Fungal Genetics and Biology 125 (2019) 84–92

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Genomic comparisons of the laurel wilt pathogen, *Raffaelea lauricola*, and related tree pathogens highlight an arsenal of pathogenicity related genes

Jorge R. Ibarra Caballero[a], Jongbum Jeon[b], Yong-Hwan Lee[b], Stephen Fraedrich[c], Ned B. Klopfenstein[d], Mee-Sook Kim[e], Jane E. Stewart[a,*]

[a] Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO 80523, USA
[b] Department of Agricultural Biotechnology, Seoul National University, Seoul, Republic of Korea
[c] USDA Forest Service, Southern Research Station, Athens, GA 30602, USA
[d] USDA Forest Service, Rocky Mountain Research Station, Moscow, ID 83843, USA
[e] USDA Forest Service, Pacific Northwest Research Station, Corvallis, OR 97331, USA

ARTICLE INFO

Keywords:
Raffaelea
Grosmannia
Ophiostoma
Ophiostomatales
Tree wilting pathogens
Fungal pathogenicity-related proteins

ABSTRACT

Raffaelea lauricola is an invasive fungal pathogen and symbiont of the redbay ambrosia beetle (*Xyleborus glabratus*) that has caused widespread mortality to redbay (*Persea borbonia*) and other Lauraceae species in the southeastern USA. We compare two genomes of *R. lauricola* (C2646 and RL570) to seven other related Ophiostomatales species including *R. aguacate* (nonpathogenic close relative of *R. lauricola*), *R. quercus-mongolicae* (associated with mortality of oaks in Korea), *R. quercivora* (associated with mortality of oaks in Japan), *Grosmannia clavigera* (cause of blue stain in conifers), *Ophiostoma novo-ulmi* (extremely virulent causal agent of Dutch elm disease), *O. ulmi* (moderately virulent pathogen that cause of Dutch elm disease), and *O. piceae* (blue-stain saprophyte of conifer logs and lumber). Structural and functional annotations were performed to determine genes that are potentially associated with disease development. *Raffaelea lauricola* and *R. aguacate* had the largest genomes, along with the largest number of protein-coding genes, genes encoding secreted proteins, small-secreted proteins, ABC transporters, cytochrome P450 enzymes, CAZYmes, and proteases. Our results indicate that this large genome size was not related to pathogenicity but was likely lineage specific, as the other pathogens in *Raffaelea* (*R. quercus-mongolicae* and *R. quercivora*) had similar genome characteristics to the *Ophiostoma* species. A diverse repertoire of wood-decaying enzymes were identified in each of the genomes, likely used for toxin neutralization rather than wood degradation. Lastly, a larger number of species-specific, secondary metabolite, synthesis clusters were identified in *R. lauricola* suggesting that it is well equipped as a pathogen, which could explain its success as a pathogen of a wide range of lauraceous hosts.

1. Introduction

Raffaelea is a genus of asexual fungi comprised within the Ophiostomatales (Harrington, 2005). Members of the Ophiostomatales consist primarily of symbionts of ambrosia beetles that carry these fungi within specialized structures known as mycangia (Vanderpool et al., 2018). Several Ophiostomatales species are also known to cause diseases, such as vascular wilts and vascular staining. One of these species, *Raffaelea lauricola* T.C. Harr., Fraedrich & Aghayeva, causes laurel wilt disease that is responsible for widespread mortality of redbay (*Persea borbonia*) and other lauraceous species in the southeastern United States (USA). This laurel wilt pathogen is a symbiont of the redbay ambrosia beetle, *Xyleborus glabratus* (Harrington et al., 2008; Fraedrich et al., 2008). *Raffaelea lauricola* and *X. glabratus* are considered to be native in Asia (Japan, Taiwan, India, and Myanmar) and were likely introduced to the USA via packing materials (Harrington et al., 2011). Since their introduction at Port Wentworth, Georgia, USA, around 2002, the spread of the beetle vector and fungal pathogen have resulted in laurel wilt disease throughout the southeastern coastal plains and westward as far as Texas, USA. Laurel wilt disease is now found in nine states of the USA (Barton et al., 2016), where it can affect numerous hosts including redbay, swampbay (*P. palustris*), sassafras (*Sassafras albidum*), avocado (*P. americana*), pondspice (*Litsea aestivalis*), silkbay (*P. humilus*), spicebush (*Lindera benzoin*), and European bay laurel (*Laurus nobilis*) (Fraedrich et al., 2015, 2016; Hughes et al., 2011, 2012, 2014; Mayfield et al., 2008; Smith et al., 2009a, 2009b). *Raffaelea lauricola* has also been found in association with other ambrosia beetle species that attack lauraceous species (Fraedrich et al., 2011; Ploetz et al., 2017), which

* Corresponding author.
E-mail address: Jane.Stewart@colostate.edu (J.E. Stewart).

https://doi.org/10.1016/j.fgb.2019.01.012
Received 11 May 2018; Received in revised form 9 January 2019; Accepted 31 January 2019
Available online 01 February 2019
1087-1845/

Table 1
Genome assembly statistics of *Raffaelea lauricola* and related beetle-vectored fungi as generated in the manuscript.

	Raffaelea lauricola C2646*	*R. lauricola RL570**	*R. aquacate*#	*R. quercus-mongolicae**	*R. quercivora**	*Grosmannia clavigera**	*Ophiostoma piceae*#	*O. novo-ulmi**	*O. ulmi**
Genome size (Mb)	35.40	34.64	35.18	27.00	26.41	29.79	32.84	31.85	31.46
Scaffold number	1535	207	414	43	23	289	45	10	792
Contig number	1566	578	1006	200	23	333	343	161	2634
Scaffold N50 (Mb)	0.13	2.87	0.45	2.2	3.69	1.99	1.45	3.66	1.00
G-C contents (%)	55.32	55.34	57.54	54.25	57.02	53.36	53.35	50.10	50.52
Rep. DNA (%)	12.68	9.75	5.75	4.04	7.02	10.73	3.83	5.15	3.88

* These species have been described as causal agents of wilt diseases in woody plants.
\# These species are known to be insect vectored, but not known as pathogens.

