

Transcriptome analysis of a powdery mildew pathogen (*Podosphaera pannosa*) infecting *Eucalyptus urophylla*: De novo assembly, expression profiling and secretome prediction

Natália R. Fonseca¹ | Jorge Ibarra Caballero² | Mee-Sook Kim³ | Jane E. Stewart²  |
Lúcio M. S. Guimarães⁴ | Acelino C. Alfenas⁴  | Ned B. Klopfenstein⁵ 

¹Instituto Federal de Minas Gerais – Campus São João Evangelista, São João Evangelista, Brazil

²Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, Colorado

³USDA Forest Service - Pacific Northwest Research Station, Corvallis, Oregon

⁴Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, Brazil

⁵USDA Forest Service - Rocky Mountain Research Station, Moscow, Idaho

Correspondence

Natália R. Fonseca, Instituto Federal de Minas Gerais – Campus São João Evangelista, São João Evangelista, Brazil.
Email: natalia.fonseca@ifmg.edu.br

and
Jane E. Stewart, Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO.
Email: jane.stewart@colostate.edu

Funding information

USDA Forest Service Rocky Mountain Research Station; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

Editor: Carl Gunnar Fossdal

Abstract

Podosphaera pannosa is the causal agent of powdery mildew on eucalypt in Brazil. This powdery mildew disease is important in nurseries causing leaf and shoot distortion, shoot discoloration and reduction in growth, which decreases mini-cutting production. Improved RNA sequencing (RNA-Seq) technologies have allowed increased information about the transcriptome of several pathogens and hosts, enabling a better understanding of their interaction at the gene level. For this study, we analysed the transcriptome of *P. pannosa* during leaf infection of *Eucalyptus urophylla* leaves using RNA-Seq and de novo transcriptome assembly. The transcriptome was Illumina sequenced and assembled de novo, generating over 178 million RNA-Seq reads assembled onto 200,473 contigs. After filtering steps, the resulting 12,106 (6%) transcripts were identified as the *P. pannosa* transcriptome data set. The 10 most abundant transcripts included genes encoding enzymes likely involved in fungal establishment and growth, such as dihydrofolate reductase, putative methyltransferases, acyl-desaturase, glycoside hydrolase and dehydrogenases. In addition, genes putatively encoding an aquaporin and an orthologue to the effector protein, GoEC2 of *Golovinomyces orontii* were identified. The predicted secretome consisted of 1,899 translated transcripts, of which 310 exhibited homology to proteins described in the PHI database, 144 of these showing homology to fungal PHI accessions that affected pathogenicity or that are described as effectors. In addition, 81 transcripts encoded secreted proteins homologous to effectors described in the Erysiphales. These results provide a basis for continued studies to better understand the *P. pannosa*-eucalypt (*Eucalyptus* spp.) pathosystem and could parallel studies of the eucalypt transcriptome to help determine host resistance mechanisms.

KEYWORDS

biotroph, leaf disease, microsatellites, secreted proteins

1 | INTRODUCTION

Powdery mildews are diseases caused by a group of biotrophic, plant pathogenic fungi belonging to the Erysiphales. These diseases are easily recognized by superficial white powdery patches or film produced by anamorphic mycelia, conidiophores, and conidia affecting leaves, stems, flowers and fruits of almost 10,000 host species (Amano, 1986; Braun & Cook, 2012). On eucalypt (*Eucalyptus* spp.), powdery mildew is a nursery disease causing leaf and shoot distortion, shoot discoloration and reduction in growth, which consequently reduces mini-cutting production (Alfenas, Zauza, Mafia, & Assis, 2009; Keane, Kile, Podger, & Brown, 2000). The eucalypt powdery mildew pathogen in Brazil was traditionally identified as *Oidium eucalypti* Rostr. (1902) or generically as *Oidium* sp. (Alfenas et al., 2009) based on its anamorph morphology. Recently, studies based on ITS and 28S rDNA sequences concluded that *Podosphaera pannosa* (Wallr.:Fr.) de Bary (1870) was the cause of powdery mildew on eucalypt in Brazil (Fonseca, Guimaraes, Pires, Klopfenstein, & Alfenas, 2017). However, studies on the powdery mildew pathosystem are limited, and definitive information is needed to develop technologies for managing the disease.

