# Enemy or ally: a genomic approach to elucidate the lifestyle of *Phyllosticta citrichinaensis*

Valerie A. Buijs[1,2]*, Johannes Z. Groenewald[1], Sajeet Haridas[3], Kurt M. LaButti[3], Anna Lipzen[3], Francis M. Martin[4], Kerrie Barry[3], Igor V. Grigoriev[3,5], Pedro W. Crous[1,2], Michael F. Seidl[6]*

[1]Evolutionary Phytopathology, Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

[2]Laboratory of Phytopathology, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[3]US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

[4]Institut National de la Recherche Agronomique, UMR INRA-université de Lorraine "Interaction Arbres/Microorganismes", Champenoux, France

[5]Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, CA, USA

[6]Theoretical Biology & Bioinformatics, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.

*Corresponding authors.

V. A. Buijs, Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands. *E-mail address*: v.buijs@wi.knaw.nl

M. F. Seidl, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands. *E-mail address:* m.f.seidl@uu.nl

## Abstract

Members of the fungal genus *Phyllosticta* can colonize a variety of plant hosts, including several *Citrus* species such as *Citrus sinensis* (orange), *Citrus limon* (lemon), and *Citrus maxima* (pomelo). Some *Phyllosticta* species have the capacity to cause disease, such as Citrus Black Spot, while others have only been observed as endophytes. Thus far, genomic differences underlying lifestyle adaptations of *Phyllosticta* species have not yet been studied. Furthermore, the lifestyle of *Phyllosticta citrichinaensis* is ambiguous, as it has been described as a weak pathogen but Koch's postulates may not have been established and the presence of this species was never reported to cause any crop or economic losses. Here, we examined the genomic differences between pathogenic and endophytic *Phyllosticta* spp. colonizing *Citrus* and specifically aimed to elucidate the lifestyle of *Phyllosticta citrichinaensis*. We found several genomic differences between species of different lifestyles, including groups of genes that were only present in pathogens or endophytes. We also observed that species, based on their carbohydrate active enzymes, group independent of their phylogenetic association, and this clustering correlated with trophy prediction. *Phyllosticta citrichinaensis* shows an intermediate lifestyle, sharing genomic and phenotypic attributes of both pathogens and endophytes. We thus present the first genomic comparison of multiple citrus-colonizing pathogens and endophytes of the genus *Phyllosticta*, and therefore provide the basis for further comparative studies into the lifestyle adaptations within this genus.

## 1 – Introduction

Fungal and oomycete phytopathogens are a major threat to global food security (Fisher *et al.* 2012). Despite many technological and methodological developments, such as the development of disease-resistant crops, this threat remains a pressing concern for humankind due to emergence of new or adapted species, and a lack of in-depth understanding of disease mechanisms and their genomic basis (Fudal *et al.* 2009; Singh *et al.* 2011; Fisher *et al.* 2012).

Plant associated fungi and oomycetes can be broadly classified as pathogens, endophytes, or saprotrophs, i.e., they are classified based on their capacity to cause disease symptoms. Furthermore, these microbes can be linked to five different trophic classes based on their specific feeding behavior (Kabbage *et al.* 2015). Necrotrophs are characterized as pathogens that feed on dead tissue, biotrophs as pathogens that feed on living tissue, and hemibiotrophs are pathogens that go through an initial biotrophic phase before switching to a necrotrophic phase (Oliver and Ipcho 2004). In the same classification model, non-pathogenic species that live within a plant are classified as endophytes, while species that live only on decaying plant material are referred to as saprotrophs. This classification model, which is mainly based on observational data, clearly has limitations, for instance when one species is classified as a necrotroph when interacting with one host but as biotroph when interacting with another (Veloso and Van Kan 2018). Consequently, much research in recent years has focused on establishing the genomic basis underlying differences between species that exhibit different lifestyles. Uncovering these genomic signatures would provide a more reliable method of classification and an increased understanding of host colonization and disease mechanisms, which is of significant importance in developing more effective disease management strategies (Haridas et al., 2020; Möller and Stukenbrock, 2017; O'Connell et al., 2012; Ohm et al., 2012; Plissonneau et al., 2017; Spanu, 2012).

A common feature of various investigations into the genomic basis of pathogenicity is the identification of specific adaptations present in one lifestyle but absent or reduced in the other (Klosterman *et al.* 2011; Gardiner *et al.* 2012; Kim *et al.* 2016). A major current focus is the study of effectors, which are often defined as small secreted proteins that play an important role in establishing the interaction with the host, for instance by degrading the host cell wall or shielding the pathogen from detection by the host immune system (Rovenich *et al.* 2014; Lo Presti *et al.* 2015; Fouché *et al.* 2018). Effectors are often shared by strains and sometimes by species that colonize the same host (Chiapello *et al.* 2015; van Dam *et al.* 2017), and on rare occasions are even passed on to a separate species through horizontal gene transfer

(Gardiner *et al.* 2012; van Dam and Rep 2017). The identification of known effectors and other genes that are present only in species of a specific lifestyle can therefore provide useful information when studying the genomic basis of pathogenicity (Gibriel *et al.* 2016). However, as hosts rapidly evolve mechanisms to recognize effectors to re-establish immunity, effectors frequently mutate resulting in rapid effector diversification to avoid detection by the host immune system. Thus, effector repertoires in separate fungal lineages may differ significantly (Rovenich *et al.* 2014; Lo Presti *et al.* 2015).

Carbohydrate active enzymes (CAZymes) play diverse roles in degradation and biosynthesis of carbohydrates such as those found in plant cell walls. For example, plant pathogens can utilize CAZymes to penetrate the host cell wall to establish symbiosis and to liberate carbohydrates from host tissues for growth and reproduction (van den Brink and de Vries 2011; Kubicek *et al.* 2019). Thus, CAZymes can also contribute to virulence, and differences in CAZyme repertoires can mediate microbial lifestyle differences (ten Have *et al.* 2002; King *et al.* 2011; Hane *et al.* 2020). Consequently, CAZymes have been used to propose new lifestyle classification models for oomycete and fungal species (Hane et al., 2020). For instance, Hane and colleagues recently proposed five new trophic classes based on primary nutrient source preferences as approximated by the presence and abundance of CAZymes directly predicted from genome assemblies (Hane et al., 2020): polymertrophs correspond best to necrotrophs, and have received their name due to a preference for polymeric carbohydrates as primary nutrient sources. In contrast, monomertrophs, which correspond best to symbionts and biotrophs, prefer monomeric primary nutrient sources. Mesotrophs are an intermediate group utilizing both monomeric and polymeric nutrient sources, and correspond best to hemibiotrophs. Vasculartrophs are similar to hemibiotrophs, but also include species commonly classified as wilts, anthracnoses, and rots. The saprotrophic class remains a separate group, encompassing species that feed mainly on decaying plant material. These new trophic classes were proposed based on a broad and phylogenetically diverse set of phytopathogenic fungi and oomycetes (Hane et al. 2020), including several members of

*Dothideomycetes*, a diverse class including many plant-associated fungi as well as fungi adapted to other lifestyles such as marine or soil environments (Haridas *et al.* 2020).

Within the *Dothideomycetes*, members of the genus *Phyllosticta* are particularly well suited to study the genomic basis of lifestyle adaptation and phytopathogenicity. *Phyllosticta* contains at least 50 species that are able to associate with a broad range of plant hosts (Wikee *et al.* 2013b), but species that colonize *Citrus* are of particular interest as they comprise both endophytes and pathogens while being phylogenetically closely related (Wikee *et al.* 2013b; Guarnaccia *et al.* 2019). The most well-known species of this genus is *Phyllosticta citricarpa* which causes Citrus Black Spot, a disease causing significant economic losses worldwide and which therefore has a quarantine status in Europe (Kotzé 2000; European Food Safety Authority 2014; Eustáquio Lanza *et al.* 2018). *Phyllosticta paracitricarpa* is closely related to *P. citricarpa*, bears a strong morphological resemblance and appears to cause similar disease symptoms on citrus (Guarnaccia *et al.* 2017). Other pathogens include *P. citriasiana*, described from several citrus hosts in Asia, *P. citrimaxima*, described from *Citrus maxima* in Thailand, and *P. citrichinaensis*, described as a weak pathogen from several citrus hosts in China (Wulandari *et al.* 2009; Wang *et al.* 2012; Wikee *et al.* 2013b). Endophytic species within the genus are *P. capitalensis* with a very broad host range and present on all continents, *P. paracapitalensis*, known from citrus in Europe, and *P. citribraziliensis*, currently known only from citrus in Brazil (Glienke *et al.* 2011; Wikee *et al.* 2013a; Guarnaccia *et al.* 2017).

Genomes were recently published for *P. capitalensis*, *P. citriasiana*, *P. citribraziliensis*, *P. citricarpa*, *P. citrichinaensis*, and *P. paracitricarpa*, with genome sizes ranging between 29–32 Mb (Guarnaccia *et al.* 2019). These genomes pave the way for comparative genomic studies aimed to disentangle lifestyle differences within the genus *Phyllosticta* (Guarnaccia *et al.* 2019). Although both mating types (MAT1-1 and MAT1-2) are reported to exist for both endophytic and pathogenic species (Guarnaccia *et al.* 2019; Petters-Vandresen *et al.* 2020), the pathogenic strains for which genomes have been published are all heterothallic and only

of the MAT1-2 mating type. In contrast, the sequenced strain of the endophytic *P. citribraziliensis* is heterothallic and of the MAT1-1 mating type, while the other sequenced endophyte*, P. capitalensis*, is homothallic and therefore contains both mating types genes (Guarnaccia *et al.* 2017, 2019; Petters-Vandresen *et al.* 2020). *Phyllosticta citrichinaensis* is also homothallic, but the MAT1-2 idiomorph in *P. citrichinaensis* is present in a separate location from the mating type locus (Petters-Vandresen *et al.* 2020). This phenomenon sets *P. citrichinaensis* apart from the other species in the genus as the configuration of the mating type locus is typically very conserved amongst *Phyllosticta* species (Petters-Vandresen *et al.* 2020).

Previous comparative analyses between pathogenic and endophytic *Phyllosticta* spp. have been hampered by the quality of genomes and the availability of only a single endophyte genome (*P. capitalensis*), which was relatively distantly related to the species it was compared to, and consequently genomic adaptations towards these two broad lifestyles remained unclear (Wikee *et al.* 2013b; Rodrigues *et al.* 2019; Wang *et al.* 2020). Therefore, a comparison of new and high-quality genomes which includes multiple species of different lifestyles could provide the necessary foundation to finally discovering the genomic underpinning for phytopathology in *Phyllosticta*, which is essential for the development of better disease management strategies.

*Phyllosticta citrichinaensis* was originally described as a weakly aggressive pathogen of several citrus hosts in China as it was isolated from freckles or spots on fruits or leaves of citrus (Wang *et al.* 2012). However, lesions never exhibited typical pycnidia, the presence of this species was never reported to cause any crop or economic losses and Koch's postulates may not have been established (Wang *et al.* 2012), and thus its lifestyle remains ambiguous. *P. citribraziliensis* is a very close relative of *P. citrichinaensis* and has been described only as an endophyte from Brazil. Therefore, if these two species were certain to be of two different lifestyles, these species would be ideal to study pathogenicity in *Phyllosticta*. As the genome

of *P. citrichinaensis* has not been included in earlier comparative work focused on lifestyle differences (Rodrigues *et al.* 2019; Wang *et al.* 2020), a thorough study of its genome and comparison to the genomes of the other species in this genus could provide valuable information on this species' lifestyle as well as genomic underpinning of disease mechanisms of other species in this genus. Here, we present the first comparative genomics study using multiple complete genomes of two endophytic and three phytopathogenic *Phyllosticta* species and established genomic differences between species of different lifestyles within this genus. In addition, we use these data in an attempt to elucidate the lifestyle of the ambiguous *P. citrichinaensis*.