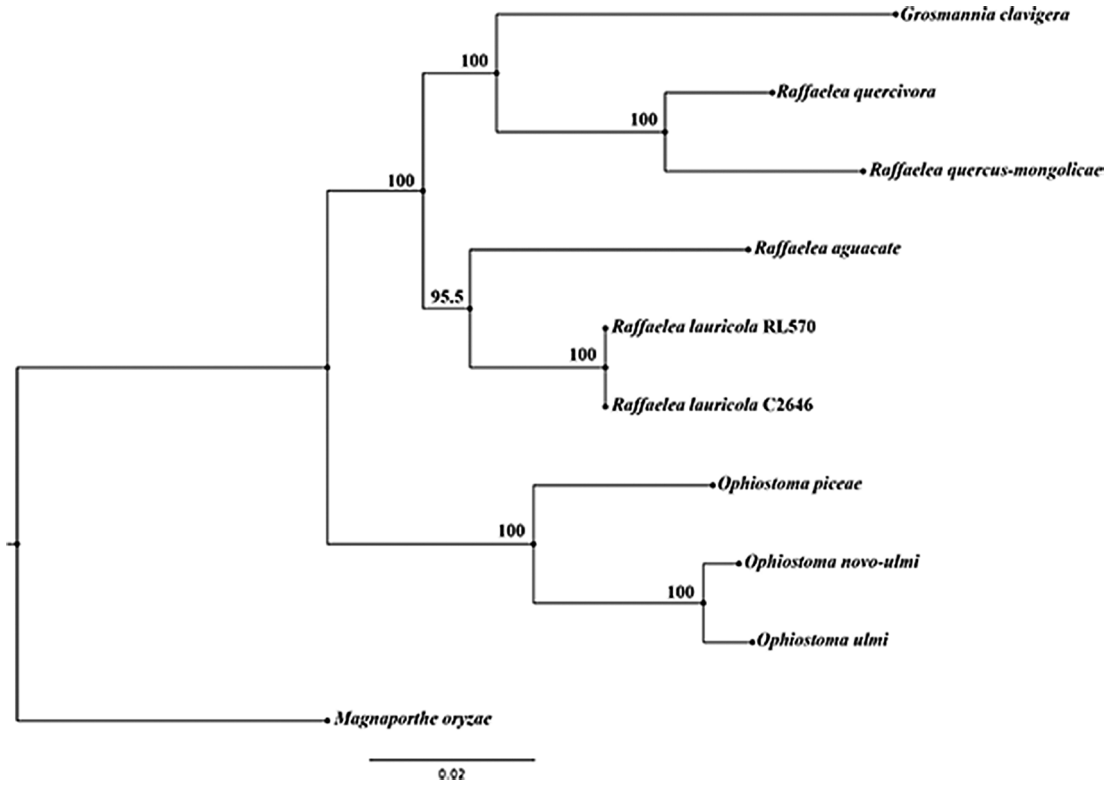

Fig. 1. Whole genome phylogenetic tree of Ophiostomatales species and other Sordarioimycetes as implemented in PhyML using the likelihood methodologies. Numbers on each node indicated bootstrap values.

likely acquired the pathogen following attack of infected hosts.

Raffaelea lauricola systemically colonizes the sapwood of host plants and induces the formation of tyloses in xylem vessels, which are outgrowths from adjacent parenchyma cells. Tyloses plug the vessels and reduce hydraulic conductivity, which may limit pathogen spread in some hosts (Inch et al., 2012). Tyloses are commonly observed in other wilt diseases such as Dutch elm disease caused by *Ophiostoma ulmi* and *O. novo-ulmi* (Durkovic et al., 2014). Other ophiostomatoid fungi (*Grosmannia clavigera*, *O. piceae*) cause "blue stain" in the host wood, and when these fungi colonize the phloem and/or sapwood, they must cope with phenolic and terpenoid compounds produced by the host (Haridas et al., 2013). In general, species in the Ophiostomatales have limited capabilities to degrade the cellulose and lignin components of wood, and primarily obtain nutrients from the living cells in the sapwood (Lundell et al., 2014; Lee et al., 2002).

Herein, we report on a draft assembly of the *R. lauricola* genome and compare to genomes of closely related ophiotomatiod fungi to identify genes that are potentially associated with pathogenicity. Structural and functional annotations and analyses were performed for the assembled genome to determine genes that are potentially associated with disease development of this wilt pathogen; coded proteins were predicted and those likely related to pathogenicity were identified. The total number of putative genes within these protein families were compared to those of another *R. lauricola* isolate (Vanderpool et al., 2018) and seven other related fungi within the Ophiostomatales that range from virulent pathogens to saprophytes: *R. aguacate* (nonpathogenic close relative of *R. lauricola*), *R. quercus-mongolicae* (associated with mortality of oaks in Korea), *R. quercivora* (associated with mortality of oaks in Japan), *G. clavigera* (cause of blue stain in conifers), *O. novo-ulmi* (highly virulent causal agent of Dutch elm disease), *O. ulmi* (moderately virulent pathogen that cause of Dutch elm disease), and *O. piceae* (blue-stain saprophyte of conifer logs and lumber) (Saucedo-Carabez et al., 2018; Jeon et al., 2017; Kubono and Ito, 2002; DiGuistini et al., 2011; Comeau et al., 2014; Khoshraftar et al., 2013; Haridas et al., 2013).

2. Materials and methods

2.1. Biological material

Raffaelea lauricola isolate C2646 (CBS129006) was used for genome assembly in this study. This isolate was originally collected in Taiwan from a female redbay ambrosia beetle that is known to vector *R. lauricola* in the USA (Harrington et al., 2011). The isolate was identified by

Table 2
Total number of genes identified for genome structural and functional statistics of *Raffaelea lauricola* and related beetle-vectored fungi. The number in parentheses indicates the number of predicted secreted proteins. Gene numbers were obtained using the Maker pipeline, as described in the method section.

	Raffaelea lauricola C2646[*]	*R. lauricola RL570*[*]	*R. aquacate*[#]	*R. quercus-mongolicae*[*]	*R. quercivora*[*]	*Grosmannia clavigera*[*]	*Ophiostoma piceae*[#]	*O. novo-ulmi*[*]	*O. ulmi*[*]
tRNA genes	140	156	152	147	154	190	248	254	261
Protein coding genes	11,173	11,069	11,932	9,049	8,917	8,855	9,277	8,837	8,846
Secreted proteins	676	654	671	501	477	456	524	511	528
Small secreted proteins	305	293	270	193	174	186	215	223	245
ABC transporters	57	56	59	41	40	45	34	32	32
Cytochrome P450	75 (0)	71 (0)	67 (0)	54 (0)	46 (0)	57 (0)	46 (0)	45 (0)	44 (0)
Laccase	15 (8)	14 (8)	11 (11)	14 (7)	11 (5)	12 (5)	11 (6)	11 (7)	10 (7)
Peroxidase[A]	13 (2)	13 (2)	13 (2)	13 (5)	13 (4)	11 (3)	16 (2)	15 (3)	15 (3)
Tannase/feruloyl esterase	5 (4)	5 (4)	11 (11)	5 (4)	4 (4)	2 (1)	3 (2)	4 (0)	5 (0)
CAZYmes[B]	313 (1 1 9)	308 (1 1 5)	351 (1 2 5)	275 (1 1 6)	267 (1 0 4)	223 (88)	244 (1 0 8)	246 (1 0 5)	285 (1 0 8)
Proteases	201 (39)	198 (40)	208 (45)	183 (35)	184 (36)	182 (39)	169 (36)	171 (34)	169 (37)
Polyketide synthases[C,D]	17[1][E]	18[E][1]	8[F] [1]	15[G] [1]	13[G] [1]	11 [1]	8 [1]	8 [1]	8 [1]
Non-ribosomal peptide synthases[D]	9[E]	9[E]	3[F]	5[G]	6 [G]	2[H]	1	1	1

[A] This number includes lignin peroxidases (AA2) and other peroxidases.
[B] This number does not include glycosyl transferases, carbohydrate-binding modules or laccases.
[C] The number of Type I PKs clusters. The number of Type III PKs clusters are in [ ].
[D] The number and type of peptide synthases and non-ribosomal peptide synthases identified in each genome: [E]5 T1pks-Nrps; [F]1 T1pks-Nrps; [G]2 T1pks-Nrps; [H]1 Terpene-Nrps.
[*] These species have been described as causal agents of wilt diseases in woody plants.
[#] These species are known to be insect vectored, but not known as pathogens.

its cultural morphology, mucoid growth, size and shape of its conidiophores and budding conidia, and sequencing of the large subunit of the rDNA (HQ688666) (Harrington et al., 2011).

2.2. DNA extraction and sequencing

DNA was extracted from *R. lauricola* isolate C2646 using cetyltrimethylammonium bromide (CTAB) extraction protocol. The isolate was grown at room temperature (ca. 22 °C) on a bench shaker (70 rpm) for 7 d in flask with 2% malt extract broth. Similar to Li et al. (2017), ca. 500 mg of mycelium finely ground in liquid nitrogen were added to 17.5 ml CTAB lysis buffer. The lysis buffer included: 6.5 ml of Buffer A (0.35 M sorbitol; 0.1 M Tris–HCl, pH 9; and 5 mM EDTA, pH 8), 6.5 ml of Buffer B (0.2 M Tris–HCl, pH 9; 50 mM EDTA, pH 8; 2 M NaCl; 2% CTAB), 2.6 ml of Buffer C (5% sarkosyl), 1.75 ml (0.1%), polyvinylpyrrolidone (PVP), and 1.25 µl (25 mg = 750 U) proteinase K. Using a 2010 Geno/Grinder (SPEX SamplePrep, Metuchen, NJ, USA) the mixture was agitated with two 5-mm glass beads (VWR Soda Lime, Radnor, PA, USA) at 1750 rpm for 2 min. After adding 5.75 ml 5 M potassium acetate, tubes were inverted 10 times, incubated on ice for 30 min, and centrifuged for 20 min at 14,000g. The supernatant was added to chloroform-isoamyl alcohol (v/v 24:1), mixed, centrifuged for 10 min at 14,000g, and then 100 µl RNase A (10 mg ml^{-1}, 50 U mg) was added to the upper phase. The solution was incubated at 37 °C for 120 min. Isopropanol was added at equal volume and 1/10 vol 3 M sodium acetate was added, the mixture was incubated at 25 °C for 5 min, and then centrifuged at 14,000g for 30 min, after which the supernatant was discarded. After rinsing twice with 70% ethanol and air-drying overnight, the DNA pellet was dissolved in 100 µl deionized H_2O.