Podosphaera pannosa is a pathogen with a broad host range, including *Rosa* spp. (Rosaceae), *Prunus* spp. (Rosaceae), *Cotinus coggygria* (Anacardiaceae), *Forsythia* spp. (Oleaceae), *Eucalyptus* spp. (Myrtaceae) (Braun & Cook, 2012) and *Corymbia citriodora* (Myrtaceae) (Cho, Park, Lee, Lee, & Shin, 2016). On eucalypt, *P. pannosa* has been reported in several countries including Argentina, Brazil, Australia, Denmark, Italy, New Zealand, Poland, Portugal, United Kingdom, South Africa and recently South Korea (Cho et al., 2016; Old, Wingfield, & Qing Yuan, 2003).

Like other obligate/biotrophic fungal pathogens, *P. pannosa* cannot be cultured on artificial media. These pathogens live in intimate cellular contact with the host plants via haustoria and depend on an ability to secrete molecules into host cells that manipulate host physiology and defence responses, which allow the pathogen to avoid host recognition (Godfrey et al., 2010). While necrotrophs and hemibiotrophs mainly secrete secondary metabolites and cell-wall-degrading enzymes during their necrotrophic phase (Kemen, Agler, & Kemen, 2015), obligate biotrophs possess large repertoires of effector proteins, as has been demonstrated for the barley powdery mildew pathogen, *Blumeria graminis* (Godfrey et al., 2010; Spanu et al., 2010) and rust fungi (Bruce et al., 2014; Lorrain, Hecker, & Duplessis, 2015). Identification of these effector proteins is a fundamental step towards developing disease management strategies.

Whole transcriptome sequencing using next-generation sequencing technologies or RNA Sequencing (RNA-Seq) has increased the quality and utility of transcriptome analysis through the sequencing of entire transcriptome of an organism under a given condition. RNA-Seq is a powerful and relatively cost-effective, high-throughput sequencing method that uses deep sequencing to produce millions of short-sequence reads. These reads can be aligned to a reference genome when available or assembled de novo without the genomic reference to produce a genome-scale transcription map comprising both

the transcriptional profile and/or level of expression for each gene (Wang, Gerstein, & Snyder, 2009). For non-model organisms, such as *P. pannosa*, the use of RNA-Seq is a practical approach, because a reference genome is not required. Whole-genome sequencing and assembly for powdery mildew pathogens are hampered by the obligate nature of the pathogen and by large numbers of repetitive elements throughout the genomes (Spanu et al., 2010). Transcriptome sequencing by RNA-Seq provides fundamental information for gene discovery and quantification of gene expression (Huynh, Page, Richardson, & Udall, 2015; Kim et al., 2014), comparative genomic studies (Bindschedler, Panstruga, & Spanu, 2016; Pérez, Romero, Restrepo, Gutiérrez, & Castaño, 2015; Wang, Shi, & Rinehart, 2015), secretome analysis and prediction of fungal candidate effectors (Bruce et al., 2014; Guyon, Balagué, Roby, & Raffaele, 2014; Liu et al., 2015; Meinhardt et al., 2014). Thus, the aim of our study was to develop genomic resources, through transcriptome profiling of epiphytic hyphae/conidia and haustoria and identification of potential microsatellites for population genetic studies of *P. pannosa*, infecting an economically important tree crop *Eucalyptus urophylla*.

2 | MATERIALS AND METHODS

2.1 | Fungal and plant material

To obtain powdery mildew pathogen materials for RNA extraction, 10 *E. urophylla* plants (clone 1,183) were prepared for inoculation. Ninety-day-old cuttings were transplanted to 2-L pots containing the commercial substrate MecPlant® (MECPREC commercial industries, Telêmaco Borba, PR, Brazil) supplemented with 3 kg/m³ of superphosphate [Ca (H₂PO₄)₂] and 1.5 kg/m³ of Osmocote® (19-6-12; The Scotts Co., Marysville, OH, USA). Plants were maintained under greenhouse conditions and inoculated 30 days after transplanting. Isolate LPF 615 of the eucalypt powdery mildew pathogen, collected in Viçosa, MG, Brazil and maintained on eucalypt cuttings in a growth chamber free of other inoculum sources, was used for inoculations. Inoculations were performed with a small soft brush by dusting conidia from an infected eucalypt leaf onto newly expanded leaves of *E. urophylla* plants. Inoculated plants were placed in growth chamber at 19 ± 2°C with a 12-hr photoperiod and light intensity of 40 µmol s⁻¹ m⁻². Four weeks later, *P. pannosa*-infected *E. urophylla* leaves were collected, and mycelia, conidia and haustoria within plant tissues were removed by gentle scraping, and immediately placed into 2.0-ml microcentrifuge tubes containing 1.0 ml of RNeasy® Stabilization Solution (ThermoFisher Scientific, NY, USA). The tubes were sealed and maintained at -80°C until RNA extraction.

