOS Biology* **15**, e2001793 (2017).
19. Terhonen, E., Blumenstein, K., Kovalchuk, A. & Asiegbu, F. O. Forest Tree Microbiomes and Associated Fungal Endophytes: Functional Roles and Impact on Forest Health. *Forests* **10**, 42 (2019).
20. Vandermeer, J., Perfecto, I. & Philpott, S. Ecological Complexity and Pest Control in Organic Coffee Production: Uncovering an Autonomous Ecosystem Service. *BioScience* **60**, 527–537 (2010).
21. Pyhäjärvi, T., Kujala, S. T. & Savolainen, O. 275 years of forestry meets genomics in Pinus sylvestris. *Evolutionary Applications* **13**, 11–30 (2020).
22. Blumenstein, K., Langer, G., Bußkamp, J., Langer, E. & Terhonen, E. The Opportunistic Pathogen Sphaeropsis Sapinea is Found to be one of the Most Abundant Fungi in Symptomless and Diseased Scots Pine in Central-Europe. (2020).
23. Fabre, B., Piou, D., Desprez-Loustau, M.-L. & Marçais, B. Can the emergence of pine Diplodia shoot blight in France be explained by changes in pathogen pressure linked to climate change? *Global Change Biology* **17**, 3218–3227 (2011).
24. Bosso, L. *et al.* Predicting current and future disease outbreaks of Diplodia sapinea shoot blight in Italy: species distribution models as a tool for forest management planning. *Forest Ecology and Management* **400**, 655–664 (2017).
25. SantamaríaOscar, R, S. & R, S. Interaction between Diplodia pinea or Diplodia scrobiculata and fungal endophytes isolated from pine shoots. *Canadian Journal of Forest Research* (2012) doi:10.1139/x2012-132.
26. Bußkamp, J., Langer, G. J. & Langer, E. J. Sphaeropsis sapinea and fungal endophyte diversity in twigs of Scots pine (Pinus sylvestris) in Germany. *Mycol Progress* **19**, 985–999 (2020).
27. Vornam, B. *et al.* Response of Scots pine (Pinus sylvestris) seedlings subjected to artificial infection with the fungus Sphaeropsis sapinea. *Plant Mol Biol Rep* **37**, 214–223 (2019).
28. Theobald, S. *et al.* Uncovering secondary metabolite evolution and biosynthesis using gene cluster networks and genetic

dereplication. *Scientific Reports* **8**, 17957 (2018).
29. Keller, N. P. Translating biosynthetic gene clusters into fungal armor and weaponry. *Nature Chemical Biology* **11**, 671–677 (2015).
30. 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).
31. 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).
32. Charlop-Powers, Z. *et al.* Global biogeographic sampling of bacterial secondary metabolism. *eLife* **4**, e05048 (2015).
33. 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).
34. 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).
35. 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).
36. Arnold, A. E. *et al.* Fungal endophytes limit pathogen damage in a tropical tree. *PNAS* **100**, 15649–15654 (2003).
37. 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).
38. Illumina. Indexed Sequencing Overview Guide (15057455). *Illumina Proprietary Document #15057455v04* 15 (2020).
39. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996–998 (2013).
40. 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).
41. Bills, G. F. & Gloer, J. B. Biologically Active Secondary Metabolites from the Fungi. *Microbiology Spectrum* **4**, (2016).
42. Bills, G. *et al.* New insights into the echinocandins and other fungal non-ribosomal peptides and peptaibiotics. *Nat. Prod. Rep.* **31**, 1348–1375 (2014).
43. 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).
44. Schläpfer, P. *et al.* Genome-Wide Prediction of Metabolic Enzymes, Pathways, and Gene Clusters in Plants. *Plant Physiology* **173**, 2041–2059 (2017).
45. 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.
46. 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).
47. 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).
48. 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).
49. 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).
50. Nofiani, R. *et al.* Strobilurin biosynthesis in Basidiomycete fungi. *Nature Communications* **9**, 3940 (2018).
51. 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).
52. Gessler, N., Egorova, A. & Belozerskaya, T. Fungal anthraquinones. *Applied Biochemistry & Microbiology* **49**, 85–99 (2013).
53. Mapari, S. A. S., Meyer, A. S. & Thrane, U. Evaluation of *Epicoccum nigrum* for growth, morphology and production of natural colorants in liquid media and on a solid rice medium. *Biotechnol Lett* **30**, 2183–2190 (2008).
54. Jacobson, E. S. & Tinnell, S. B. Antioxidant function of fungal melanin. *Journal of Bacteriology* **175**, 7102–7104 (1993).
55. Shcherba, V. V., Babitskaya, V. G., Kurchenko, V. P., Ikonnikova, N. V. & Kukulyanskaya, T. A. Antioxidant properties of fungal melanin pigments. *Appl Biochem Microbiol* **36**, 491–495 (2000).
56. Singaravelan, N. *et al.* Adaptive Melanin Response of the Soil Fungus Aspergillus niger to UV Radiation Stress at “Evolution Canyon”, Mount Carmel, Israel. *PLOS ONE* **3**, e2993 (2008).
57. Wu, Y., Shan, L., Yang, S. & Ma, A. Identification and antioxidant activity of melanin isolated from Hypoxylon archeri, a companion fungus of Tremella fuciformis. *Journal of Basic Microbiology* **48**, 217–221 (2008).
58. Wachowiak, W., Trivedi, U., Perry, A. & Cavers, S. Comparative transcriptomics of a complex of four European pine species. *BMC Genomics* **16**, 234 (2015).
59. Diplodia sapinea CMW 190 (ID 215898) - BioProject - NCBI. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA215898.
60. Diplodia sapinea strain:CMW39103 (ID 242796) - BioProject - NCBI. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA242796.
61. Zheng, X. *et al.* The regulatory network for petal anthocyanin pigmentation is shaped by the MYB5a/NEGAN transcription factor in Mimulus. *Genetics* **217**, (2021).
62. Conesa, A. *et al.* A survey of best practices for RNA-seq data analysis. *Genome Biology* **17**, 13 (2016).
63. Reichman, J. R. *et al.* Douglas-Fir (*Pseudotsuga menziesii* (Mirb.) Franco) Transcriptome profile changes induced by diesel emissions generated with CeO2 nanoparticle fuel borne catalyst. *Environmental Science & Technology* (2018) doi:10.1021/acs.est.8b02169.
64. Neale, D. B. *et al.* The Douglas-Fir genome sequence reveals specialization of the photosynthetic apparatus in Pinaceae. *G3: Genes, Genomes, Genetics* g3.300078.2017 (2017) doi:10.1534/g3.117.300078.
65. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**, 57–63 (2009).
66. Hook, S. E. *et al.* 454 pyrosequencing-based analysis of gene expression profiles in the amphipod Melita plumulosa: Transcriptome assembly and toxicant induced changes. *Aquatic Toxicology* **153**, 73–88 (2014).
67. PlantCV. https://plantcv.readthedocs.io/en/stable/.

68. Weber, T. *et al.* antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* **43**, W237–W243 (2015).
69. Owen, J. G. *et al.* Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. *PNAS* **110**, 11797–11802 (2013).
70. Germer, J., Cerveau, N. & Jackson, D. J. The Holo-Transcriptome of a Calcified Early Branching Metazoan. *Front. Mar. Sci.* **4**, (2017).