Genomic DNA was submitted to the Georgia Genomics Facility (Athens, GA, USA) for next-generation sequencing library preparation. Illumina sequencing reactions followed the NextSeq platform, based on a paired end 150-bp (PE150) protocol (Illumina Inc. San Diego, CA, USA).

2.3. Genome assembly and genome statistics

Genome assembly was performed using A5-miseq, version 20160825 (Coil et al., 2015). Assembly metrics were obtained using the QUAST web server (Gurevich et al., 2013).

2.4. Evaluation of genome assembly and annotation

Completeness of the genome assembly and structural annotation was evaluated using BUSCO 2.0b2 (Simão et al., 2015). The "fungi lineage" dataset (fungi-odb9), that includes the appropriate set of orthologs for the fungi examined, was selected for the evaluation. BLASTp (Camacho et al., 2009) was also used to compare predicted proteins in *R. lauricola* isolate C2646 and *R. aguacate* against all fungal proteins in the Uniprot database to determine the percentage of proteins with significant matches to already described proteins.

2.5. Phylogenetic analysis

A whole genome phylogenetic tree was created using Realphy 1.12 (Bertels et al., 2014) with Bowtie2 2.3.3.1 (Langmead and Salzberg, 2012) and PhyML (Guindon et al., 2010), including the genome assembly of *R. lauricola* C2646 (GenBank QDHB00000000) and the following whole genome assemblies: *R. lauricola* strain RL570, GenBank PCDG00000000.1; *R. aguacate*, GeneBank PCDF00000000.1; *R. quercus-mongolicae*, GenBank NIPS00000000.1; *R. quercivora*, GenBank PCDE00000000.1; *G. clavigera* kw1407, GenBank ACXQ00000000.2; *O. piceae* UAMH 11346, from JGI MycoCosm (Haridas et al., 2013); *O. novo-ulmi* H327, from JGI MycoCosm (Comeau et al., 2014); and *O. ulmi*, from the *Ophiostoma ulmi* resource browser (http://www.moseslab.csb.utoronto.ca/o.ulmi/). The number of bootstrap replicates in PhyML was set to 200.

2.6. Structural annotation

To obtain a structural annotation of the genome assembly, the Maker v.2.31.8 pipeline was used (Cantarel et al., 2008). First, a custom set of repetitive sequences was obtained using RepeatModeler v.1.0.11 (Smit and Hubley, 2017). Then, the following programs were included in the Maker pipeline: (1) RepeatMasker v.4.0.6 (Smit et al., 1996) to mask interspersed repeats and low complexity DNA sequences using the repeats set obtained by RepeatModeler; (2) three gene predictors: GeneMark-ES (Ter-Hovhannisyan et al., 2008), which does not require curated training sets to predict genes in fungal genomes; SNAP (Zaharia et al., 2011), a high-performance gene finder; and Augustus (Keller

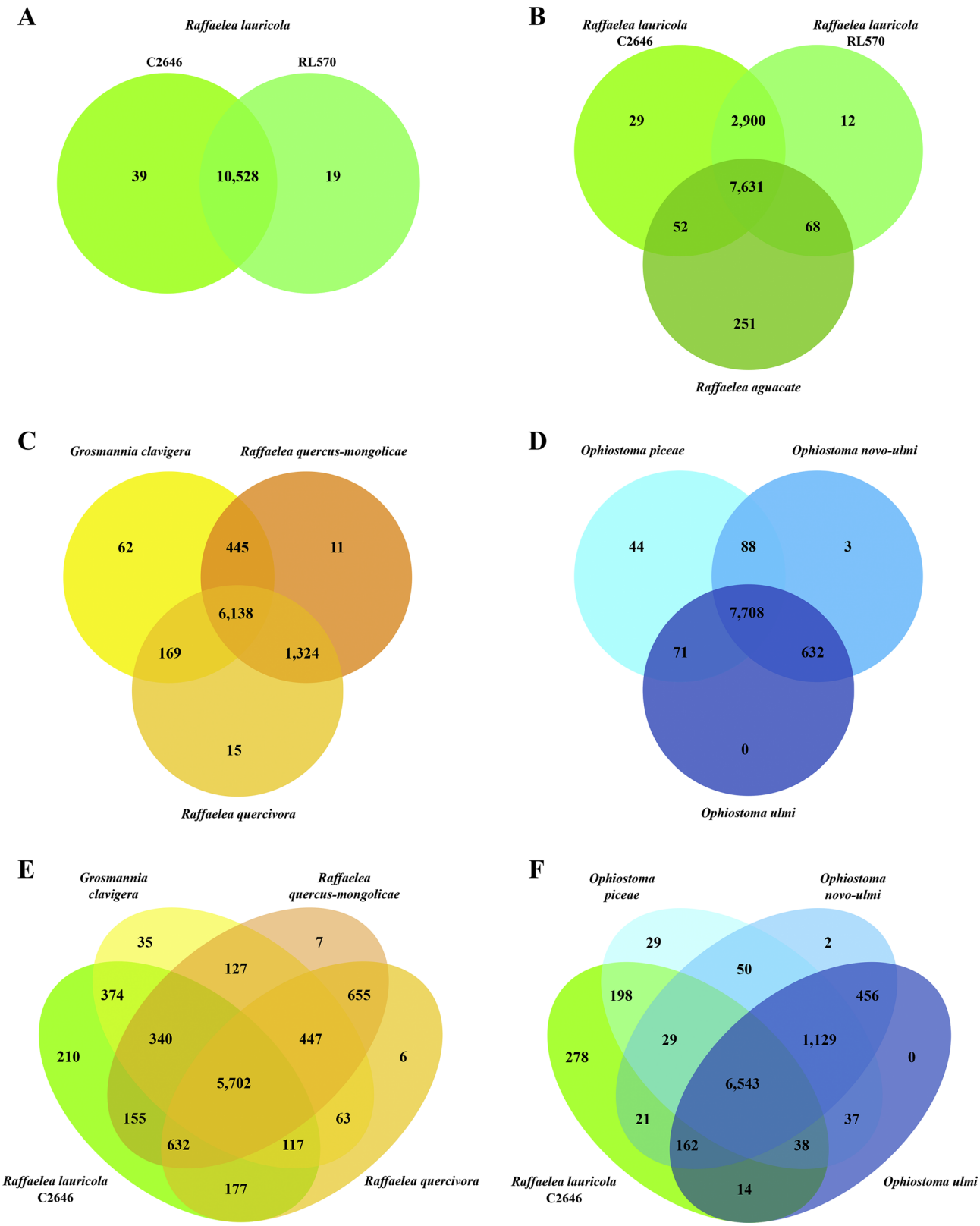

Fig. 2. Distribution of orthologous and non-orthologous proteins comparing A: *Raffaelea lauricola* C2646 and *R. lauricola* RL570, B: *R. lauricola* C2646, *R. lauricola* RL570 and *R. aguacate*, C: *R. quercus-mongolicae*, *R. quercivora*, and *Grosmannia clavigera*, D: *Ophiostoma piceae*, *O. novo-ulmi*, and *O. ulmi*, E: *R. lauricola* C2646 to *R. quercus-mongolicae*, *R. quercivora*, and *Grosmannia clavigera*, and F: *R. lauricola* C2646 to *R. quercus-mongolicae*, *R. quercivora*, and *Grosmannia clavigera.* Each species were color-coded. Numbers represent the number of protein clusters either unique to a species or common between or among species. The central number in each comparison shows the number of proteins shared with all species.