2.2 | RNA extraction

Total RNA was extracted from *P. pannosa*-infected eucalypt leaves using the ZR Fungal/Bacterial RNA Mini Prep Kit (Zymo Research, Orange, CA, USA) with the following modifications: 1 ml sterilized MiliQ water was added to tubes containing RNeasy®, followed by vortexing and centrifugation at 20,800×g. The supernatant was

discarded and the pellet resuspended in 800 µl RNA Lysis Buffer; the mixture was divided between two 2.0-ml FastPrep tubes (MP Biomedicals, Solon, Ohio, USA) containing 0.18 g of garnet matrix (MP Biomedicals) and two 6.25-mm ceramic spheres (MP Biomedicals). The tubes were processed in a FastPrep FP120 cell disrupter (Thermo Savant; Holbrook, NY, USA) at speed 5.5 for 30 s; tubes were placed on ice for 5 min and processed again using the same methods and conditions. All centrifugation steps were performed at 15,300×g (centrifuge model S417R; Eppendorf, Hauppauge, NY, USA). RNA was eluted in 15 µl DNase-/RNase-free water. RNA concentration was measured using a NanoDrop 2,000 Spectrophotometer (ThermoFisher Scientific), and RNA integrity number (RIN) was checked using Agilent RNA Screen Tape System (Agilent Technologies, Germany) in a 2,200 TapeStation (Agilent Technologies).

2.3 | Illumina sequencing and de novo assembly

Illumina next-generation sequencing was performed at Macrogen Korea (Seoul, Republic of Korea). First, the rRNA in total RNA was depleted using a Ribo-Zero kit (Illumina Inc., San Diego, CA, USA). The enriched mRNA samples were subjected to Illumina cDNA library construction using TruSeq stranded mRNA (Microbe) kit (Illumina Inc.). The RNA was purified, fragmented and primed for cDNA synthesis. The RNA fragments were transcribed into first-strand cDNA using reverse transcriptase and random hexamers, followed by second-strand cDNA synthesis. These fragments were prepared for sequencing with an end-repair process and addition of a single "A" base at the 3' end. Paired-end adapters were ligated to the ends of these 3' adenylated cDNA fragments. Products were purified and enriched using PCR to create the final cDNA library. The cDNA library was sequenced using Illumina HiSeqTM 2,000 (Illumina Inc.) producing a read length of 101 bp.

The resulting sequence reads were subjected to a quality control check using FastQC. Sequences were trimmed, and the adapters removed using Trimmomatic V0.32 with settings at ILLUMINACLIP: TruSeq3-PE-2.fa:2:151:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. De novo transcriptome assembly of the sequence reads was performed using the software package, Trinity (version r20140717), a de Bruijn graph-based assembler (Grabherr et al., 2011). Assembled sequences with a minimum length of 201 bp were recorded as a contig.

2.4 | Identification of *P. pannosa* transcripts and functional annotation

Assembled contigs were utilized for similarity searches against the NCBI non-redundant (nr) protein database (<http://www.ncbi.nlm.nih.gov/>) with BLASTX algorithm with an *e*-value $\leq 1e^{-5}$. Because RNA was extracted from fungal tissue removed from infected eucalypt leaves, it was expected that eucalypt, insect and/or other microbial transcripts would be present in the samples. To remove transcripts that were not produced by *P. pannosa*, the results were filtered based on BLASTX searches and only those transcripts

matching previously described proteins from Erysiphales were used for further analysis of the *P. pannosa* transcriptome. Gene annotation on Gene Ontology Functional Classification System was performed using the Blast2GO program (Conesa et al., 2005), which retrieved Gene Ontology (GO) terms in the categories, molecular function (MF), biological process (BP) and cellular component (CC) assigned to the powdery mildew pathogen genes based on their homologies to the NCBI nr database.

Transcript abundance was estimated using RSEM 1.2.15 (RNA-Seq by Expectation-Maximization), an accurate software tool for quantifying transcript abundance from RNA-Seq data (Li & Dewey, 2011). Since the results were produced in a unit of a transcript, normalized expression values (i.e. FPKM: Fragments Per Kilobase of transcript per Million mapped reads) were used to quantify the expression (relative abundance) of each contig.

2.5 | Prediction of secretory proteins

Secreted proteins were predicted using DeepLoc-1.0. This software uses "deep neural networks" to predict protein subcellular localization based only on the protein sequence information (Almagro Armenteros, Kaae Sønderby, Kaae Sønderby, Nielsen, & Winther, 2017). The proteome (i.e. set of proteins encoded within transcriptome) was obtained by identifying the longest open reading frame of each transcript using the "getorf" subcommand in Hmmer2Go 0.17.8 (Staton, 2018), setting the lower limit to 50 aa.