## 2 - Materials and Methods

### 2.1 – Sequencing, annotation, genome quality and availability

All non-*Phyllosticta* genomes were previously published (Haridas *et al.* 2020) and are available on MycoCosm (https://mycocosm.jgi.doe.gov/Dothideomycetes; Grigoriev et al., 2014). The database identifiers (DBIDs) that are given by the Joint Genome Institute (JGI) to identify specific genomes, and which can be used to access the genome's online portal (https://mycocosm.jgi.doe.gov/DBID, e.g., https://mycocosm.jgi.doe.gov/Aaoar1) are listed in Suppl. Table 1. Seven of the eight *Phyllosticta* genomes included in our analyses were also previously published (Guarnaccia *et al.* 2019), and are available on MycoCosm (mycocosm.jgi.doe.gov/Phyllosticta).

*Phyllosticta citrichinaensis* liquid cultures (250 ml Malt peptone broth in 500 ml Erlenmeyer flasks) were incubated at 25 ºC and 180 rpm for 10 to 14 days, after which genomic DNA was isolated using the Qiagen Genomic-tip 100/G kit and the Qiagen Genomic DNA Buffer Set. The genome assembly of *P. citrichinaensis* genome (CBS 129764) was generated by the JGI using the PacBio long-read sequencing technology. Long-read sequencing data was assembled using Flye and the genome assembly was annotated using the JGI Annotation pipeline (Grigoriev *et al.* 2014). The genome assembly and annotation are

194 available via the MycoCosm platform (https://mycocosm.jgi.doe.gov/Pcit129764). Quality
195 assessments were performed using BUSCO 4.1.4 (Manni *et al.* 2021) and QUAST 5.0.2
196 (Gurevich *et al.* 2013) using default parameters. One-to-one whole-genome comparisons
197 were performed using PROmer (default settings), which is part of the MUMMer 3.25 conda
198 package (Marçais *et al.* 2018) and plotted with mummerplot using the --filter and --fat
199 parameters. We used OrthoFinder 2.2.6 (Emms and Kelly 2019) to identify ortholog groups
200 (OGs) across all 116 fungal genome annotations (Suppl. Table 2A). Ortholog groups unique
201 to species of a specific lifestyles were identified using UpSetR (Suppl. Table 2B, Conway et
202 al., 2017).
203
204 **2.2 – Secreted proteins and effectors**
205 We used SignalP 5.0b Linux x86_64 (Almagro Armenteros et al., 2019) to predict secreted
206 proteins in the predicted proteomes of all 116 fungal genomes, and subsequently applied
207 EffectorP 2.0 (Sperschneider *et al.* 2016) to predict effectors within the secretomes. We
208 visualized the distribution of OGs of which 50% or more of the genes were predicted to be a
209 secreted protein or an effector, by generating a clustered heatmap in R using the
210 ComplexHeatmap package (Gu *et al.* 2016).
211
212 **2.3 – Carbon utilization and CATAstrophy**
213 Carbon growth studies were performed as described previously (Buijs *et al.* 2021). In short,
214 1-mm-diameter plugs from 2-week-old colony edges of *Phyllosticta* species were inoculated
215 on 35 different carbon sources and incubated at 25 ºC until the largest colony reached the
216 edge of a 35-mm-diameter plate. As different *Phyllosticta* species demonstrate different
217 growth speeds, this moment fell on different days after inoculation (between five and ten days).
218 When the largest colony of a species reached the edge of a plate, colony diameters were
219 measured on all sources and images were taken using a standard camera setup. This
220 approach was chosen to be able to compare species with different growth speeds. All growth
221 studies were performed in duplicate. Measurements were averaged and used to generate a

clustered heatmap using the ComplexHeatmap package (Gu *et al.* 2016) in R (R Core Team, 2021).

We used CATAstrophy to predict lifestyles from CAZyme repertoires (Hane *et al.* 2020). To this end, we first used hmmpress to generate a local HMMER database of dbCAN 8 (Zhang *et al.* 2018). We then queried all 116 predicted proteomes with the local dbCAN HMMs database using hmmscan with the –domtblout parameter to create a domain table for each proteome. CATAstrophy was then ran on all 116 domain tables using parameters -p, -c, --model v8 and --format hmmer_domtab. The heatmap was created by identifying all OGs (as created previously using OrthoFinder) that contained a CAZyme, counting the number of genes in each OG for each species, and then generating a heatmap using the ComplexHeatmap package in R. Empty columns, e.g. CAZyme families for which no genes present in these species, were filtered out of the heatmap for visualization purposes, but are present in the original data (Suppl. Table 5).

## 3 - Results

### 3.1 – *Phyllosticta* genome assemblies are of high quality

Lifestyle differences are often driven by genomic adaptations (Ohm *et al.* 2012; Kabbage *et al.* 2015; Haridas *et al.* 2020), and we hypothesize that this also applies to *Phyllosticta* species with different lifestyles. Taxonomically, *Phyllosticta* belongs to *Dothideomycetes*, a fungal class with extensive genomic resources (Wikee *et al.* 2013b; Haridas *et al.* 2020). To enable studies in lifestyle differences in *Phyllosticta,* we made use of eight *Phyllosticta* genome sequences, seven of which were assembled and published previously (Guarnaccia *et al.* 2017, 2019). Here, we performed genome sequencing of *P. citrichinaensis* (CBS 129764), which is the second genome of this species to be sequenced, thereby enabling us to also evaluate intra-species variation. We included two genome assemblies of *P. citrichinaensis*, as well as the genome assemblies of the endophyte *P. citribraziliensis*, the closest relative of *P. citrichinaensis*, those of the two pathogenic species *P. citricarpa* and *P. paracitricarpa*, which

are very closely related (Fig. 1A, also see Guarnaccia et al., 2019), of the pathogenic species *P. citriasiana*, and of the endophyte *P. capitalensis*, which is phylogenetically the least related to the other *Phyllosticta* species (Fig. 1A). As these genome assemblies include multiple species of both lifestyles (pathogens and endophytes), comparative genomics may reveal the genomic underpinning for lifestyle adaptations within this genus, and ultimately aid in determining the lifestyle of *P. citrichinaensis*.

To determine genome assembly size, fragmentation, and annotation completeness of the *Phyllosticta* genomes, we used QUAST (Gurevich *et al.* 2013) and BUSCO (Manni *et al.* 2021), and compared the results to 100 previously published *Dothideomycete* genomes as well as eight genomes of fungal species outside of *Dothideomycetes* (Fig. 1A**;** Suppl. Table 1, Haridas et al., 2020). Compared to the other fungal species, *Phyllosticta* genome assemblies are of a slightly smaller size (29–32 Mb) as opposed to an average of 40 Mb in other *Dothideomycetes* (Fig. 1A). The *Phyllosticta* genome assemblies have a low number of contigs, namely 14 to 152 contigs compared with on average 471 contigs, and BUSCO scores between 93.3% and 95.8%, suggesting that the *Phyllosticta* genome assemblies are *en par* or above average quality in terms of genome contiguity and completeness compared to the other *Dothideomycetes* genomes, which should facilitate further comparative analyses.

### 3.2 – *Phyllosticta* species show mesosynteny to species within the *Dothideomycetes*

*Dothideomycetes* are known to show different patterns of genome conservation as compared to other fungal classes: intra-chromosomal rearrangements lead to a seemingly random reshuffling of gene content within individual chromosomes, while genes themselves remain well conserved (Hane *et al.* 2011; Ohm *et al.* 2012), a pattern that has previously been named mesosynteny (Hane *et al.* 2011). To study the gene order conservation in *P. citrichinaensis* and in other *Phyllosticta* species compared with other fungal genomes, we performed whole-genome alignments of *P. citrichinaensis* to all other *Phyllosticta* and to two more distantly related fungi (Fig. 1B). When comparing the two *P. citrichinaensis* strains, we observed a clear

278 pattern of macrosynteny with only a few inversions and translocations (Fig. 1B); a similar
279 pattern could also be observed when comparing *P. citrichinaensis* to *P. citricarpa* (Fig. 1B)
280 and *P. citribraziliensis* (Suppl. Fig. 1A). Interestingly, when comparing *P. citrichinaensis* to *P.*
281 *capitalensis*, which is a more distantly related member of this genus (Fig. 1A), we observed
282 an increased frequency of chromosomal rearrangements (mainly intra-chromosomal
283 inversions with only few inter-chromosomal translocations), suggesting that these species are
284 mesosyntenic (Fig. 1B). As expected, when moving (phylogenetically) further away from *P.*
285 *citrichinaensis* and comparing its genome sequence to the one of *Lophium mytilinum*, a
286 relatively obscure fungus from the order *Mytilinidiales* (*Dothideomycetes*) (Fig. 1A), a clear
287 pattern of mesosynteny can be observed (Fig. 1B). When comparing *P. citrichinaensis* to
288 *Neurospora crassa*, a species from the class *Sordariomycetes*, we only observed weak
289 mesosynteny with the pattern apparently dissipating (Suppl. Fig. S1). Thus, our results clearly
290 demonstrate gradually changing patterns of mesosynteny across progressively more distantly
291 related species within the *Dothideomycetes*, but possibly also outside of this taxonomic class.
292 Furthermore, we show that species within the genus *Phyllosticta* spur the mesosyntenic mode
293 of evolution that has previously been described for other *Dothideomycetes* (Hane *et al.* 2011;
294 Ohm *et al.* 2012).
295