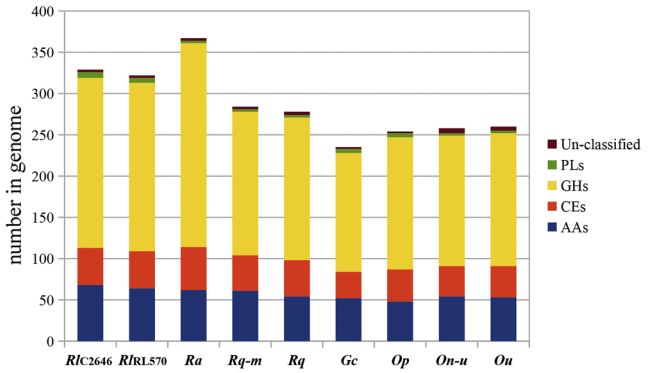

Fig. 3. Distribution of CAZYmes in the nine analyzed genomes: *Raffaelea lauricola* (RlC2646); *Raffaelea lauricola* (RlRL570), *R. aguacate* (Ra), *R. quercus-mongolicae* (Rqm); *R. quercivora* (Rq), *Grosmannia clavigera* (Gc); *Ophiostoma piceae* (Op), *O. novo-ulmi* (On-u) and *O. ulmi* (Ou). Carbohydrate enzyme gene families are indicated by AA: auxiliary activity, CE: carbohydrate esterase, GH: glycoside hydrolase, and PL: polysaccharide lyase.

et al., 2011), the widely used eukaryotes gene predictor; and (3) tRNAscan (Lowe and Eddy, 1997) to identify tRNA genes in the genomic sequence.

Transcript sequences obtained from the NCBI were included as "alternate organism EST Evidence". Ophiostomatales protein sequences from the RefSeq database were included as "Protein Homology Evidence". For each species, the following was included as "same organism EST Evidence": *G. clavigera*: transcripts from NCBI BioProject PRJNA40239; *R. quercus-mongolicae*: RNASeq assembled transcripts (Kim, M.-S. unpublished, Genbank Accession GHDV00000000); *O. piceae*: transcripts from the genome annotation from JGI MycoCosm (Haridas et al., 2013); *O. novo-ulmi* and *O. ulmi*: transcripts from supplementary material in Comeau et al. (2014); *R. lauricola* RL570 isolate, *R. quercivora* and *R. aguacate*: corresponding transcripts from Vanderpool et al. (2018).

2.7. Functional annotation

For all the isolates, functional annotation of the predicted proteins of ≥50 amino acids was obtained using InterProScan 5.24–63.0 (Jones et al,. 2014) including the program Pfam to search for matches of amino acid sequences in the corresponding database. Previously developed pipelines (Park et al., 2008; Choi et al., 2014) were used to identify cytochrome P450 and peroxidases, along with the InterProScan output. Putative secreted proteins were identified using DeepLoc v.1.0 (Almagro Armenteros et al., 2017), a program that uses deep neural networks based on the protein sequence information. Small secreted proteins were defined as those smaller than 300 amino acids. Secondary metabolism genes were detected using antiSMASH 4.0 (Blin et al., 2017), including both the genome assembly and gff3 annotation (from Maker) as input files. The cluster backbones were derived via "detailed annotation" from antiSMASH output in two separate processes: keeping only the core biosynthetic genes from the main cluster (similar to Sbaraini et al., 2017) or including all genes (i.e., also containing other genes interspersed between the core biosynthetic genes). Putative Carbohydrate-Active Enzymes (CAZYmes), including laccases, were identified by obtaining InterPro signatures known to be present in CAZYmes from the predicted proteins and then combining that set of proteins with the one obtained after all predicted proteins were submitted to the web server and database dbCAN (Yin et al., 2012). dbCAN was run using hmmscan. Tannases were detected based on the presence of corresponding Pfam (PF07519) and InterPro (IPR011118) domains. Proteases were identified based on corresponding Interpro signatures.

2.8. Orthologous proteins

The sets of predicted proteins from each of the eight species (nine isolates) included in this work were used to identify orthologous and non-orthologous (species exclusive) groups using the OrthoVenn web server (Wang et al., 2015). Default values were used for the two parameters that can be adjusted when using the OrthoVenn web server: e-value and inflation value (1e−5 and to 1.5, respectively) as suggested by Wang et al. (2015).

3. Results and discussion

The *R. lauricola* (C2646) genome assembly via A5-miseq resulted in 1535 scaffolds with a total length of 35.40 Mb (35,404,324 bp) (Genbank Accession QDHB00000000) and under Bioproject PRJNA513385. The largest contig contained 536,604 bp. The N50 score was 136,309 bp and the GC content was 55.32%. The total length of the genome was larger than other *Raffaelea* species found within the NBCI database including *R. quercivora* (GenBank BCFZ00000000.1) with a genome size of 25.41 Mb and *R. quercus-mongolicae* (GenBank NIPS00000000.1) with a genome size of 27.00 Mb (Table 1), but only slightly larger than *R. aguacate* (GenBank PCDF00000000.1) with a genome size of 35.18 Mb. The genome of *R. lauricola* was also larger than genomes of *G. clavigera*, *O. piceae*, *O. novo-ulmi*, and *O. ulmi*, which had sizes ranging from 29.79 to 31.46 Mb. Completeness of the genome assembly, assessed by BUSCO, indicated a high-quality assembly with a score of 97.9%. Close to 13% of the genome consisted of low complexity and interspersed repetitive DNA, which was greater than that found in the other genome assemblies (Table 1; Supplementary Table 1).

The use of a whole genome phylogenetic approach highlights relationships among the eight Ophiostomatales species studied, using *Magnaporthe oryzae* as outgroup species (Fig. 1). Of the total 40,643 sites used to generate the phylogeny, 6,915 (17%) were polymorphic. The two *R. lauricola* genomes and *R. aguacate* were clustered into a well-supported group, whereas *R. quercus-mongolicae*, *R. quercivora*, and *G. clavigera* were together in a separate well-supported group, and the three *Ophiostoma* spp. were clustered within another distinct group. These results concur with other studies comparing the phylogenetic placement of *Raffaelea* and *Ophiostoma* using multiple loci (de Beer and Wingfield, 2013; Dreaden et al., 2014; Vanderpool et al., 2018).

The two *R. lauricola* genomes (C2646 and RL570) and the *R. aguacate* genome had the largest number of protein-coding genes, ranging from 11,069–11,932, compared to the genomes of the other Ophiostomatales species, which ranged from 8,837–9,277; however, fewer tRNA genes were found in *R. lauricola* C2646 compared to *R. lauricola* RL570 and the other Ophiostomatales species (Table 2). Completeness of *R. lauricola* C2646 structural annotation was 97.6%, again indicating high quality in the genome annotation. Herein, a larger number of proteins were identified in the *R. lauricola* genomes, C2646 and RL570, than was previously published by Vanderpool et al. (2018). Annotations were based solely on genomic data or included transcript data collected from other sources (see methods). We did not perform transcriptomic analyses, and therefore, this represents a potential limitation with our data. Further, the Vanderpool et al. (2018) protein data set was restricted to those with RNAseq, protein evidence, or Pfam domains, whereas we kept all gene models predicted by Maker. This likely resulted in less conservative results (i.e., more predicted proteins) for all the genomes (DiGuistini et al., 2011; Sbaraini et al., 2017).