2.6 | Prediction of pathogenicity proteins

To identify proteins involved in pathogenicity, the proteome was also used as a query for BLASTP (*e*-value ≤ 0.05) search against the pathogen-host interaction database (PHI-base v4.5) that catalogues experimentally verified pathogenicity, virulence and effector genes from fungal, oomycete and bacterial pathogens (Winnenburg et al., 2008).

Searches for transcripts encoding proteins described as putative effectors were also performed using BLASTP (*e*-value ≤ 0.05) of the proteome against a database of Erysiphales proteins obtained from the NCBI nr database.

2.7 | Screening for simple sequence repeat loci

Searches for polymorphic tandem repeat [i.e. microsatellites/simple sequence repeats (SSR)] motif regions from the contig data set of *P. pannosa* were performed using the Batch Primer3 web software [<http://probes.pw.usda.gov/cgi-bin/batchprimer3/batchprimer3.cgi>]. The same web server was used to design primer pairs. The pattern types screened were perfect tri-, tetra-, penta- and hexa-nucleotide motifs with minimum number of SSR patterns repeat set for three, four, five and six, respectively. Primer design parameters were set as follows: PCR product size range = 100 to 300 bp with 150 bp as optimum; primer size range = 18 to 23 nucleotides, with 21 as optimum; annealing temperature (*T*_m) optimum = 55°C; and

TABLE 1 Basic statistics of de novo transcriptome assembly in *Podosphaera pannosa*

	All transcript contigs	Only longest isoform per "gene"
Total Trinity transcripts	200,473	180,413
N50 (b)	586	497
Maximum contig length (b)	13,361	13,361
Minimum contig length (b)	201	201
Median contig length (b)	310	300
Average contig length (b)	493.7	453.5
Total assembled bases	98,973,062	81,813,563

GC content ranging from 40% to 60%, with 50% as optimum, accordingly with default parameters.

3 | RESULTS

3.1 | Illumina sequencing and de novo assembly

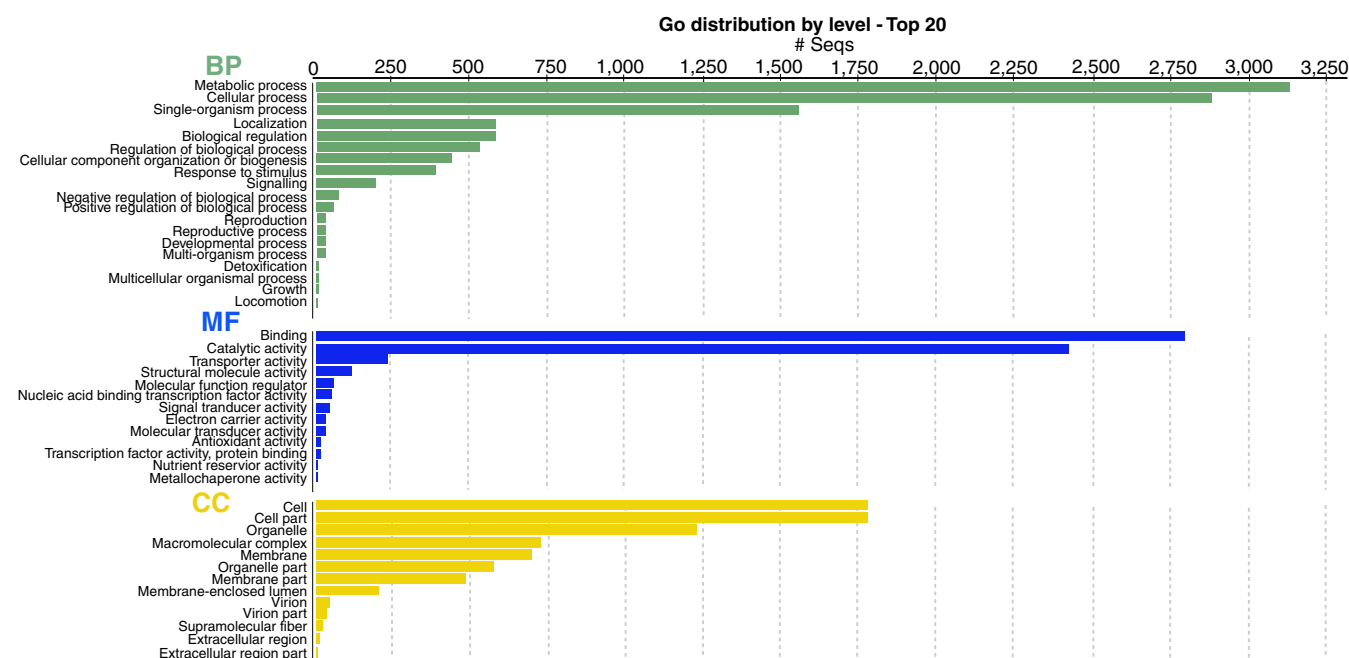
In total, we obtained 185,128,706 reads containing 18,697,999,306 bp from the cDNA library. After adaptor sequences and low-quality reads were removed, over 178×10^6 RNA-Seq ranging from 36 to 101 bp, paired-end reads remained with a Q30 of 93.6% and 49.3% GC. The de novo assembly produced 200,473 contigs, which were assembled from 98,973,062 bases with a median contig length of 493.7 bp and an N50 metric of 586 bp. Considering only the longest isoform per gene, 180,413 contigs were identified with a median contig length of 300 bp and an N50 metric of 497 bp (Table 1).