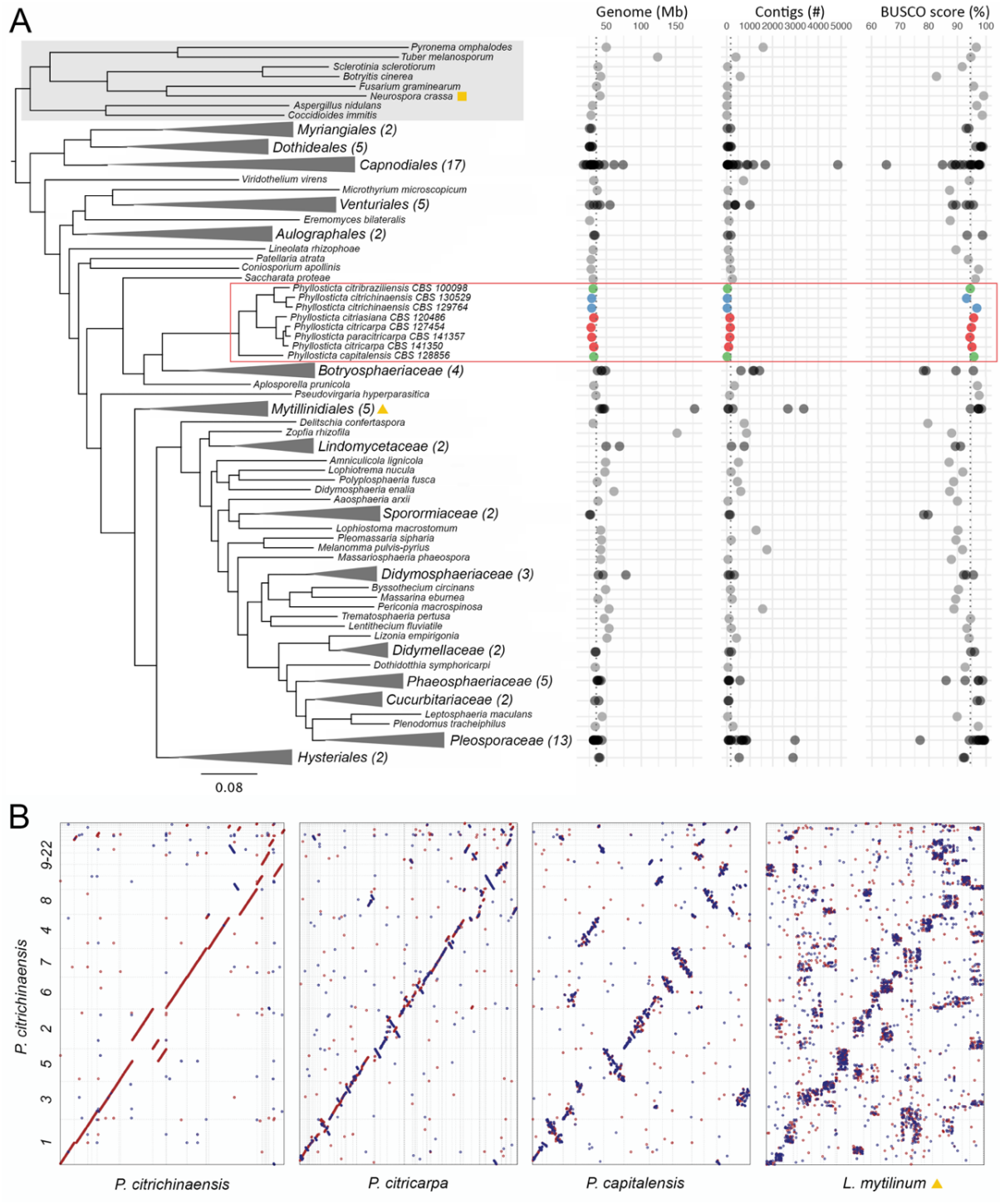



**Figure 1. *Phyllosticta* genomes are of high quality and show mesosynteny. A.** The phylogenetic relationship of more than 100 fungal species (116 genomes) is shown. The phylogenetic tree was generated using OrthoFinder, and sub-trees were collapsed manually to enhance readability. Grey box indicates species outside the *Dothideomycetes*, red rectangle indicates the genus *Phyllosticta*. Red orbs = *Phyllosticta* pathogens, green orbs = *Phyllosticta* endophytes, blue orbs = *P. citrichinaensis.* Yellow triangle and square indicate the locations of *Lophium mytilinum* and *Neurospora crassa*, respectively. **B.** Whole-genome alignments between *P. citrichinaensis* and *P. citricarpa, P. capitalensis*, and *L. mytilinum*, respectively, were generated using PROmer; the red line/dot = homologous area and the blue line/dot = reversed homologous area.

### 3.3 – *Phyllosticta* differ in gene number and functional annotation in a lifestyle dependent manner

Species of similar lifestyles often share (groups of) genes (Lo Presti *et al.* 2015; Kim *et al.* 2016). Consequently, we hypothesized that the presence or absence of specific genes may provide information about the (predominant) lifestyle of *P. citrichinaensis*. To be able to compare gene content over different species and strains, we used OrthoFinder (Emms and Kelly 2019) to identify ortholog groups (OGs) across all 116 predicted proteomes. Orthofinder identified 35,379 OGs containing 88.1% of all genes (Suppl. Table 2A). The 11.9% of genes that were not assigned to any OG likely constitute species-specific genes, which is to be expected given the taxonomically diverse set of fungal species considered in our study (Fig. 1A). In total, 1,794 OGs (5.1%) contained genes from all species, representing a fungal core genome. Of all OGs, 32.2% (11,352) contained at least one gene from a *Phyllosticta* species, and of those, 57.8% (6,558, Fig. 2) contained genes from all *Phyllosticta* species and 33.2% (3,764) were unique to *Phyllosticta* (i.e., they only contained *Phyllosticta* genes). The latter percentage is quite high because of the close taxonomic relation of some of the genomes: a rather large fraction of the *Phyllosticta* unique OGs contain two or three genes (2,734, 72.6%), as these often contain one gene from each of the two *P. citricarpa* genomes and one more gene from *P. paracitricarpa*. Genes that are unique to a species, i.e., sequences that are sufficiently different from other sequences, do not form a separate OG on their own and consequently are not considered in these statistics. Since the *P. citricarpa* and *P. paracitricarpa* genomes are so closely related, many of their “unique genes” are assigned to an OG, which causes the fraction of *Phyllosticta* unique OGs to be quite large.

To discover lifestyle-associated genes in *Phyllosticta*, we compared OG content across the eight *Phyllosticta* species and sought to identify differences between species of different lifestyles. The number of OGs that is shared by *Phyllosticta* species of a specific lifestyle (45–222, Fig. 2 and Suppl. Table 2B) is much smaller than the total number of OGs that all species in this genus share (6,558). The OGs shared by *Phyllosticta* spp. of a specific

lifestyle might also contain genes from species outside of this genus. *Phyllosticta* pathogens shared 222 OGs that are not present in *Phyllosticta* endophytes, a much higher number than the total of 45 OGs that are shared by endophytes and not present in pathogens, which is likely due to the larger phylogenetic distance between the two endophytes compared with the pathogenic *Phyllosticta* species (Fig. 1A). In addition to OGs that are unique to species of a specific lifestyle, OGs that are present in all species but differ in their abundance in species that share a lifestyle (e.g., there are more genes in species of one lifestyle compared to the others) may provide information on how species adapt to their lifestyles. We thus identified OGs for which the average number of genes in species of one lifestyle was higher than the average number of genes for species of the other lifestyle; we did not consider OGs that contained large outliers, i.e., one species having a much larger number of genes as compared to the other species. This resulted in a total of 87 OGs: 73 OGs that had more genes in endophytes and 14 OGs that had more genes in pathogens (Suppl. Table 2C), suggesting that achieving the endophytic lifestyle requires additional genes.

As functions of genes in OGs that are unique to, or enriched in, species of a certain lifestyle may underly lifestyles adaptations, and this may help to uncover which lifestyle *P. citrichinaensis* has, we looked into the annotations of the lifestyle-related OGs. In addition to the annotations of *Phyllosticta* genes, we also included the annotations of 50 out of the 108 species outside the genus *Phyllosticta* for which annotation data was available on the JGI database. The number of OGs in which at least one gene was functionally annotated (other than as hypothetical or expressed protein) varied widely between different lifestyle-related groups: while nearly 50% of the OGs that had more genes in endophytes received an annotation, less than 3% of the pathogen-only group did (Suppl. Table 3A–F). We further divided the individual functional annotations into KOG-classes (from the EuKaryotic Orthologous Groups tool, https://mycocosm.jgi.doe.gov/help/kogbrowser.jsf), which provide a high-level classification system to group genes with comparable activities. Of all KOG-classes,

363 the class 'Secondary metabolites biosynthesis, transport and catabolism' was most often
364 found to be associated with lifestyle-related OGs: only two of six lifestyle-related groups did
365 not contain genes in this class, suggesting this group of genes may be useful to distinguish
366 species of different lifestyles (Suppl. Table 3I).
367
368 To study secondary metabolite biosynthesis genes in more detail we used antiSMASH
369 (Blin *et al.* 2019) to identify biosynthetic gene clusters (BGCs) in the eight *Phyllosticta*
370 genomes. The total number of predicted BGCs varied from 20 in *P. paracitricarpa* to 24 in *P.*
371 *citricarpa*, with no apparent differences between species of different lifestyles. Interestingly,
372 one of the terpene clusters was predicted as a squalestatin in all pathogenic species, while it
373 received no functional prediction in the endophytic species or in *P. citrichinaensis*, suggesting
374 there is a difference in this cluster between species of different lifestyles. Squalestatins are
375 predicted to be inhibitors of squalene synthase, which produces squalene, a sterol
376 biosynthetic intermediate that is reported to play a role in mediating interactions between fungi
377 and their plant hosts (Lindo *et al.* 2020). Therefore, further characterization of this BGC and
378 others in *Phyllosticta* will be worthwhile for future studies into pathogenicity of *Phyllosticta*
379 species.
380
381 Ortholog groups with more genes in *Phyllosticta* endophytes were more often
382 functionally annotated, suggesting that these are generally better characterized and likely
383 evolutionary conserved. We did not find any particular KOG-class to be annotated in higher
384 abundance in endophytes, but nonetheless found a few interesting annotated OGs, such as
385 six OGs that were annotated to belong to the 'carbohydrate transport and metabolism' class
386 including a CAZyme family (GH55) gene and several transporters, suggesting a role for
387 carbohydrate transport in lifestyle (Suppl. Table 3A–F, I). One endophyte-only OG contained
388 the MAT1-1 gene, which was the result of all sequenced pathogenic strains having MAT1-2
389 mating types (Petters-Vandresen *et al.* 2020). In addition, although the *P. citrichinaensis*
390 MAT1-1 gene is not present in this OG, we did find the MAT1-1 gene in the *P. citrichinaensis*

genome assembly and found it to be highly similar to that of *P. citribraziliensis*, its closest relative.

The pathogen-only group was poorly functionally annotated; out of 222 OGs, only five received a functional annotation. The fact that such a large fraction of OGs could not be assigned a functional annotation is of interest as some of the non-annotated OGs in the pathogen-only group could contain putative effectors or genes that are otherwise involved in virulence, as effectors often remain unannotated in standard annotation pipelines (Sperschneider *et al.* 2015). In addition, two OGs received functional annotations that have previously been implied to be virulence factors and could therefore be interesting targets for future functional studies; a pectin lyase fold (Yang *et al.* 2018), and a cytochrome p450 (Siewers *et al.* 2005; Shin *et al.* 2017).

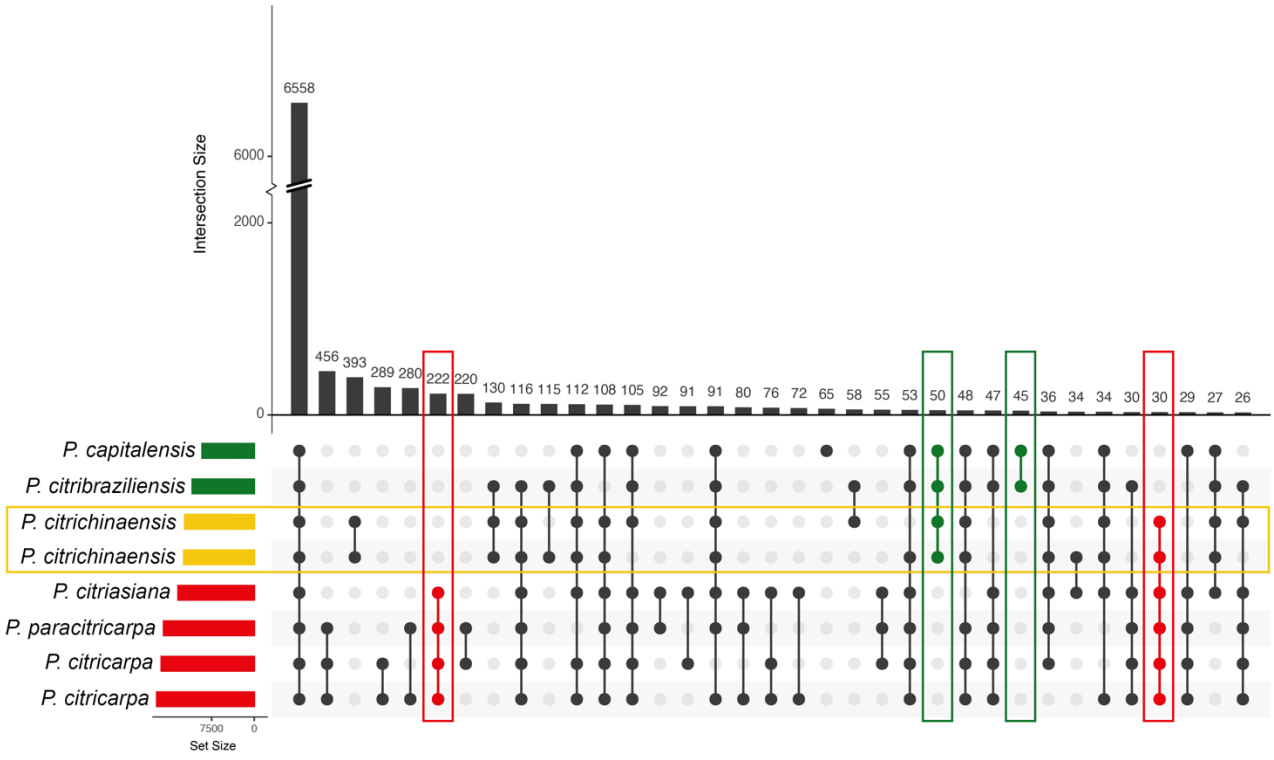



**Figure 2. *Phyllosticta citrichinaensis* shares more unique OGs with endophytes than with pathogens, but pathogens share more OGs with each other.** Yellow rectangles = *P. citrichinaensis*, red rectangles = pathogens, green rectangles = endophytes. Figure generated using UpSetR.