The annotated protein sets for each species were used to perform pairwise analyses of orthologous proteins using OrthoVenn (Wang et al., 2015), beginning with comparisons within each well-supported group observed in the phylogenetic tree (Fig. 1). Similar to the phylogenetic tree, species within a group were more similar to each other compared with species comprised in other groups. As expected, only few differences were observed between the two *R. lauricola* isolates (Fig. 2A). Comparing *R. lauricola* and *R. aguacate*, the two *R. lauricola* isolates showed much more similarity to each other than to *R. aguacate* (Fig. 2B). Within the *R. quercivora/R. quercus-mongolicae/G. clavigera* cluster, the two *Raffaelea* species had a higher number of orthologous protein clusters compared to

Table 3
Biosynthetic gene cluster (BGC) conservation in the Ophiostomataceae family for *Raffaelea lauricola* (RlC2646, RlRL570), *R. aguacate* (Ra), *R. quercus-mongolicae* (Rq-m); *R. quercivora* (Rq), *Grosmannia clavigera* (Gc), *Ophiostoma piceae* (Op), *O. novo-ulmi* (On-u) and *O. ulmi* (Ou). Domains for polyketide synthase (Pks) and non-ribosomal peptide synthase (Nrps) backbone genes are found in Supplementary Table 2.

Cluster (Nrps/Pks)	RlC2646	RlRL570	Ra	Rq-m	Rq	Gc	Op	On-u	Ou	**Similar known cluster** (top match shown)
GcNrps1										PR_toxin (50% of genes show similarity) *Penicillium roqueforti*
GcNrps2										
RqNrps1										
RqNrps2										
RqNrps3										
RqNrps4										
RqNrps5										
RlNrps1										
RlNrps2										
RlNrps3										
RlNrps4										
RaNrps1										
RaNrps2										Acetylaranotin (30%) *Aspergillus terreus*
OphiNrps1										
OphiNrps2										
GcPks1										
GcPks2										Emericellin (28%) *Aspergillus nidulans*
GcPks3										Asperfuranone (18%) *A. nidulans*
RqPks1										Sorbicillin (85%) *Penicillium rubens*
RqPks2										
RqPks3										Clapurines (63%) *Claviceps purpurea*
RqPks4										Equisetin (27%) *Fusarium heterosporum*
RqPks5										NG-391 (50%) *Metarhizium anisopliae*
RqPks6										Equisetin (27%) *F. heterosporum*
RqPks7										Viridicatumtoxin (31%) *P. aethiopicum*
RqPks8										Emericellin (42%) *A. nidulans*
RqPks9										
OpPKS8*										Fujikurins (50%) *F. fujikuroi* (On-u, Ou) Betaenone C / betaenone A_ (50%, 37%) *Phoma betae* (*Rq-m*, *Rq*)
OpPKS1*										
OpPKS7* OpPKS3*										Fusaric_acid (50%) *F. verticillioides*
RlPks1										Depudecin (33%) *Alternaria brassicicola*
RlPks2										Asperfuranone (36%) *A. nidulans*
RlPks3										Depudecin (33%) *A. brassicicola*
RlPks4										
RlPks5										
RlPks6										Asperfuranone_ (36%) *Aspergillus nidulans*
RlPks7										Depudecin (33%) *Alternaria brassicicola*
RaPks1										
RaPks2										
RlPks8										
RlPks9										PR_toxin_ (50%) *P. roqueforti*
RlPks10										NG-391_ (50%) *Metarhizium anisopliae*
RlPks11										
RlPks12										Asperfuranone (27%) *Aspergillus nidulans*
RlPks13										Desmethylbassianin (80%) *Beauveria bassiana*
OpPKS5*										
RlPks14										
OpPKS2*										
OpPKS10*										

* Clusters identified and named in Sbaraini et al. (2018).

[*] Clusters identified and named in Sbaraini et al. (2017).

each other, than compared to *G. clavigera* (Fig. 2C). No non-orthologous proteins were identified in *O. ulmi* when compared to *O. novo-ulmi* and *O. piceae*, although *O. novo-ulmi* had only three non-orthologous clusters, compared to 44 in *O. piceae* (Fig. 2D). *Raffaelea lauricola* C2646 had a large number of non-orthologous clusters when compared to *R. quercivora/R. quercus-mongolicae/G. clavigera* (Fig. 2E) and *O. novo-ulmi/O. ulmi/O. piceae* (Fig. 2F) with 210 and 278, respectively.

Of the annotated genes in *R. lauricola* C2646, 65% (7272) had at least one InterProScan (Pfam) match allowing classification into a gene family, and 90% (10,077 with evalues ≤ 0.05; 10,044 ≤ 0.01) had a BLASTp hit to fungal sequences in the Uniprot database. The largest number of annotated proteins was found in *R. aguacate* where 88% (10,487) had a BLASTp hit. We found 676 predicted secreted proteins in the *R. lauricola* C2646 genome, similar to *R. lauricola* RL570 and *R.*

aguacate where 654 and 671 secreted proteins, respectively, were identified. Species within this cluster had more secreted proteins than other examined species in the Ophiostomatales (Table 2). Secreted proteins are necessary to overcome host defense responses and/or scavenge food resources, and the release of secreted proteins outside the fungal cell is finely regulated (McCotter et al., 2016). Of the secreted proteins (secretome) identified in the *R. lauricola* C2646, 305 of these could be considered small-secreted proteins (SSP), compared to 293 in *R. lauricola* RL570 and 270 in *R. aguacate*. In the other ophiostomatoid species, the number of SPPs (range: 174–245) was considerably less (Table 2). SSPs are considered important factors in the interaction of fungal pathogens with their hosts, although many of them have yet to be characterized (Kim et al., 2016; de Sain and Rep, 2015). CAZYmes and proteases, may also play important roles in fungal wilt diseases (de Sain and Rep, 2015), and both proteins were found in considerable numbers among secreted members in all species examined (Table 2).

ABC transporters are considered important in fungi as key components of resistance to antifungal agents (Coleman and Mylonakis, 2009). An ABC transporter has been shown to confer resistance to monoterpenes in *G. clavigera* (Wang et al., 2012), and both *G. clavigera* and *O. piceae* were shown to express high levels of a similar ABC transporter transcript when treated with monoterpenes (Haridas et al., 2013). Besides their role in exporting xenobiotics out of fungal cells, ABC transporters are also involved in the uptake of nutrients from the environment, and the secretion of fungal compounds that interact with hosts (Perlin et al., 2014). The number of ABC transporters was again considerably greater in *R. lauricola* and *R. aguacate* compared to the other species. Although *Ophiostoma* spp. had the fewest ABC transporters, they still had more than 30 (Table 2). The same trend was observed for cytochrome P450 genes. Cytochrome P450 proteins are involved in the metabolism of diverse compounds found in fungi and plants (Chen et al., 2014). For example, cytochrome P450 enzymes have been shown to be important in the breakdown of lignin in white rot fungi (Syed and Yadav, 2012). Peroxidases and laccases are other enzymes associated with lignin degradation (Levasseur et al., 2013), and both were present in similar numbers among the eight species (Table 2). These finding are noteworthy because *R. lauricola* and other ophiostomatoid fungi are not considered wood degraders (Lundell et al., 2014). However, fungal cytochrome P450 enzymes have widely ranging catalytic activities that affect numerous substrates and processes, including the detoxification of host-defense compounds (xenobiotics) (Durairaj et al., 2016; Syed and Yadav, 2012). In *G. clavigera*, for example, certain cytochrome P450 genes were shown to be differentially induced in the presence of terpenoids and phenolic compounds (Lah et al., 2013). Similarly, cytochrome P450 enzymes in the other ophiostomatoid species evaluated here could have important functions related to the detoxification of host defense chemicals, which could enable these fungi to become established in host tissues.

Based on the versatility of the enzymatic activities of laccases, peroxidases, and tannases (Claus, 2004; Banci et al., 1999; Baik et al., 2014), we hypothesize that the presence of these enzymes in *R. lauricola* and the other species (Table 2) may also serve a primary function in detoxifying host-produced, defensive compounds, instead of degrading wood components. Evidence of this kind of activity has been shown for the three types of enzymes in fungal pathogens of grapes, which accumulate phenolic compounds in their woody vines as a defense mechanism (Bruno and Sparapano, 2006). Both laccases and peroxidases have been used to detoxify wood hydrolysates to make them less inhibitory to fermentation by yeast (Jönsson et al., 1998), and more recently, laccases have been used to detoxify various other substrates (Plácido and Capareda, 2015).