3.2 | Identification of *P. pannosa* transcripts and functional annotation

Among all assembled transcripts, 146,546 (73.1%) were successfully annotated using BLASTX algorithm against NCBI nr protein database with an e-value threshold of $1e^{-5}$. The 53,927 (26.9%) transcripts with no hits to nr database were saved as a separate library for future investigations. After annotation, transcripts were subjected to a manual filtering to retrieve only the most reliable pathogen transcripts. Of the 146,546 annotated transcripts, 12,106 (8.3%) transcripts were retrieved with significant (>95%) similarity to proteins from Erysiphales (i.e. *B. graminis*, *Erysiphe necator*, *G. orontii*, *P. macularis* and *P. fusca*). The resulting transcript data set was considered as the *P. pannosa* transcriptome for further analysis. Reference transcriptome of *P. pannosa* (12,106 contigs > 200 nt) was submitted to the NCBI Transcriptome Shotgun Assembly (TSA) database under the accession number GHDE000000000 (BioProject PRJNA474547).

Gene annotation on Gene Ontology Functional Classification System (GO terms) resulted in 42.8% of total transcripts (5,181 out of 12,106) successfully annotated, 23.6% (2,851 of 12,106) of which were assigned to at least one GO term. Of the total 12,106 transcripts, 5,101 (42.1%) had similarity with proteins defined only as hypothetical. Within contigs that had significant similarity to Erysiphales, a BLASTX top-hit species distribution of gene annotations showed highest similarity to *B. graminis* (51.3%), followed by *E. necator* (48.1%). In addition, 42 (0.35%) transcripts had homology with *G. orontii* protein-encoding sequences, and 16 (0.13%) assembled transcripts aligned with *P. macularis* and *P. fusca* protein-encoding sequences.

Three GO annotation categories are provided by the GO database: biological process (BP), molecular function (MF) and cellular

**FIGURE 1** Summary of Top 20 functional annotations of *Podosphaera pannosa* transcriptome by Gene Ontology (GO) terms divided into three main domains, biological process (BP), molecular function (MF) and cellular component (CC)

component (CC). In this study, the distribution among these three GO domains was sorted based on level 2 classification. BP domains comprised the majority of assignments (9,755; 80.6%), with metabolic process (3,187; 26.3%), cellular process (2,889; 23.9%) and single-organism process (1,554; 12.8%) representing the most abundant categories. MF domains were prominently represented by binding (2,801; 23.1%) and catalytic activity (2,432; 20%). GO terms with most assigned sequences under the domains of CC (7,255; 60%) were cell (1,783; 14.7%) and cell parts (1,770; 14.6%) (Figure 1).

To quantify the expression of each transcript, the reads were assembled back to the contigs and reported FPKM values. Of all 12,106 transcripts, the most abundant expressed gene for the *P. pannosa* transcriptome was homologous to dihydrofolate reductase (DHFR) (c169657_g1_i1), an enzyme found in many organisms. Other expressed enzyme-encoding genes found on the most expressed transcripts of the *P. pannosa* transcriptome were genes encoding transferase, hydrolase and enzymes belonging to oxidoreductase group, such as desaturase and dehydrogenase. A highly expressed transcript (c11058_g1_i1), homologous to effector protein EC2 of *G. orontii*, was also identified in this study (Table 2).