### 3.4 – *Phyllosticta citrichinaensis* shares more lifestyle-specific OGs with endophytes, but follows an intermediate pattern in other lifestyle-associated OGs

The lifestyle of *P. citrichinaensis* is currently ambiguous (Wang *et al.* 2012), but the number of OGs that it shares with species of either lifestyle may provide clarity not only about the lifestyle of *P. citrichinaensis* itself but also about the differences between species of different lifestyles within this genus. The two *P. citrichinaensis* strains share more lifestyle specific OGs with endophytes (50) than they do with pathogens (30). In addition, the number of OGs that is shared by only the two endophytes and *P. citrichinaensis* is larger than the number of endophyte-only OGs that are not shared with *P. citrichinaensis* (Fig. 2, green rectangles), suggesting that *P. citrichinaensis* indeed compared well to endophytic species. However, for 30 out of 87 OGs that contained more genes in either pathogenic or endophytic species, the *P. citrichinaensis* gene numbers corresponded best to the endophytic numbers, while in 32 OGs, they corresponded to the pathogenic numbers. In 25 OGs, the number of genes of *P. citrichinaensis* corresponded to neither lifestyle (Suppl. Table 2C). As opposed to the numbers of OGs specific to species of one lifestyle, where *P. citrichinaensis* shared more with endophytes (Fig. 2), the numbers of *P. citrichinaensis* genes in OGs that contained more genes in either pathogenic or endophytic species thus indicate that it compares equally well to species of either lifestyle. Thus, this data suggests that presence/absence and/or gene abundance differences are not sufficient to provide insights into the lifestyle of *P. citrichinaensis.*

While we did not observe clear patterns in the types of functions that *P. citrichinaensis* shares with either endophytes or pathogens, we nevertheless could observe some interesting functional patterns in *P. citrichinaensis* (Suppl. Table 3G–H). For instance, two groups were annotated as heat shock proteins: an Hsp40 (DNAJC17) that had more genes in pathogens (3 in pathogens vs 1–2 in endophytes) as well as an Hsp70 that had more genes in endophytes (4–5 in pathogens vs 6–7 in endophytes). *Phyllosticta citrichinaensis* contains fewer genes in both groups, suggesting it may respond differently to stress. Furthermore, one group that had

more genes in pathogens as well as in *P. citrichinaensis* (1 in endophytes versus 2–3 in pathogens) contains genes annotated as peroxiredoxin-1 or peroxiredoxin-6. Peroxiredoxins are necessary for full virulence in several fungal pathogens such as *Magnaporthe oryzae* and *Aspergillus fumigatus* as they offer an antioxidant defense against reactive oxygen species produced by the host as part of host defense responses (Mir *et al.* 2015; Rocha *et al.* 2018). It is thus possible that these additional genes in *Phyllosticta* pathogens and in *P. citrichinaensis* contribute to their virulence.

### 3.5 – There is little difference between *Phyllosticta* endophytes and pathogens in the numbers of putative secreted proteins and putative effectors

Secreted proteins, including effectors, play an important role in lifestyle and virulence (de Wit *et al.* 2009; Lo Presti *et al.* 2015; Plissonneau *et al.* 2017). Based on the comparison of OGs in species of different lifestyles, we concluded that pathogenic species contain a large number of unannotated genes in pathogen-specific OGs (217 OGs). We hypothesized that some of these genes may be effectors or other secreted proteins, and that the presence of putative secreted proteins and effectors in the genomes of species may be an indicator for lifestyle differences. We therefore assessed the presence of OGs that contain 50% or more putative secreted or effector proteins in all 116 species used in this study. A total of 3,942 OGs (11% of 35,379 OGs) consisted of at least 50% secreted proteins (Fig. 3A, Suppl. Table 4A) and of these, 1,048 OGs consisted of at least 50% effectors (Fig. 3B, Suppl. Table 4B). Notably, only 62 OGs containing putatively secreted genes were present in all 116 species, and only one OG containing putative effectors contained genes from all 116 species, corroborating that secreted proteins and effectors are typically not shared between different species and especially effectors are rather species and/or strain specific (Stergiopoulos *et al.* 2012).

*Phyllosticta* genes were present in a total of 762 putatively secreted OGs, with 432 of those containing genes from all *Phyllosticta* species included in this study. With an average of 786 genes per species, *Phyllosticta* species contain less genes in OGs encoding secreted

proteins compared with other *Dothideomycetes* (an average of 1,033 genes, Suppl. Table 4). In addition, *Phyllosticta* species contained on average 110 putatively secreted genes that were not in an OG or that were the only protein predicted to be secreted in an OG (singletons), which is also less than we observed for the overall average of 128 singleton secreted genes (Suppl. Table 4). When assessing the total number of predicted secreted genes (both in OGs and singletons), *Phyllosticta* endophytes have more putative secreted proteins (average of 913) as compared to *Phyllosticta* pathogens (average of 897). However, when considering these as a percentage of the total predicted proteome size, this difference becomes negligible, with endophytes having a slightly smaller percentage of proteins that is secreted (7.89%) as compared to pathogens (8.01%, Suppl. Table 4C). We found 247 putatively secreted OGs to be unique to *Phyllosticta* (Fig. 3A), which mainly represent genes encoding proteins without predicted functions such as those often found to be putative effectors (Suppl. Table 4G, Lo Presti *et al.* 2015). Namely, only four out of 247 OGs contained genes that were functionally annotated, and interestingly, three out of four were predicted to function in carbon utilization (Suppl. Table 4G).

*Phyllosticta* species contain less genes in OGs encoding effectors with an average of 74 genes compared to the overall average of 93 genes. In addition, *Phyllosticta* species contained on average 30 putative effector genes that were not in an OG or that were the only effector in an OG (singletons), which is lower than the average of 44 singleton effector genes for the other *Dothideomycetes*. Comparable to secreted proteins, when assessing the total number of predicted effector genes (both in OGs and singletons), *Phyllosticta* endophytes appear to have slightly more putative effector genes (average of 109) as compared to *Phyllosticta* pathogens (average of 106.5), but when taken as percentage of the total number of predicted proteomes per species, this difference is negligible (0.95% vs 0.94%) (Fig 3B–D, Suppl. Table 4D). A total of 30 effector OGs contained genes from all *Phyllosticta* species while also containing genes from other species. One of these had higher gene numbers in endophytes, but received no annotation, while none had higher gene numbers in pathogens.

493 Effector genes are often hypothesized to be species and/or strain specific (Lo Presti *et al.*
494 2015; Sperschneider *et al.* 2015). We identified in total 67 OGs containing effectors that are
495 unique to *Phyllosticta*, seven of which were present in all pathogens but not in endophytes,
496 two were present in all endophytes but not in pathogens, and one had higher gene numbers
497 in endophytes (Suppl. Table 4E). Genes of the seven pathogen-only OGs were not colocalized
498 as described in some fungal phytopathogens (Ma *et al.* 2010; Dong *et al.* 2015) but rather are
499 spread across separate scaffolds. None of these unique effector genes were functionally
500 annotated, suggesting that these have yet undescribed functions.
501
502 The occurrence of *P. citrichinaensis* effector genes in lifestyle-associated OGs could
503 provide further evidence for its lifestyle. One of the seven effector OGs that only occurred in
504 pathogens contained a gene from one of the *P. citrichinaensis* strains. In contrast, both
505 endophyte-only effector OGs contained genes from *P. citrichinaensis*, in one case only from
506 one strain, and in the other case from both strains. In the effector OG that had more genes in
507 endophytes, *P. citrichinaensis* followed an intermediate pattern: one strain contained the same
508 number of genes as pathogens, while the other contained the same number as endophytes.
509 These data thus suggest that *P. citrichinaensis* follows an intermediate lifestyle.
510

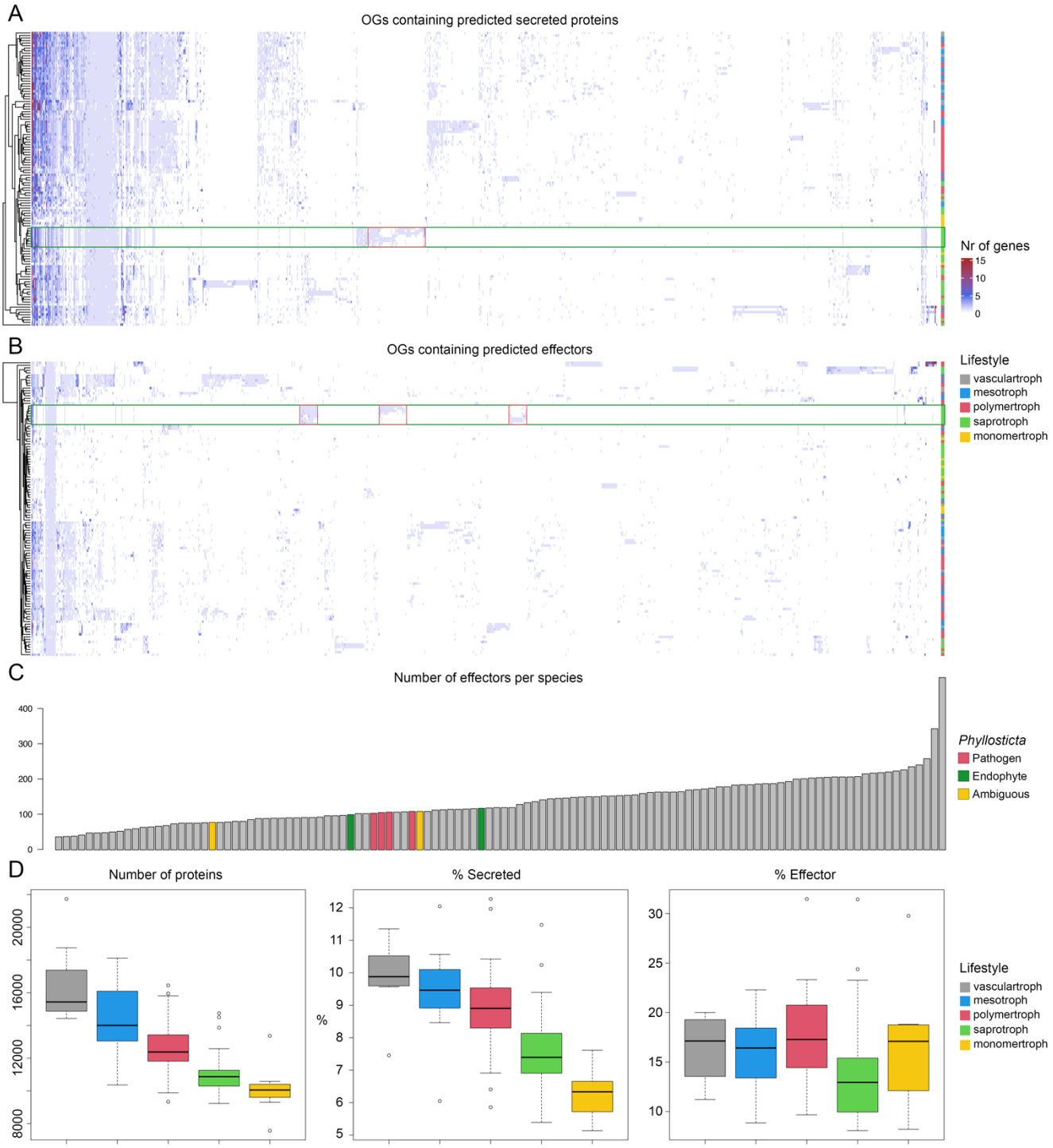



**Figure 3. Species cluster independent of their taxonomic relationship based on presence of secreted or effector proteins.** Clustered heatmaps of the number of genes in OGs that were predicted to contain **A.** secreted proteins and **B.** effectors. Lifestyles as predicted by CATAstrophy (Fig. 4). *Phyllosticta* species are highlighted with green rectangles, *Phyllosticta*-unique OGs are highlighted in red rectangles. **C.** Number of effectors per species (both those in OGs and singletons). *Phyllosticta* species are highlighted in red (pathogens), green (endophytes), and yellow (*P. citrichinaensis*). **D.** The total number of proteins, the percentage of proteins that is predicted to be secreted, and the percentage of secreted proteins that is predicted to be an effector, in species of different lifestyles (as predicted by CATAstrophy, Fig. 4).