Other plant cell wall components (e.g., cellulose, hemicellulose, and pectin) and plant carbohydrates (e.g., starch and mannans) can be degraded by a diverse array of enzymes. These enzymes are categorized as glycoside hydrolases (GHs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), and auxiliary activities (AAs) (Levasseur et al., 2013; Zhao et al., 2013). Herein, these enzymes were grouped as Carbohydrate-Active Enzymes (CAZYmes) (Fig. 3). We excluded glycosyl transferases that are primarily involved in carbohydrate synthesis and carbohydrate binding modules. A larger number of genes encoding these enzymes were found in *R. lauricola* (C2464: 313, RL570: 308) and *R. aguacate* (351) when compared to the other ophiostomatoid species, which had similar numbers of these enzymes, ranging from 223 to 285 (Table 2). The distribution of the different categories of CAZYmes considered is shown in Fig. 3. *Raffaelea lauricola* (C2546 and RL570), *R. aguacate*, and *R. quercus-mongolicae* had similar numbers of AAs ranging from 68 to 61, respectively, compared to the other species. Interestingly, the putative non-pathogen *R. aguacate* had the largest number of CEs and GHs compared to the other species. The largest number of PLs were identified in the two *R. lauricola* isolates (C2646 and RL570) with 7 and 6, respectively. These enzymes, especially those that are secreted, could be responsible for obtaining nutrients from the living cells present in the cambium and sapwood of the host trees.

Numerous proteolytic enzymes (proteases) were predicted to be secreted in the Ophiostomatales species compared here: the lowest number corresponding to *O. novo-ulmi* (34) and the highest to *R. aguacate* (45) (Table 2). Proteases can play a vital role in disease development caused by phytopathogenic fungi because they can degrade plant cell wall-associated components to facilitate penetration by fungal hypha, inactivate plant chitinases and plant proteases that serve as a plant defense mechanisms, scavenge nitrogen compounds for nutrition, as well as perform other functions (Chandrasekaran et al., 2016). It has been suggested that the types of proteases secreted by fungi may reflect their life style (Valueva et al., 2016), but this remains to be investigated in the Ophiostomatales.

Polyketides and non-ribosomal peptides have been previously shown to act as virulence factors in other plant-fungi, animal-fungi, and insect-fungi interactions (Macheleidt et al., 2016). The presence of polyketide synthase and non-ribosomal peptide synthase genes was also verified in the Ophiostomatales genomes studied here using antiSMASH 4.0 (Blin et al., 2017). For non-ribosomal protein synthesis clusters (Nrps clusters), *R. lauricola* isolates had a larger number (9), especially compared to *R. aguacate*, *G. clavigera*, and *Ophiostoma* species, which had only had three, two, and one, respectively (Table 2). Most of the Nrps clusters were species-specific except two clusters (RqNrps3 and RqNrps4) were shared among *R. quercivora* and *R. quercus-mongolicae* and one cluster (OphiNrps2) was shared by *O. novo-ulmi* and *O. ulmi*. All four Nrps clusters were the same in both *R. lauricola* isolates (Table 3).

Both *R. lauricola* isolates had a larger number of polyketide synthase clusters (Pks clusters) compared to *R. aguacate*, and a similar number compared to *R. quercus-mongolicae* and *R. quercivora*. The fewest Pks clusters were observed in *O. novo-ulmi* and *O. ulmi* (Table 3) ("Terpene" and "Other" categories predicted by antiSMASH were excluded from analyses). Interestingly, several of the Pks clusters were species-specific (Table 3 and Supplementary Table 2), although some were identified within and among species. Three clusters were shared among all species (OpPKS3, RlPks1, OpPKS10); another (OpPKS7) shared by all but *O. piceae*. Ten Pks clusters were identified as unique to *R. lauricola* (C2646 and/or RL570). Several of these were found similar to secondary metabolite clusters known in other fungal species, for example RlPks9 had 50% of the genes similar to the PR toxin biosynthetic cluster identified in *Penicillium roqueforti*, a common fungus used to produce gourmet blue cheeses (Table 3).

Our results show some differences with those of Sbaraini et al. (2017) likely attributable in part to our use of different methodologies for gene annotation, including a newer version of the antiSMASH program (version 4.0 vs. version 3.0). For example, clusters OpPKS1 and OpPKS7 were not present in *O. piceae* in our analysis. It is also important to mention that the OpPKS8 fujikurin-like gene cluster that previously reported to be present only in *O. novo-ulmi* and *O. ulmi* (Sbaraini et al., 2017) was also found in *R. quercivora* and *R. quercus-*

mongolicae in our analyses. However, for *R. quercivora* and *R. quercus-mongolicae*, the OpPKS8 cluster was predicted to be a betaenone C/ betaenone A biosynthetic gene cluster characterized in *Phoma betae*. This difference can be explained by examining the genes adjacent to the main backbone genes that are also considered as part of the biosynthetic cluster (Supplementary Table 2). As mentioned in the methods, we identified the Nrps and Pks clusters either as the main backbone or including the adjacent and interspersed genes (Table 3 and Supplementary Table 2). When those adjacent and/or interspersed genes are considered, a greater diversity of Pks clusters is observed and their resulting products can vary. The polyketide synthase and non-ribosomal peptide synthase clusters present in any given species could contribute greatly to their ecological activities (Macheleidt et al., 2016; Sbaraini et al., 2017) and their significance to these fungi warrants further investigation.

Vascular wilt pathogens, including fungal pathogens, are very destructive to trees in natural and urban settings (Yadeta and Thomma, 2013). Understanding the biology and pathogenicity of the causal agents at molecular level is critical for developing management strategies for wilt diseases of trees. Overall, the genome assembly and annotation of *R. lauricola* are useful resources for continued studies on *R. lauricola* and related species. Mining into these genomes, we found that *R. lauricola* is well equipped for its role as a pathogen which likely explains its success as a highly destructive pathogen of a diverse group of lauraceous hosts indigenous to the southeastern USA.

Acknowledgements

We would like to thank Z. Abdo for computational resources, T.C. Harrington for the *R. lauricola* C2646 isolate, Ki-tae Kim for his help in the analysis of secreted proteins, and two anonymous reviewers whose comments and suggestions improved the manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fgb.2019.01.012.

References

Almagro Armenteros, J.J., Kaae Sønderby, C., Kaae Sønderby, S., Nielsen, H., Winther, O., 2017. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33 (21), 3387–3395.

Baik, J.H., Suh, H.J., Cho, S.Y., Park, Y., Choi, H.-S., 2014. Differential activities of fungi-derived tannases on biotransformation and substrate inhibition in green tea extract. J. Biosci. Bioeng. 118 (5), 546–553.

Banci, L., Ciofi-Baffoni, S., Tien, M., 1999. Lignin and Mn peroxidase-catalyzes oxidation of phenolic lignin oligomers. Biochemistry 38, 3205–3210.

Barton, C., Bates, C., Cutrer, B., Eickwort, J., Harrington, S., Jenkins, D., McReynolds, Stones D., Reid, L., Riggins, J.J., Trickel, R., 2016. Distribution of counties with laurel wilt disease by year of initial detection. Available online. <https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/fseprd523011.pdf>.

Bertels, F., Silander, O.K., Pachkov, M., Rainey, P.B., van Nimwegen, E., 2014. Automated reconstruction of whole genome phylogenies from short sequence reads. Mol. Biol. Evol. 31 (5), 1077–1088.