3.3 | Secretome and effector prediction

Analyses of the *P. pannosa* transcriptome were performed to identify potential secretory proteins and effectors (Figure 2). The search for extracellular proteins using DeepLoc resulted in 1,899 transcripts coding for putative secreted proteins. A total of 3,195 proteins had a match in the PHI-base, but only 310 were predicted as extracellular. Of those predicted as extracellular, 171 were reported as “effector_plant avirulence determinant” or “reduced virulence” or “loss

of pathogenicity” (the rest were “unaffected pathogenicity,” “lethal” or “increased virulence”), and of those, 144 were reported in fungi, while the remaining ca. 27 included mostly plant pathogenic bacteria (Supporting Information Table S1). The BLASTP of the proteome against a database of Erysiphales proteins resulted in 263 proteins with highest similarity to putative effectors, including many that were assigned to the EKA (Effectors homologous to the *Avr k 1* and *Avr a 10* avirulence) gene family (Amselem et al., 2015), and of these, 81 were also predicted as extracellular (Supporting Information Table S2).

Six of the putative effectors found after the BLASTP search had significant similarity to proteins in the final set of 144 PHI-base accessions mentioned above. Four of these proteins were previously described in *Magnaporthe oryzae*, one in *Botrytis cinerea*, and one in *Cryptococcus neoformans* (Table 3).

3.4 | Microsatellite marker identification and testing

All 12,106 contigs were used to mine potential SSRs defined as tri- to hexa-nucleotide motifs with a minimum of four repeats for tri-nucleotide motifs and three repeats for tetra- to hexa-nucleotide motifs. A total of 1,571 putative SSRs were identified, and of these, 1,389 (88.3%) had sufficient flanking sequences to permit the primer design. Tri-nucleotide SSR motifs were the most abundant, 711 were detected (45.2%), followed by tetra- and hexa-nucleotide motifs (36.2% and 11.5%, respectively). Penta-nucleotide motifs were found less frequently, with only 107 (Table 4). The length of SSR motifs ranged from 12 to 48 nucleotides. If the thresholds were raised from four to five repeats for tri-nucleotide motifs and from three to four repeats for tetra-, penta- and

TABLE 2 Summary of the 10 most abundant expressed genes in the transcriptome of *Podosphaera pannosa* according to FPKM values

Contig ID	FPKM	Accession	Annotation	Source
c169657_g1_i1	532,61	CCU82453.1	Dihydrofolate reductase	<i>Blumeria graminis</i>
c74378_g1_i1	495,70	KHJ35908.1	Putative fungal protein	<i>Erysiphe necator</i>
c130120_g1_i1	294,08	KHJ33619.1	Putative methyltransferase	<i>Erysiphe necator</i>
c92601_g1_i1	288,67	CAD66431.1	Aquaporin 1	<i>Blumeria graminis</i>
c149315_g1_i1	288,18	KHJ34064.1	Putative acyl-desaturase	<i>Erysiphe necator</i>
c149758_g1_i1	276,43	KHJ30843.1	Putative glycoside hydrolase	<i>Erysiphe necator</i>
c152156_g1_i1	273,43	KHJ31592.1	Hypothetical protein	<i>Erysiphe necator</i>
c11058_g1_i1	269,02	AEQ16464.1	Effector protein EC2	<i>Golovinomyces orontii</i>
c130884_g1_i1	248,64	KHJ35817.1	Putative dehydrogenase	<i>Erysiphe necator</i>
c91266_g1_i1	236,90	CCU78156.1	Allergen F4-like/hypothetical protein	<i>Blumeria graminis</i>

Notes. FPKM, Fragments per kilobase of transcript per million fragments mapped.

The information of accession, annotation and source was from BLASTX against NCBI non-redundant protein database.