### 3.6 – *Phyllosticta citrichinaensis* clusters with pathogens based on carbon growth data

CAZymes enable fungi to utilize different carbon sources (van den Brink and de Vries 2011; Lombard *et al.* 2014), and are thought to be involved in fungal pathogenicity (ten Have *et al.* 2002; King *et al.* 2011; Kubicek *et al.* 2014; Hane *et al.* 2020). We have previously shown that carbon utilization capabilities differ between *Phyllosticta* spp. and uncovered a clear distinction in the ability of pathogens and endophytes to grow in the presence of sugar beet pulp; while the growth of endophytes was unchanged, pathogens were strongly inhibited (Buijs *et al.* 2021). To assess if *P. citrichinaensis* displays similar growth behavior to pathogens or endophytes, we grew *P. citrichinaensis* on 35 different carbon sources including sugar beet pulp (Fig. 4). Interestingly, growth of *P. citrichinaensis* is not inhibited by the presence of sugar beet pulp (Fig. 4A), suggesting that *P. citrichinaensis* behaves comparable to endophytic *Phyllosticta* spp. To further substantiate this observation, we performed hierarchical clustering of seven *Phyllosticta* strains based on their growth on all 35 carbon sources. Unanticipatedly, *P. citrichinaensis* clustered together with the pathogenic species rather than with the endophytes (Fig. 4B). Thus, although *P. citrichinaensis* is not inhibited by sugar beet pulp, it generally displays carbon utilization capabilities comparable with pathogens.

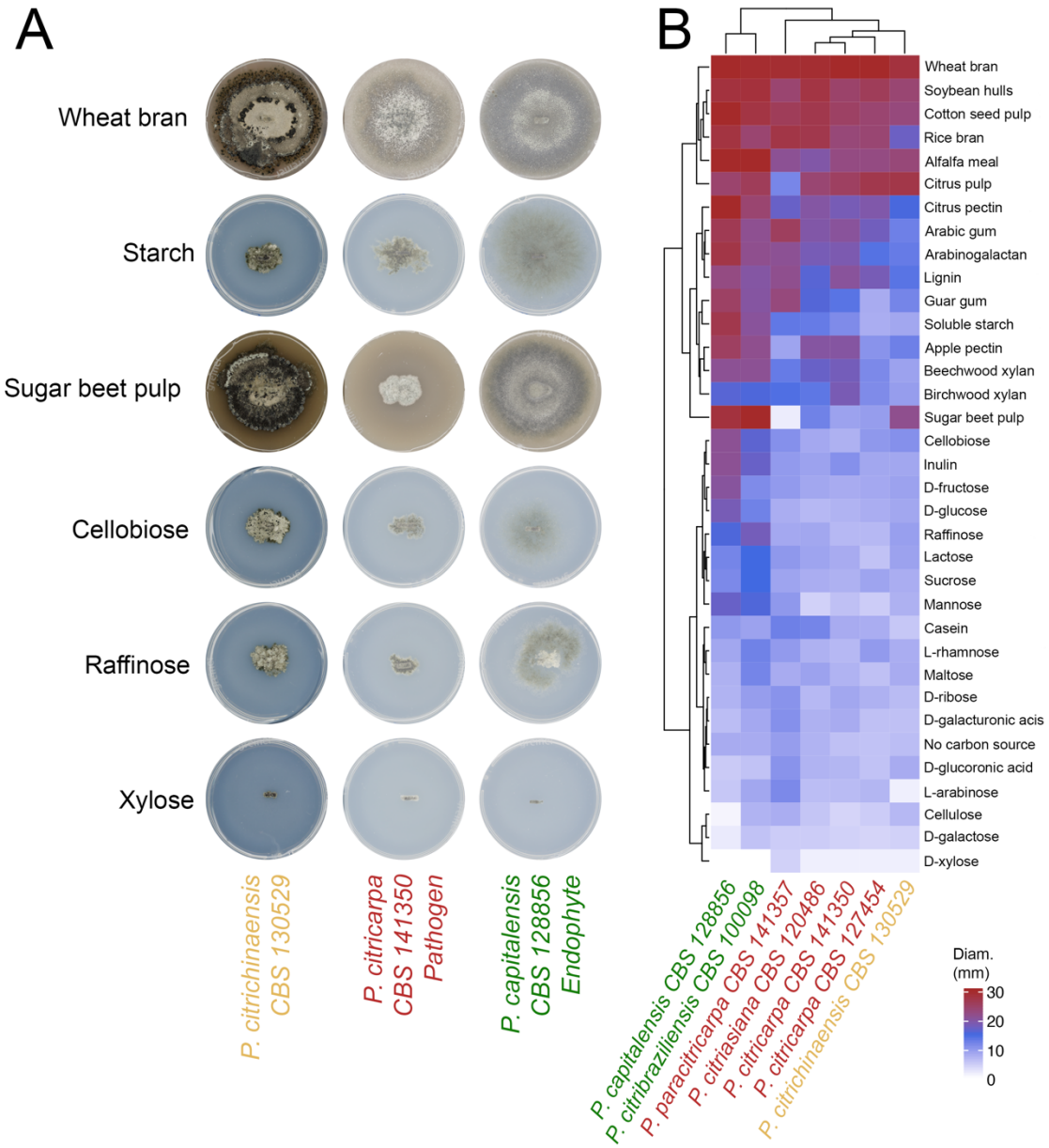
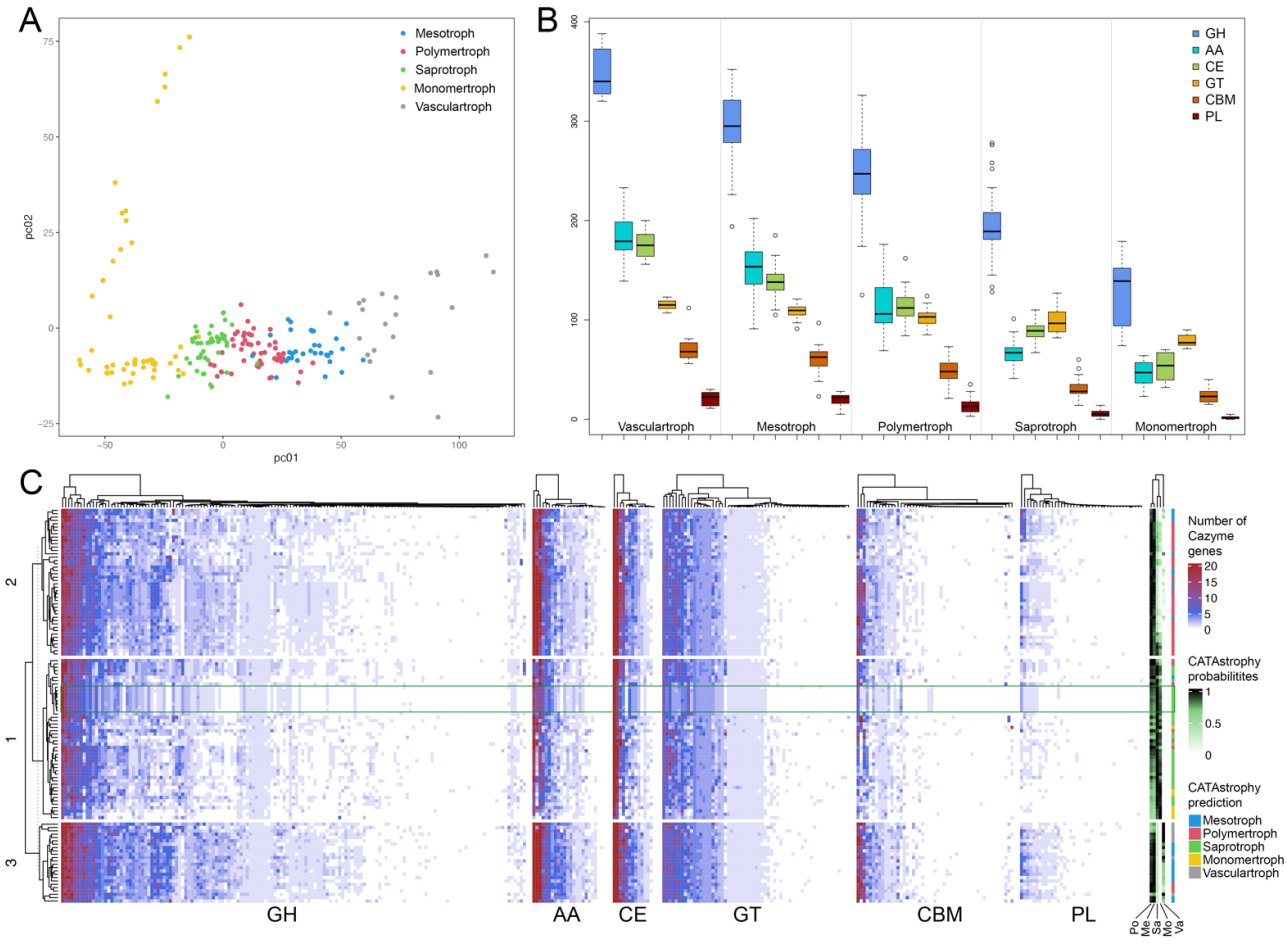



**Figure 4. *Phyllosticta citrichinaensis* clusters with pathogens based on growth on 35 carbon sources, but behaves like an endophyte in the presence of sugar beet pulp. A.** Images of *Phyllosticta* species growing on a selection of different carbon sources. **B.** Clustering of *Phyllosticta* species based on their growth on different carbon sources. All species were grown on 35 different carbon sources, colony diameters were measured, and images takes on all sources when the biggest colony of a species reached the edge of its plate. All species grew fastest on wheat bran.

548 **3.7 – Genomes of *Dothideomycetes* can be clearly distinguished based on CAZyme**
549 **content, but this does not correlate well with lifestyles described in literature**

550 The genetic basis for the ability to utilize different carbon sources is often caused by
551 differences in CAZymes repertoires (van den Brink and de Vries 2011; Lombard *et al.* 2014).
552 Interestingly, the abundance and diversity of CAZymes encoded in a genome is also related
553 to lifestyle and consequently enables to predict the tropic classification of a species (Lo Presti
554 *et al.* 2015; Hane *et al.* 2020). To determine the trophic classification for *P. citrichinaensis*, we
555 used CATAstrophy to annotate CAZyme genes for all 116 predicted proteomes and to perform
556 a principal component analysis (PCA) to distinguish species with different trophic classes
557 based on their CAZyme repertoire. CATAstrophy clearly separated species of different trophic
558 classes based on the first principal component (PC1) (Fig. 5A and Suppl. Table 5). The second
559 principal component (PC2) mainly separated species based on phylogeny; as previously
560 observed for fungi and oomycetes (Hane *et al.* 2020), oomycetes clustered together separated
561 along the PC2 axis from fungi. The different trophic classes differ considerably in the numbers
562 of genes per CAZyme family (Fig. 5B). For example, GH families differ clearly between trophy
563 class, and GH are also the CAZyme families with the highest number of genes per species.
564 In contrast, almost no difference can be observed in the PL family, which generally has the
565 smallest number of genes per species. The CATAstrophy gene predictions were used to
566 identify CAZyme-containing OGs, for which we then obtained the number of genes present in
567 each species to generate a clustered heatmap (Fig. 5C). We observed three distinct clusters
568 that differ in their CAZyme repertoires, which typically correlate well with the CATAstrophy
569 trophy predictions (Fig. 5C and Suppl. Table 5). The CATAstrophy trophy predictions also
570 correlate well with the numbers of secreted proteins and effectors, although we did not observe
571 such a strict separation into three clusters as was observed for CAZyme genes (Fig. 5C, Fig.
572 3).
573
574 All *Phyllosticta* species were predicted to be saprotrophs by CATAstrophy (Fig. 5C)
575 and clustered according to their phylogenetic relationship, suggesting that they are generally

576 very similar in terms of CAZymes (Suppl. Table 5). Consequently, *P. citrichinaensis* clustered
577 most closely to *P. citribraziliensis*, an endophyte, as this is its closest relative (Fig. 1A). We
578 nevertheless found six CAZyme families that show consistent differences in gene number
579 between pathogens and endophytes: AA1_3, AA3, CBM18, CBM67, GH3, and PL22 (Table
580 1). In all cases except AA3, *P. citrichinaensis* follows the endophytic pattern. Together, these
581 results indicate that in terms of the presence of CAZyme genes, the *P. citrichinaensis*
582 genomes compare best to those of endophytes.