Blin, K., Wolf, T., Chevrette, M.G., Lu, X., Schwalen, C.J., Kautsar, S.A., Suarez Duran, H.G., de los Santos, E.L.C., Kim, H.U., Nave, M., Dickschat, J.S., Mitchell, D.A., Shelest, E., Breitling, R., Takano, E., Lee, S.Y., Weber, T., Mederma, M.H., 2017. antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. Nucl. Acids Res. 45 (Web Server issue), W36–W41.

Bruno, G., Sparapano, L., 2006. Effects of three esca-associated fungi on *Vitis vinifera* L.: III. Enzymes produced by the pathogens and their role in fungus-to-plant or in fungus-to-fungus interactions. Physiol. Mol. Plant Pathol. 69, 182–194.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinf. 10, 421–429.

Cantarel, B.L., Korf, I., Robb, S.M.C., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez Alvarado, A., Yandell, M., 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18 (1), 188–196.

Chandrasekaran, M., Thangavelu, B., Chun, S.C., Sathiyabama, M., 2016. Proteases from phytopathogenic fungi and their importance in phytopatogenicity. J. Gen. Plant Pathol. 82, 233–239.

Chen, W., Lee, M., Jefcoate, C., Kim, S., Chen, F., Yu, J., 2014. Fungal cytochrome P450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin. Genome Biol. Evolut. 6 (7), 1620–1634.

Choi, J., Détry, N., Kim, K.T., Asiegbu, F.O., Valkonen, J.P., Lee, Y.H., 2014. fPoxDB: fungal peroxidase database for comparative genomics. BMC Microbiol. 14, 117.

Claus, H., 2004. Laccases: structure, reactions, distribution. Micron 35, 93–96.

Coil, D., Jospin, G., Darling, A.E., 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. Bioinformatics 31 (4), 587–589.

Coleman, J.J., Mylonakis, E., 2009. Efflux in fungi: La pièce de résistance. PLoS Pathog. 5 (6), e1000486. https://doi.org/10.1371/journal.ppat.1000486.

Comeau, A.M., Dufour, J., Bouvet, G.F., Jacobi, V., Nigg, M., Henrissat, B., Laroche, J., Levesque, R., Bernier, L., 2014. Functional annotation of the *Ophiostoma novo-ulmi* genome: Insights into the phytopathogenicity of the fungal agent of Dutch elm disease. Genome Biol. Evolut. 7 (2), 410–430.

de Beer, Z.W., Wingfield, M.J., 2013. Emerging lineages in the Ophiostomatales. In: Seifert, K.A., de Beer, Z.W., Wingfield, M.J. (Eds.), Ophiostomatoid Fungi. CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands, pp. 21–46.

De Sain, M., Rep, M., 2015. The role of pathogen-secreted proteins in fungal vascular wilt diseases. Int. J. Mol. Sci. 16, 23970–23993.

DiGuistini, S., Wang, Y., Liao, N.Y., Taylor, G., Tanguay, P., Feau, N., Henrissat, B., Chan, S.K., Hesse-Orce, U., Alamouti, S.M., Tsui, C.K., Docking, R.T., Levasseur, A., Haridas, S., Robertson, G., Birol, I., Holt, R.A., Marra, M.A., Hamelin, R.C., Hirst, M., Jones, S.J., Bohlmann, J., Breuil, C., 2011. Genome and transcriptome analyses of the mountain pine beetle-fungal symbiont *Grosmannia clavigera*, a lodgepole pine pathogen. Proc. Natl. Acad. Sci. USA 108 (6), 2504–2509.

Dreaden, T.J., Davis, J.M., de Beer, Z.W., Ploetz, R.C., Soltis, P.S., Wingfield, M.J., Smith, J.A., 2014. Phylogeny of ambrosia beetle symbionts in the genus *Raffaelea*. Fungal Biol. 118 (12), 970–978.

Durairaj, P., Hur, J.-S., Yun, H., 2016. Versatile biocatalysis of fungal cytochrome P450 monooxygenases. Microb. Cell Fact. 15, 125–140.

Durkovic, J., Kacil, F., Olcak, D., Kucerova, V., Krajnakova, J., 2014. Host responses and metabolic profiles of wood components in Dutch elm hybrids with contrasting tolerance to Dutch elm disease. Ann. Bot. 114, 47–59.

Fraedrich, S.W., Harrington, T.C., Bates, C.A., Johnson, J., Reid, L.S., Best, G.S., Leininger, T.D., Hawkins, T.S., 2011. Susceptibility to laurel wilt and disease incidence in two rare plant species, pondberry and pondspice. Plant Dis. 95, 1056–1062.

Fraedrich, S.W., Harrington, T.C., McDaniel, B.A., Best, G.S., 2016. First report of laurel wilt, caused by *Raffaelea lauricola*, on spicebush (*Lindera benzoin*) in South Carolina. Plant Dis. 100 (11), 2330.

Fraedrich, S.W., Harrington, T.C., Rabaglia, R.J., Ulyshen, M.D., Mayfield III, A.E., Hanula, J.L., Eickwort, J.M., Miller, D.R., 2008. A fungal symbiont of the redbay ambrosia beetle causes a lethal wilt in redbay and other Lauraceae in the southeastern United States. Plant Dis. 92, 215–224.

Fraedrich, S.W., Johnson, C.W., Menard, R.D., Harrington, T.C., Olatinwo, R., Best, G.S., 2015. First Report of *Xyleborus glabratus* (Coleoptera: Curculionidae: Scolytinae) and laurel wilt in Louisiana, USA: the disease continues westward on sassafras. Florida Entomologist 98 (4), 1266–1268.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate Maximum-Likelihood phylogenies: assessing the performance of PhyML 3.0. Systemat. Biol. 59 (3), 307–321.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29 (8), 1072–1075.

Haridas, S., Wang, Y., Lim, L., Alamouti, S.M., Jackman, S., Docking, R., Robertson, G., Birol, I., Bohlmann, J., Breuil, C., 2013. The genome and transcriptome of the pine saprophyte *Ophiostoma piceae*, and comparison with the bark beetle-associated pine pathogen *Grosmannia clavigera*. BMC Genomics 14, 373–377.

Harrington, T.C., 2005. Ecology and evolution of mycophagous bark beetles and their fungal partners. In: Vega, F.E., Blackwell, M. (Eds.), Ecological and Evolutionary Advances in Insect-Fungal Associations. Oxford University Press, pp. 257–291.

Harrington, T.C., Fraedrich, S.W., Aghayeva, D.N., 2008. *Raffaelea lauricola*, a new ambrosia beetle symbiont and pathogen on the Lauraceae. Mycotaxon 104, 399–404.

Harrington, T.C., Yun, H.Y., Lu, S., Goto, H., Aghayeva, D.N., Fraedrich, S.W., 2011. Isolations from the redbay ambrosia beetle, *Xyleborus glabratus*, confirm that the laurel wilt pathogen, *Raffaelea lauricola*, originated in Asia. Mycologia 103 (5), 1028–1036.

Hughes, M.A., Black, A., Smith, J.A., 2014. First report of laurel wilt caused by *Raffaelea lauricola* on bay laurel (*Laurus nobilis*) in the United States. Plant Dis. 98, 1159.

Hughes, M.A., Shin, K., Eickwort, J., Smith, J.A., 2012. First report of laurel wilt disease caused by *Raffaelea lauricola* on silk bay in Florida. Plant Dis. 96 (6), 910.

Hughes, M., Smith, J.A., Mayfield, A.E., Minno, M.C., Shin, K., 2011. First report of laurel wilt disease caused by *Raffaelea lauricola* on pondspice in Florida. Plant Dis. 95 (12), 1588.

Inch, S., Ploetz, R., Held, B., Blanchette, R., 2012. Histological and anatomical responses in avocado, *Persea americana*, induced by the vascular wilt pathogen *Raffaelea lauricola*. Botany 90, 627–635.

Jeon, J., Kim, K.-T., Song, H., Lee, G.-W., Cheong, K., Kim, H., Choi, G., Lee, Y.-H., Stewart, J.E., Klopfenstein, N.B., Kim, M.-S., 2017. Draft genome sequence of the fungus associated with oak-wilt mortality in South Korea, *Raffaelea quercus-mongolicae* KACC44405. Genome Announcements 5 (34), e00797–e817.