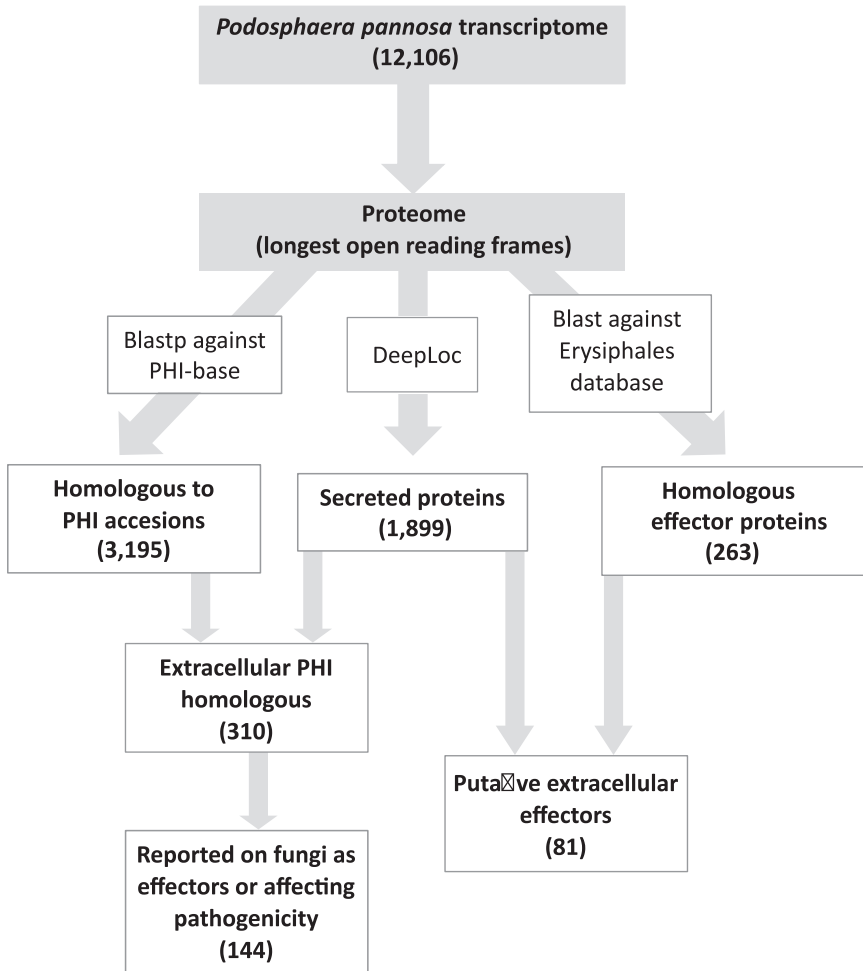


FIGURE 2 Workflow of secretome prediction for *Podosphaera pannosa* during infection of eucalypt (*Eucalyptus urophylla*) using bioinformatics tools. Number of transcripts after each methodology applied is in parenthesis

hexa-nucleotide motifs, the number of detectable SSRs decreased to 392, of which, 289 SSR sequences permitted suitable primer pair design. In total, 309 primer pairs were identified that have the potential to amplify these SSRs (Table 4).

4 | DISCUSSION

Powdery mildews are diseases on the economically important eucalypt crop in South America, but little is known about the pathogen, which was only recently identified as *P. pannosa* (Fonseca et al., 2017). The transcriptome sequencing of *P. pannosa* produced more than 178 million reads, de novo assembled into ~200,000 transcripts. The de novo transcriptome assembly method was used because a whole-genome sequence is not currently available for *P. pannosa*. This method has yielded good results for several non-model organisms, including fungi (Liu et al., 2015; Ross-Davis et al., 2013; Thakur et al., 2013; Yazawa, Kawahigashi, Matsumoto, & Mizuno, 2013) and plants (Castro et al., 2015; Li, Deng, Qin, Lui, & Men, 2012).

As expected, the BLASTX search against the NCBI nr database revealed that the majority of annotated transcripts derived from scraped *P. pannosa*-infected leaves of *E. urophylla* originated from non-*P. pannosa* organisms. After manual filtering, the resulting

12,106 transcripts were identified as the *P. pannosa* transcriptome data set. A similar strategy was used previously to recover transcripts of *Bipolaris sorghicola* from infected sorghum (*Sorghum bicolor*) leaves, which was also based on a de novo transcriptome assembly (Yazawa et al., 2013). In that previous study, the mixed transcriptome of *B. sorghicola*-infected sorghum leaves resulted in 160 pathogen transcripts. For comparison, a previous transcriptome-sequencing strategy with another powdery mildew pathogen resulted in 29,505 transcripts acquired from conidia scraped off *P. plantaginis*-infected ribwort plantain (*Plantago lanceolata*) leaves using a 454 GS-FLX sequencing platform and de novo assembly (Tollenaere et al., 2012). However, that previous study had a low depth of sequencing (1.76 reads per nucleotide) compared to the high, deep sequencing used in this study that provided an average 148-fold coverage. Previous transcriptome sequencing of *G. orontii* haustoria using deep 454 pyrosequencing resulted in 7,077 contigs assembled by mapping back to the *G. orontii* draft genome with 3,725 annotated contigs (Weßling et al., 2012). In this study, the sequencing depth and the number of transcripts of *P. pannosa* obtained demonstrated the merits of the methodology chosen to identify pathogen transcripts, which resulted in strong representation of the total transcriptome of *P. pannosa*. With this method, however, putative *P. pannosa* transcripts with no significant similarity to previously identified genes