583

584 **Figure 5. Separation of species into different trophy classes based on presence of**
585 **CAZyme genes in their genomes. A.** PCA plot of PC1 vs PC2. PC1 separates species of
586 different trophy classes, PC2 separates species on phylogeny. **B.** Number of CAZyme genes
587 in each CAZy family per CATAstrophy class. **C.** Clustered heatmap of the number of CAZyme-
588 gene-containing OGs per species. GH = Glycosyl hydrolase, AA = Auxiliary activity, CE =
589 Carbohydrate esterase, GT = Glycosyl transferase, CBM = Carbohydrate binding module, PL
590 = Pectin lyase. Po = Polymertroph, Me = Mesotroph, Sa = Saprotroph, Mo = Monomertroph,
591 Va = Vasculartroph.

## 4 – Discussion

Lifestyle adaptations are thought to be driven by differences in gene content, and especially CAZymes are assumed to be crucial (ten Have *et al.* 2002; King *et al.* 2011; Hane *et al.* 2020). Here we aimed to elucidate the genomic differences between endophytes and pathogens within the *Phyllosticta* genus occurring on citrus, and aimed to determine the lifestyle of *P. citrichinaensis*. Based on the results we uncovered several differences between species with different lifestyles. For instance, endophytes more frequently contain higher numbers of genes in OGs, and these genes are more often annotated than in pathogenic species. In addition, pathogenic species share more un-annotated lifestyle-specific OGs as compared to endophytic species. Furthermore, we show that species cluster independently of phylogeny based on the CAZyme content of their genomes, and this clustering correlated well with trophy prediction by CATAstrophy. The ambiguous species *P. citrichinaensis* showed characteristics that matched with endophytes in some cases, with pathogens in other cases, and sometimes it did not match with either lifestyle, suggesting it may exhibit an intermediate lifestyle not accounted for in the current definitions.

We previously observed that only four CAZyme families showed a consistent difference between endophytic and pathogenic species (Buijs *et al.* 2021), one of which contained AA1_3/CBM18 (seven in pathogens vs eight in endophytes), which were mixed in one orthologue group, and another one contained family CBM18 only. In this study, we found in total six CAZyme families with a consistent difference between endophytes and pathogens. These included two separate OGs that contained the AA1_3 and CBM18 CAZyme families, which both consistently contained more genes in endophytes compared to pathogens. CATAstrophy predicts on average about 90 CAZyme genes extra compared to our previous results (Buijs *et al.* 2021). The biggest difference can be found in the CE (carbohydrate esterase) family, where between 83–89 genes are predicted instead of 15–17. Most of these are in the CE10 family (47–53). The CE10 family is no longer listed as carbohydrate active enzyme by the cazy.org database (used by JGI) because most of the members of this family

act on non-carbohydrate substances (Lombard *et al.* 2014). If we manually remove these, CATAstrophy still predicts 20 extra genes in the CE family. In addition, there are some 15–20 extra genes predicted in the AA family, five to ten more in the GH family, and approximately five more in the GT family. Numbers in the PL family are practically identical. In the CBM family, CATAstrophy predicts about five to ten genes less. In the JGI annotation pipeline, many of the genes in the CBM family contained multiple domains and were therefore counted multiple times, which might have led to an overestimation. However, for most families it seems that CATAstrophy predicts more genes compared to JGI. As most of the CAZyme genes predicted by JGI are based on some experimental validation (cazy.org), it would be wise to experimentally validate the CAZyme genes that were predicted by CATAstrophy and not by JGI.

Genes related to secondary metabolite biosynthesis but not the number of biosynthetic gene clusters (BGCs) as predicted by antiSMASH differ between species. It is important to note here that BGC prediction by antiSMASH is based primarily on the presence of a gene to produce the metabolite 'backbone', such as a polyketide synthase (PKS) (Blin *et al.* 2019). Once such a gene is found, antiSMASH takes an area in the genome of up to 20 kb (depending on the type of backbone gene) on either side of the gene and checks for the presence of tailoring genes, which are then all automatically included in the BGC. This means that tailoring genes that are altered or inactive are still included in the predicted BGC, and that alterations in genes may not result in an altered BGC prediction. In contrast, although the cluster may look very similar, alterations in tailoring genes may lead to the production of a very different compound: a good example is the synthesis of the toxins dothistromin, aflatoxin, and sterigmatocystin, which are all synthesized in a very similar manner, with only the very last tailoring steps being different (Schwelm and Bradshaw, 2010; for a review see Hüttel and Müller, 2021). We therefore conclude that although antiSMASH did not detect differences in BGC numbers, alterations in biosynthetic genes may be responsible for the differences

between species of different lifestyles in the genus *Phyllosticta*, and this will be an interesting subject for future studies.

CATAstrophy was able to clearly separate species based on the number of CAZyme genes present in the genomes and was able to separate species into different trophy predictions. Apart from closely related species that cluster together, a phylogenetic pattern cannot be observed in the clustering of the heatmap or in the trophy predictions, which suggests that there is a strong signal that links genome content to lifestyle. CATAstrophy also allows for trophy-overlap for species that are bordering two trophies, such as for *Sclerotinia sclerotiorum*, which received very high scores for both the polymertroph and the saprotroph class (Suppl. Fig. 5). This is consistent with literature, as *S. sclerotiorum* has been described to exhibit a necrotrophic phase that is followed by a saprotrophic phase (Hegedus and Rimmer 2005). *Sclerotinia sclerotiorum* is a much-researched model organism, and the descriptions in literature of its lifestyle are therefore well-developed. However, for many fungal species, this is not the case, and circumscriptions in literature are often limited or conflicting. Indeed, we see that the trophy predictions by CATAstrophy do not always correlate well with lifestyles described (or supposed) in literature: each trophy class includes species that are described as pathogens, endophytes, symbionts, or saprotrophs, or which have been described to exhibit multiple of these lifestyles. An underlying cause for this fact is that lifestyles that are described in literature may be inaccurate as the border between species of different lifestyles such as necrotrophs, hemibiotrophs, or biotrophs is not very strict, or need very specific conditions to manifest. For instance, *Phytophthora infestans* has been placed in all three classes (Oliver and Ipcho 2004). Similarly, *Botrytis cinerea* has been placed in different classes as the symptoms it causes differ widely in their severity, depending on the exact interaction with the host (Veloso and Van Kan 2018). Within the genus *Phyllosticta*, some obscurity with respect to lifestyle is present for instance for *P. capitalensis*, which is a widespread endophyte of many hosts including citrus (Wikee *et al.* 2013a), but may cause disease in other hosts such as guava (Arafat 2018). In addition, non-pathogens can evolve a

675 pathogenic lifestyle, and vice-versa, as can be observed with pathogenicity on pea
676 by *Neocosmospora solani*, which is dependent on the presence of only a few genes, or with
677 pathogenicity of *Fusarium oxysporum* on cucurbit species, which is determined by the
678 absence or presence of a mobile pathogenicity chromosome (Temporini and Vanetten 2002;
679 Dong *et al.* 2015; van Dam *et al.* 2017; Möller and Stukenbrock 2017). The possibility for
680 species to be categorized into multiple trophies in the CAZyme-based classification system
681 therefore presents an advantage over the traditional classification system as it allows for a
682 more correct, double classification of species that exhibit multiple lifestyles depending on the
683 host and other environmental parameters.
684
685 We compared genomes of species with different lifestyles within the genus *Phyllosticta*
686 specifically, and found several distinctions. We observed that endophytes more often have
687 higher numbers of genes in specific OGs as compared to pathogens. This suggests that the
688 ability to be an endophyte necessitates the presence of additional genes. The ancestral
689 *Dothideomycete* was likely a saprotroph, however, the most common ancestor of the
690 *Botryosphaeriales*, the order in which *Phyllosticta* resides, was probably a plant pathogen, as
691 determined by ancestral state reconstruction (Abdollahzadeh *et al.* 2020; Haridas *et al.* 2020).
692 The evolution of *Phyllosticta* endophytes from *Phyllosticta* pathogens through a gain of genes
693 and thereby gain of abilities is therefore a plausible scenario. With respect to lifestyle definition
694 in *Phyllosticta*, *P. citrichinaensis* is the most ambiguous citrus-related species within the
695 genus. By comparing the genome of this species with those of other species with different
696 lifestyles within the genus *Phyllosticta*, we aimed to elucidate the lifestyle of this species. In
697 several aspects, we found *P. citrichinaensis* to be most similar to endophytes. For instance,
698 *P. citrichinaensis* shares more OGs specific to species of one lifestyle with endophytic species
699 (50) than it does with pathogenic species (30), and none of its biosynthetic gene clusters was
700 predicted to produce a squalestatin, which was also the case for all of the endophytes, but not
701 for the pathogens. In addition, CAZyme families that had more genes in endophytes than in
702 pathogens, often also had more genes in *P. citrichinaensis*. Furthermore, growth of *P.*

703 *citrichinaensis* was not inhibited by the presence of sugar beet pulp, similarly to endophytic
704 species. In contrast, its broader carbon utilization capabilities were more comparable to those
705 of pathogenic species. Another aspect in which *P. citrichinaensis* was comparable to
706 pathogens, was the presence of more putative peroxiredoxin genes in pathogens and *P.*
707 *citrichinaensis* as compared to endophytes. On other aspects, *P. citrichinaensis* did not
708 compare well with species of either lifestyle, such as the number of OGs that had more genes
709 in species of either lifestyle: in almost a third of the cases, *P. citrichinaensis* did not match the
710 gene numbers of either lifestyle. The number of effector genes that *P. citrichinaensis* shared
711 with species of either lifestyle also suggests an intermediate pattern. The lifestyle of *P.*
712 *citrichinaensis* cannot be univocally determined without performing pathogenicity assays, but
713 as these are currently not available for this species, the data presented here give a good
714 estimation that shows that *P. citrichinaensis* is an intermediate taxon, not perfectly fitting into
715 any of the currently defined lifestyle definitions.
716
717 Research performed in recent years has shown with increasing confidence that
718 borders between lifestyles simply are not very strict and in fact are subject to constant change.
719 Examples such as *B. cinerea* and *Phytophthora infestans*, which both have been placed in
720 multiple lifestyle classes depending on the host and other environmental parameters,
721 demonstrate that our current classification systems are not always adequate to separate
722 species into different lifestyles (Oliver and Ipcho 2004; Veloso and Van Kan 2018). In addition,
723 the ability of species to be pathogenic to a specific host may change with the gain of only a
724 few genes (Temporini and Vanetten 2002; van Dam *et al.* 2017). Classifying plant-associated
725 microbes into different lifestyles is an important area of research as it allows for the
726 identification of genomic parameters that are required for pathogenicity, and therefore aids in
727 the search of a remedy against such pathogens. However, a classification system is only
728 valuable if it allows for the accurate separation of species; an incorrectly classified organism
729 may lead to incorrect conclusions and could cause much confusion. A classification such as
730 the one proposed by Hane *et al*., which is based on the number of CAZyme genes in a species'