Jones, P., Binns, D., Chang, H., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30 (9), 1236–1240.

Jönsson, L.J., Palmqvist, E., Nilvebrant, N.-O., Hahn-Hägerdal, B., 1998. Detoxification of wood hydrolysates with laccase and peroxidase from the white-rot fungus *Trametes versicolor*. Appl. Microbiol. Biotechnol. 49, 691–697.

Keller, O., Kollmar, M., Stanke, M., Waack, S., 2011. A novel hybrid gene prediction

method employing protein multiple sequence alignments. Bioinformatics 27 (6), 757–763.
Kim, K.T., Jeon, J., Choi, J., Cheong, K., Song, H., Choi, G., Kang, S., Lee, Y.H., 2016. Kingdom-wide analysis of fungal small secreted proteins (SSPs) reveals their potential role in host association. Front. Plant Sci. 7, 186.
Khoshraftar, S., Hung, S., Khan, S., Gong, Y., Tyagi, V., Parkinson, J., Sain, M., Moses, A.M., Christendat, D., 2013. Sequencing and annotation of the *Ophiostoma ulmi* genome. BMC Genomics 14, 162–172.
Kubono, T., Ito, S., 2002. *Raffaelea quercivora* sp. nov. associated with mass mortality of Japanese oak, and the ambrosia beetle (*Platypus quercivorus*). Mycoscience 43, 255–260.
Lah, L., Haridas, S., Bohlmann, J., Breuil, C., 2013. The cytochromes P450 of *Grosmannia clavigera*: Genome organization, phylogeny, and expression in response to host chemicals. Fungal Genet. Biol. 50, 72–81.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.
Lee, S., Kim, S.H., Breuil, C., 2002. The use of the green fluorescent protein as a biomarker for sapstain fungi. Forest Pathol. 32 (3), 153–161.
Levasseur, A., Drula, E., Lombard, V., Coutinho, P.M., Henrissat, B., 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol. Biofuels 6, 41–54. https://doi.org/10.1186/1754-6834-6-41.
Li, H.-X., Gotilla, T.M., Brewer, M.T., 2017. Organization and evolution of mating–type genes in three *Stagonosporopsis* species causing gummy stem blight of cucurbits and leaf spot and dry rot of papaya. Fungal Biol. 121, 849–857.
Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl. Acids Res. 25, 955–964.
Lundell, T.K., Mäkelä, M.R., de Vries, R.P., Hildén, K.S., 2014. Genomics, lifestyle and future prospects of wood-decay and litter-decomposing basidiomycota. Adv. Bot. Res. 70, 329–370.
Macheleidt, J., Mattern, D.J., Fischer, J., Netzker, T., Weber, J., Schroeckh, V., Valiantem, V., Brakhage, A.A., 2016. Regulation and role of fungal secondary metabolites. Annu. Rev. Genet. 50, 371–392. https://doi.org/10.1146/annurev-genet-120215-035203.
Mayfield, A.E., Smith, J.A., Hughes, M., Dreaden, T.J., 2008. First report of laurel wilt disease caused by a *Raffaelea* sp. on avocado in Florida. Plant Dis. 92 (6), 976.
McCotter, S.W., Horianopoulos, L.C., Kronstad, J.W., 2016. Regulation of the fungal secretome. Curr. Genet. 62 (3), 533–545.
Park, J., Lee, S., Choi, J., Ahn, K., Park, B., Park, J., Kang, S., Lee, Y.H., 2008. Fungal cytochrome P450 database. BMC Genomics 9, 402.
Perlin, M.H., Andrews, J., Toh, S.S., 2014. Essential letters in the fungal alphabet: ABC and MFS transporters and their roles in survival and pathogenicity. Adv. Genet. 85, 201–253.
Plácido, J., Capareda, S., 2015. Ligninolytic enzymes: a biotechnological alternative for bioethanol production. Bioresour. Bioprocess. 2, 23–34.
Ploetz, R.C., Konkol, J.L., Narvaez, T., Duncan, R.E., Saucedo, R.J., Campbell, A., Mantilla, J., Carrillo, D., Kendra, P.E., 2017. Presence and prevalence of *Raffaelea lauricola*, cause of laurel wilt, in different species of ambrosia beetle in Florida, USA. J. Econ. Entomol. 110 (2), 347–354.
Saucedo-Carabez, J.R., Ploetz, R.C., Konkol, J.L., Carrillo, D., Gazis, R., 2018. Partnerships between ambrosia beetles and fungi: lineage-specific promiscuity among vectors of the laurel wilt pathogen. Microb. Ecol. 76, 925–940.
Sbaraini, N., Andreis, F.C., Thompson, C.E., Guedes, R.L.M., Junges, Â., Campos, T., Staats, C.C., Vainstein, M.H., Ribeiro de Vasconcelos, A.T., Schrank, A., 2017. Genome-Wide Analysis of Secondary Metabolite Gene Clusters in *Ophiostoma ulmi* and *Ophiostoma novo-ulmi* Reveals a Fujikurin-Like Gene Cluster with a Putative Role in Infection. Front. Microbiol. 8, 1063–1075.
Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31 (19), 3210–3212.
Smit, A.F.A., Hubley, R., Green, P., 1996. RepeatMasker at <http://repeatmasker.org>.
Smit, A.F.A., Hubley, R., 2017. RepeatModeler at <http://repeatmasker.org>.
Smith, J.A., Dreaden, T.J., Myfield, A.E., Boone, A., Fraedrich, S.W., Bates, C., 2009a. First report of laurel wilt disease Caused by *Raffaelea lauricola* on sassafras in Florida and South Carolina. Plant Dis. 93 (10), 1079.
Smith, J.A., Mount, L., Mayfield, A.E., Bates, C.A., Lamborn, A., Fraedich, S.W., 2009b. First report of laurel wilt disease caused by *Raffaelea lauricola* on camphor in Florida and Georgia. Plant Dis. 93 (2), 198.
Syed, K., Yadav, J.S., 2012. P450 monooxygenases (P450ome) of the model white rot fungus *Phanerochaete chrysosporium*. Crit. Rev. Microbiol. 38 (4), 339–363.
Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y., Borodovsky, M., 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res. 18, 1979–1990.
Valueva, T.A., Zaichik, B.T., Kudryavtseva, N.N., 2016. Role of proteolytic enzymes in the interaction of phytopathogenic microorganisms with plants. Biochemestry (Moscow) 81 (13), 1709–1718.
Vanderpool, D., Bracewell, R.R., McCutcheon, J.P., 2018. Know your farmer: Ancient origins and multiple independent domestications of ambrosia beetle fungal cultivars. Mol. Ecol. 43, 2077–2094.
Wang, Y., Coleman-Derr, D., Chen, G., Gu, Y.Q., 2015. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. Nucl. Acids Res. 43 (W1), W78–W84.
Wang, Y., Lim, L., DiGuistini, S., Robertson, G., Bohlmann, J., Breuil, C., 2012. A specialized ABC efflux transporter GcABC-G1 confers monoterpene resistance to *Grosmannia clavigera*, a bark beetle-associated fungal pathogen of pine trees. New Phytol. 197, 886–898.
Yadeta, K.A., Thomma, B.P.H.J., 2013. The xylem as battleground for plant hosts and vascular wilt pathogens. Front. Plant Sci. 4: article, 97.
Yin, Y., Mao, X., Yang, J.C., Chen, X., Mao, F., Xu, Y., 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucl. Acids Res. 40 (W1), W445–W451.
Zaharia, M., Bolosky, W.J., Curtis, K., Fox, A., Patterson, D., Shenker, S., Stoica, I., Karp, R.M., Sittler, T., 2011. Faster and More Accurate Sequence Alignment with SNAP. arXiv: 1111.5572.
Zhao, Z., Liu, H., Wang, C., Xu, J., 2013. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. BMC Genomics 14, 274–288.