731 genome, is a significant improvement since it allows for overlap between lifestyles. Further
732 development of such classifications for instance by the addition of other genomic parameters
733 such as the presence of effectors could lead to the development of a more accurate and useful
734 classification system in the future.
735
736 **Data availability statement**
737 The whole genome sequencing data including annotations for the newly sequenced
738 *Phyllosticta citrichinaensis* genome are publicly available at the JGI genome portal:
739 https://mycocosm.jgi.doe.gov/Pcit129764. The authors affirm that all other data necessary for
740 confirming the conclusions of the article are present within the article, figures, and
741 (supplementary) tables.
742
743 **Acknowledgements**
744 This work was funded by the Dutch Applied Science division (TTW) of NWO and the
745 Technology Program of the Ministry of Infrastructure and Water Management under project
746 15807 of the Research Programme I&W Biotechnology and Safety. The work conducted by
747 the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility,
748 is supported by the Office of Science of the U.S. Department of Energy under Contract No.
749 DE-AC02-05CH11231.
750

751 **Tables**

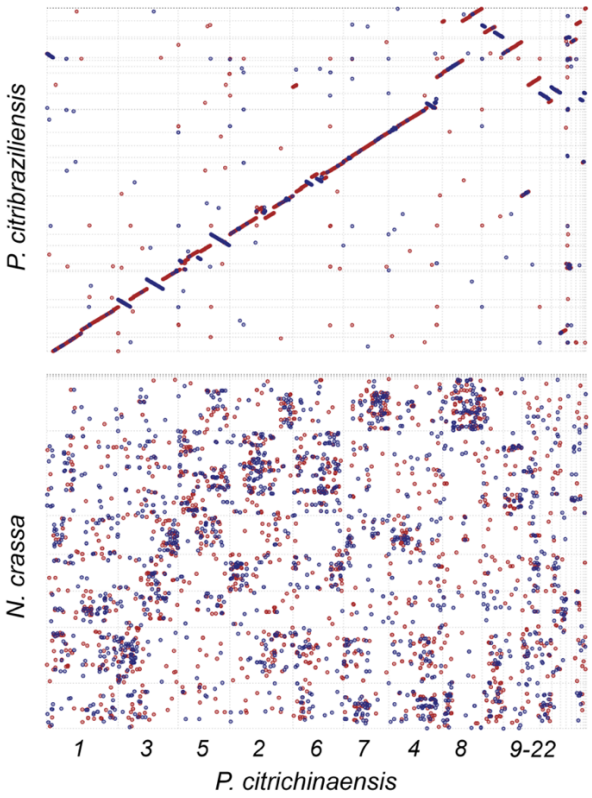
752

753 **Table 1. CAZyme families with gene abundance differences between endophytes and**
754 **pathogens. Columns with a darker fill color indicate the family which is more abundant**
755 **in pathogens, while the lighter color indicates those which are more abundant in**
756 **endophytes.**

| Species name | Lifestyle | CBM67 | AA1_3 | CBM18 | GH3 | PL22 | AA3 |
|---|---|---|---|---|---|---|---|
| *Phyllosticta capitalensis* | E | 0 | 10 | 10 | 15 | 1 | 2 |
| *Phyllosticta citribraziliensis* | E | 0 | 10 | 9 | 15 | 1 | 2 |
| *Phyllosticta citrichinaensis* | ? | 0 | 10 | 9 | 16 | 1 | 1 |
| *Phyllosticta citrichinaensis* | ? | 0 | 10 | 9 | 15 | 1 | 1 |
| *Phyllosticta citriasiana* | P | 1 | 9 | 7 | 13 | 0 | 1 |
| *Phyllosticta citricarpa* (CPC 27913) | P | 1 | 9 | 6 | 14 | 0 | 1 |
| *Phyllosticta citricarpa* (CBS 127454) | P | 1 | 9 | 7 | 14 | 0 | 1 |
| *Phyllosticta paracitricarpa* | P | 1 | 9 | 7 | 14 | 0 | 1 |

757

## Supplementary Figures




**Supplementary Fig. S1.** Whole-genome alignments between *P. citrichinaensis* and *P. citribraziliensis* (top panel) and *Neurospora crassa* (lower panel) were generated using PROmer. The red line/dot = homologous area and the blue line/dot = reversed homologous area.

## Supplementary Tables (submitted as separate excel files)

**Supplementary Table 1. Genome data**, including QUAST and BUSCO results, of all species used in this study.

**Supplementary Table 2. Orthofinder results. A**. Orthofinder statistical data. **B**. List of ortholog groups (OGs) that are unique in species of one lifestyle. **C**. List of OGs that have more genes in species of one lifestyle.

**Supplementary Table 3. Annotation data of OGs. A.** Annotation data of OGs present only in endophytes. **B**. Annotation data of OGs present in endophytes and *P. citrichinaensis* (PCC). **C**. Annotation data of OGs present only in pathogens. **D**. Annotation data of OGs present in pathogens and *P. citrichinaensis* (PCC). **E**. Annotation data of OGs that consistently contain more genes in *Phyllosticta* endophytes as compared to *Phyllosticta* pathogens. **F**. Annotation data of OGs that consistently contain more genes in *Phyllosticta* pathogens as compared to *Phyllosticta* endophytes. **G**. Annotation data of OGs in which the gene number in *P. citrichinaensis* (PCC) matches the number of endophytes. **H**. Annotation data of OGs in which the gene number in *P. citrichinaensis* (PCC) matches the number of pathogens. **I**. The number of OGs in each lifestyle-related group that is classified to each KOG class. **J**. A list of annotation files from the JGI database that were included in these analyses.

**Supplementary Table 4. Data of OGs that contain putative secreted and effector proteins. A.** Secreted OGs in listed in order corresponding to Figure 3A. **B**. Effector OGs listed in order corresponding to Figure 3B. **C**. Number of secreted genes as estimated by: number of genes in secreted OG vs signalP prediction. **D**. Number of effector genes as estimated by number of genes in effector OG vs effectorP prediction. **E**. Effector OGs that are unique to species of *Phyllosticta*. **F**. Effector OGs that are present in all *Phyllosticta* spp. **G**. Functions of secreted OGs that are unique to *Phyllosticta*.

**Supplementary Table 5. CATAstrophy results.** Including all trophy predictions and all CAZyme gene counts.

## References

Abdollahzadeh, J., J. Z. Groenewald, M. P. A. Coetzee, M. J. Wingfield, and P. W. Crous, 2020 Evolution of lifestyles in Capnodiales. Stud. Mycol. 95: 381–414.

Almagro Armenteros, J. J., K. D. Tsirigos, C. K. Sønderby, T. N. Petersen, O. Winther *et al.*, 2019 SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. 2019 374 37: 420–423.

Arafat, K., 2018 A novel isolate of Phyllosticta capitalensis causes Black Spot Disease on Guava fruit in Egypt. Asian J. Plant Pathol. 12: 27–37.

Blin, K., S. Shaw, K. Steinke, R. Villebro, N. Ziemert *et al.*, 2019 antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 47: W81–W87.

van den Brink, J., and R. P. de Vries, 2011 Fungal enzyme sets for plant polysaccharide degradation. Appl. Microbiol. Biotechnol. 91: 1477–1492.

Buijs, V. A., X. C. L. Zuijdgeest, J. Z. Groenewald, P. W. Crous, and R. P. de Vries, 2021 Carbon utilization and growth-inhibition of citrus-colonizing Phyllosticta species. Fungal Biol. 125: 815–825.

Chiapello, H., L. Mallet, C. Gué Rin, G. Aguileta, J. Lle Amselem *et al.*, 2015 Deciphering genome content and evolutionary relationships of isolates from the fungus Magnaporthe oryzae attacking different host plants. Genome Biol. Evol. 7: 2896–2912.

Conway, J. R., A. Lex, and N. Gehlenborg, 2017 UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics 33: 2938–2940.

van Dam, P., L. Fokkens, Y. Ayukawa, M. van der Gragt, A. ter Horst *et al.*, 2017 A mobile pathogenicity chromosome in Fusarium oxysporum for infection of multiple cucurbit species. Sci. Rep. 7: 9042.

van Dam, P., and M. Rep, 2017 The distribution of Miniature Impala Elements and SIX genes in the Fusarium genus is suggestive of horizontal gene transfer. J. Mol. Evol. 85: 14–25.

Dong, S., S. Raffaele, and S. Kamoun, 2015 The two-speed genomes of filamentous pathogens: waltz with plants. Curr. Opin. Genet. Dev. 35: 57–65.

825 Emms, D. M., and S. Kelly, 2019 OrthoFinder: Phylogenetic orthology inference for
826 comparative genomics. Genome Biol. 20: 238.
827 European Food Safety Authority, 2014 Scientific Opinion on the risk of Phyllosticta citricarpa
828 (Guignardia citricarpa) for the EU territory with identification and evaluation of risk
829 reduction options: Wiley-Blackwell Publishing Ltd.
830 Eustáquio Lanza, F., T. Germano Metzker, T. Vinhas, F. Behlau, and G. José Silva Junior,
831 2018 Critical fungicide spray period for Citrus Black Spot control in São Paulo State,
832 Brazil. Plant Dis. 102: 334–340.
833 Fisher, M. C., D. A. Henk, C. J. Briggs, J. S. Brownstein, L. C. Madoff *et al.*, 2012 Emerging
834 fungal threats to animal, plant and ecosystem health. Nature 484: 186–194.
835 Fouché, S., C. Mence Plissonneau, and D. Croll, 2018 The birth and death of effectors in
836 rapidly evolving filamentous pathogen genomes. Curr. Opin. Microbiol. 46: 34–42.
837 Fudal, I., S. Ross, H. Brun, A.-L. Besnard, M. Ermel *et al.*, 2009 Repeat-Induced Point
838 Mutation (RIP) as an alternative mechanism of evolution toward virulence in
839 Leptosphaeria maculans. Mol. Plant-Microbe Interact. 22: 932–941.
840 Gardiner, D. M., M. C. McDonald, L. Covarelli, P. S. Solomon, A. G. Rusu *et al.*, 2012
841 Comparative pathogenomics reveals horizontally acquired novel virulence genes in
842 fungi infecting cereal hosts (A. P. Mitchell, Ed.). PLoS Pathog. 8: e1002952.
843 Gibriel, H. A. Y., B. P. H. J. Thomma, and M. F. Seidl, 2016 The age of effectors: genome-
844 based discovery and applications. Phytopathology 106: 1206–1212.
845 Glienke, C., O. L. Pereira, D. Stringari, J. Fabris, V. Kava-Cordeiro *et al.*, 2011 Endophytic
846 and pathogenic Phyllosticta species, with reference to those associated with Citrus
847 Black Spot. Persoonia 26: 47–56.
848 Grigoriev, I. V., R. Nikitin, S. Haridas, A. Kuo, R. Ohm *et al.*, 2014 MycoCosm portal: gearing
849 up for 1000 fungal genomes. Nucleic Acids Res. 42: D699–D704.
850 Gu, Z., R. Eils, and M. Schlesner, 2016 Complex heatmaps reveal patterns and correlations
851 in multidimensional genomic data. Bioinformatics 32: 2847–2849.
852 Guarnaccia, V., T. Gehrmann, G. J. Silva-Junior, P. H. Fourie, S. Haridas *et al.*, 2019

Phyllosticta citricarpa and sister species of global importance to Citrus. Mol. Plant Pathol. 20: 1619–1635.

Guarnaccia, V., J. Z. Groenewald, H. Li, C. Glienke, E. Carstens *et al.*, 2017 First report of Phyllosticta citricarpa and description of two new species, P. paracapitalensis and P. paracitricarpa, from citrus in Europe. Stud. Mycol. 87: 161–185.

Gurevich, A., V. Saveliev, N. Vyahhi, and G. Tesler, 2013 QUAST: quality assessment tool for genome assemblies. Bioinformatics 29: 1072–1075.

Hane, J. K., J. Paxman, D. A. B. Jones, R. P. Oliver, and P. de Wit, 2020 "CATAStrophy," a genome-informed trophic classification of filamentous plant pathogens – how many different types of filamentous plant pathogens are there? Front. Microbiol. 10: 3088.

Hane, J. K., T. Rouxel, B. J. Howlett, G. H. Kema, S. B. Goodwin *et al.*, 2011 A novel mode of chromosomal evolution peculiar to filamentous Ascomycete fungi. Genome Biol. 12: 1–16.

Haridas, S., R. Albert, M. Binder, J. Bloem, K. LaButti *et al.*, 2020 101 Dothideomycetes genomes: a test case for predicting lifestyles and emergence of pathogens. Stud. Mycol. 96: 141–153.

ten Have, A., K. B. Tenberge, J. A. E. Benen, P. Tudzynski, J. Visser *et al.*, 2002 The contribution of cell wall degrading enzymes to pathogenesis of fungal plant pathogens, pp. 341–358 in *The Mycota XI: Agricultural Applications*, Springer Berlin Heidelberg, Berlin, Heidelberg.

Hegedus, D. D., and S. R. Rimmer, 2005 Sclerotinia sclerotiorum: When "to be or not to be" a pathogen? FEMS Microbiol. Lett. 251: 177–184.

Hüttel, W., and M. Müller, 2021 Regio- and stereoselective intermolecular phenol coupling enzymes in secondary metabolite biosynthesis. Nat. Prod. Rep. 38: 1011–1043.

Kabbage, M., O. Yarden, and M. B. Dickman, 2015 Pathogenic attributes of Sclerotinia sclerotiorum: switching from a biotrophic to necrotrophic lifestyle. Plant Sci. 233: 53–60.

Kim, K.-T., J. Jeon, J. Choi, K. Cheong, H. Song *et al.*, 2016 Kingdom-wide analysis of fungal Small Secreted Proteins (SSPs) reveals their potential role in host association.

881 Front. Plant Sci. 7: 186.

882 King, B. C., K. D. Waxman, N. V Nenni, L. P. Walker, G. C. Bergstrom *et al.*, 2011 Arsenal of
883 plant cell wall degrading enzymes reflects host preference among plant pathogenic
884 fungi. Biotechnol. Biofuels 4: 1–14.

885 Klosterman, S. J., K. V. Subbarao, S. Kang, P. Veronese, S. E. Gold *et al.*, 2011
886 Comparative genomics yields insights into niche adaptation of plant vascular wilt
887 pathogens (J. L. Dangl, Ed.). PLoS Pathog. 7: e1002137.

888 Kotzé, J. M., 2000 Black spot, pp. 23–25 in *Compendium of Citrus Diseases. American*
889 *Phytopathological Society Press Inc. Timmer, L.W., Garnsey, S.M. and Graham, J.H.*,
890 The American Phytopathological Society.

891 Kubicek, C. P., T. L. Starr, and N. L. Glass, 2014 Plant cell wall–degrading enzymes and
892 their secretion in plant-pathogenic fungi. Annu. Rev. Phytopathol. 52: 427–51.

893 Kubicek, C. P., A. S. Steindorff, K. Chenthamara, G. Manganiello, B. Henrissat *et al.*, 2019
894 Evolution and comparative genomics of the most common Trichoderma species. BMC
895 Genomics 20: 1–24.

896 Lindo, L., R. E. Cardoza, A. Lorenzana, P. A. Casquero, and S. Gutiérrez, 2020
897 Identification of plant genes putatively involved in the perception of fungal ergosterol-
898 squalene. J. Integr. Plant Biol. 62: 927–947.

899 Lombard, V., H. Golaconda Ramulu, E. Drula, P. M. Coutinho, and B. Henrissat, 2014 The
900 carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42: D490–
901 D495.

902 Ma, L.-J., H. C. van der Does, K. A. Borkovich, J. J. Coleman, M.-J. Daboussi *et al.*, 2010
903 Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature
904 464: 367–73.

905 Manni, M., M. R. Berkeley, M. Seppey, F. A. Simão, and E. M. Zdobnov, 2021 BUSCO
906 update: novel and streamlined workflows along with broader and deeper phylogenetic
907 coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol. Biol. Evol. 38:
908 4647–4654.

909 Marçais, G., A. L. Delcher, A. M. Phillippy, R. Coston, S. L. Salzberg *et al.*, 2018 MUMmer4:
910 A fast and versatile genome alignment system. PLOS Comput. Biol. 14: e1005944.
911 Mir, A. A., S.-Y. Park, M. A. Sadat, S. Kim, J. Choi *et al.*, 2015 Systematic characterization
912 of the peroxidase gene family provides new insights into fungal pathogenicity in
913 Magnaporthe oryzae. Sci. Rep. 5: 1–14.
914 Möller, M., and E. H. Stukenbrock, 2017 Evolution and genome architecture in fungal plant
915 pathogens. Nat. Rev. Microbiol. 15: 756–771.
916 O'connell, R. J., M. R. Thon, S. Hacquard, S. G. Amyotte, J. Kleemann *et al.*, 2012 Lifestyle
917 transitions in plant pathogenic Colletotrichum fungi deciphered by genome and
918 transcriptome analyses. Nat. Genet. 44:.
919 Ohm, R. A., N. Feau, B. Henrissat, C. L. Schoch, B. A. Horwitz *et al.*, 2012 Diverse lifestyles
920 and strategies of plant pathogenesis encoded in the genomes of eighteen
921 dothideomycetes fungi (A. Andrianopoulos, Ed.). PLoS Pathog. 8: e1003037.
922 Oliver, R. P., and S. V. S. Ipcho, 2004 Arabidopsis pathology breathes new life into the
923 necrotrophs-vs.-biotrophs classification of fungal pathogens. Mol. Plant Pathol. 5: 347–
924 352.
925 Petters-Vandresen, D. A. L., B. J. Rossi, J. Z. Groenewald, P. W. Crous, M. A. Machado *et*
926 *al.*, 2020 Mating-type locus rearrangements and shifts in thallism states in Citrus-
927 associated Phyllosticta species. Fungal Genet. Biol. 144: 103444.
928 Plissonneau, C., J. Benevenuto, N. Mohd-Assaad, S. Fouché, F. E. Hartmann *et al.*, 2017
929 Using population and comparative genomics to understand the genetic basis of
930 effector-driven fungal pathogen evolution. Front. Plant Sci. 8: 119.
931 Lo Presti, L., D. Lanver, G. Schweizer, S. Tanaka, L. Liang *et al.*, 2015 Fungal effectors and
932 plant susceptibility. Annu. Rev. Plant Biol. 66: 513–545.
933 R Core Team (2021). R: A language and environment for statistical computing. R
934 Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
935 Rocha, M. C., K. F. de Godoy, R. Bannitz-Fernandes, J. H. T. M. Fabri, M. M. F. Barbosa *et*
936 *al.*, 2018 Analyses of the three 1-Cys Peroxiredoxins from Aspergillus fumigatus reveal

that cytosolic Prx1 is central to $H_2O_2$ metabolism and virulence. Sci. Reports 2018 81 8: 1–18.

Rodrigues, C. M., M. A. Takita, N. V. Silva, M. Ribeiro-Alves, and M. A. Machado, 2019 Comparative genome analysis of Phyllosticta citricarpa and Phyllosticta capitalensis, two fungi species that share the same host. BMC Genomics 20: 554.

Rovenich, H., J. C. Boshoven, and B. P. H. J. Thomma, 2014 Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. Curr. Opin. Plant Biol. 20: 96–103.

Schwelm, A., and R. E. Bradshaw, 2010 Genetics of Dothistromin biosynthesis of Dothistroma septosporum: an update. Toxins (Basel). 2: 2680–2698.

Shin, J., D. Bui, Y. Lee, H. Nam, S. Jung *et al.*, 2017 Functional characterization of cytochrome P450 monooxygenases in the cereal head blight fungus Fusarium graminearum. Environ. Microbiol. 19: 2053–2067.

Siewers, V., M. Viaud, D. Jimenez-Teja, I. Collado, C. Gronover *et al.*, 2005 Functional analysis of the cytochrome P450 monooxygenase gene bcbot1 of Botrytis cinerea indicates that botrydial is a strain-specific virulence factor. Mol. Plant. Microbe. Interact. 18: 602–612.

Singh, R. P., D. P. Hodson, J. Huerta-Espino, Y. Jin, S. Bhavani *et al.*, 2011 The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annu. Rev. Phytopathol. 49: 465–481.

Spanu, P. D., 2012 The genomics of obligate (and nonobligate) biotrophs. Annu. Rev. Phytopathol. 50: 91–109.

Sperschneider, J., P. N. Dodds, D. M. Gardiner, J. M. Manners, K. B. Singh *et al.*, 2015 Advances and challenges in computational prediction of effectors from plant pathogenic fungi. PLoS Pathog. 11: e1004806.

Sperschneider, J., D. M. Gardiner, P. N. Dodds, F. Tini, L. Covarelli *et al.*, 2016 EffectorP: predicting fungal effector proteins from secretomes using machine learning. New Phytol. 210: 743–761.

Stergiopoulos, I., Y. A. I. Kourmpetis, J. C. Slot, F. T. Bakker, P. J. G. M. De Wit *et al.*, 2012

In silico characterization and molecular evolutionary analysis of a novel superfamily of fungal effector proteins. Mol. Biol. Evol. 29: 3371–3384.

Temporini, E. D., and H. D. Vanetten, 2002 Distribution of the pea pathogenicity (PEP ) genes in the fungus Nectria haematococca mating population VI. Curr. Genet. 41: 107–114.

Veloso, J., and J. A. L. Van Kan, 2018 Many shades of grey in Botrytis-host plant interactions. Trends Plant Sci. 23: 613–622.

Wang, X., G. Chen, F. Huang, J. Zhang, K. D. Hyde *et al.*, 2012 Phyllosticta species associated with citrus diseases in China. Fungal Divers. 52: 209–224.

Wang, M., B. Liu, R. Ruan, Y. Zeng, J. Luo *et al.*, 2020 Genomic sequencing of Phyllosticta citriasiana provides insight into its conservation and diversification with two closely related Phyllosticta species associated with Citrus. Front. Microbiol. 10: 2979.

Wikee, S., L. Lombard, P. W. Crous, C. Nakashima, K. Motohashi *et al.*, 2013a *Phyllosticta capitalensis*, a widespread endophyte of plants. Fungal Divers. 60: 91–105.

Wikee, S., L. Lombard, C. Nakashima, K. Motohashi, E. Chukeatirote *et al.*, 2013b A phylogenetic re-evaluation of *Phyllosticta* (Botryosphaeriales). Stud. Mycol. 76: 1–29.

de Wit, P. J. G. M., R. Mherabi, H. A. Van den burg, and I. Stergiopoulos, 2009 Fungal effector proteins: past, present and future. Mol. Plant Pathol. 10: 735–747.

Wulandari, N. F., C. To-anun, K. D. Hyde, L. M. Duong, J. de Gruyter *et al.*, 2009 *Phyllosticta citriasiana* sp. nov., the cause of Citrus tan spot of *Citrus maxima* in Asia. Fungal Divers. 34: 23–39.

Yang, Y., Y. Zhang, B. Li, X. Yang, Y. Dong *et al.*, 2018 A Verticillium dahliae pectate lyase induces plant immune responses and contributes to virulence. Front. Plant Sci. 9: 1271.

Zhang, H., T. Yohe, L. Huang, S. Entwistle, P. Wu *et al.*, 2018 dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 46: W95–W101.