TABLE 3 Effectors found in the transcriptome of *Podosphaera pannosa* after the BLASTP search with significant similarity to proteins in the final set of 144 PHI-base accessions (extracellular, described in fungi, and reducing virulence or described as effectors)

seq_id	e-value	blast_top_hit	Location	PHI
c69581_g6_i2	3E-33	CCU75383.1 EKA-like protein [<i>Blumeria graminis</i> f. sp. <i>hordei</i> DH14]	Extracellular	A6SR05#PHI:7594#BcCFEM1 #40559 # <i>Botrytis cinerea</i> #reduced_virulence_ unaffected_pathogenicity
c66694_g1_i1	3E-12	CCU76905.1 CSEP0219 putative effector protein [<i>Blumeria graminis</i> f. sp. <i>hordei</i> DH14]	Extracellular	G4MNL6#PHI:2092#Cch1 #318829 # <i>Magnaporthe oryzae</i> #reduced_virulence
c1614_g1_i1	4E-113	CCU77073.1 EKA-like protein [<i>Blumeria graminis</i> f. sp. <i>hordei</i> DH14]	Extracellular	Q96TN6#PHI:256#GAS1 #318829 # <i>Magnaporthe oryzae</i> #reduced_virulence
c172634_g1_i1	8E-58	CCU77073.1 EKA-like protein [<i>Blumeria graminis</i> f. sp. <i>hordei</i> DH14]	Extracellular	Q96TN6#PHI:256#GAS1 #318829 # <i>Magnaporthe oryzae</i> #reduced_virulence
c49687_g1_i2	4E-114	CCU77073.1 EKA-like protein [<i>Blumeria graminis</i> f. sp. <i>hordei</i> DH14]	Extracellular	Q96TN6#PHI:256#GAS1 #318829 # <i>Magnaporthe oryzae</i> #reduced_virulence
c127010_g1_i1	3E-28	CCU82076.1 CELP0017 effector like protein [<i>Blumeria graminis</i> f. sp. <i>hordei</i> DH14]	Extracellular	J9VHR6#PHI:3236#Aph1#5207 # <i>Cryptococcus neoformans</i> #reduced_virulence

TABLE 4 Report statistics of microsatellites (simple sequence repeats; SSRs) detection and primer design from the transcriptome of *Podosphaera pannosa*. Minimum number of repeats described as unit size/minimum number of repeats

	Tri/4x, Tetra/3x, Penta/3x, Hexa/3x	Tri/5x, Tetra/4x, Penta/4x, Hexa/4x
Total of SSRs detected	1,571	392
Number of tri-nucleotide SSRs detected	711	237
Number of tetranucleotide SSRs detected	570	121
Number of penta-nucleotide SSRs detected	107	15
Number of hexa-nucleotide SSRs detected	183	19
Number of sequences with primers pairs	1,152	289
Total primer pairs identified	1,389	309

from Erysiphales were filtered out of the transcriptome data. Thus, it is likely that some highly conserved genes and other uncharacterized genes are underrepresented in our *P. pannosa* transcriptome.

The 10 most abundant transcripts included genes encoding enzymes likely involved in fungal establishment and growth, such as DHFR, putative methyltransferases, acyl-desaturase, glycoside hydrolase and dehydrogenase. Additional transcribed genes putatively encoded an aquaporin and an effector protein.

DHFR had the highest FPKM value, representing the *P. pannosa* gene that was most transcribed during eucalypt infection. DHFR is used in cell proliferation and cell growth, and it also plays a role in regulating the amount of tetrahydrofolate in the cell, which is essential for purine and thymidylate synthesis (Schnell, Dyson, & Wright, 2004). The transcripts of a putative methyltransferase gene were also found in high number. In *M. oryzae*, this enzyme was reported as

important for DNA methylation contributing to fungal development and genome defence (Jeon et al., 2015), and in the entomopathogenic fungus *Beauveria bassiana* as contributing to spore viability, fungal development, protein secretion and virulence (Qin, Ortiz-Urquiza, & Keyhani, 2014).

Glycoside hydrolase (or glycosyl hydrolase) family proteins are known to be involved in host cell-wall degradation. A plant cell-wall-degrading enzyme was identified among the most abundant transcripts in *P. pannosa*. Although high levels of cell-wall-degrading enzyme gene transcription might be expected from necrotrophic or facultative plant pathogens, this enzyme could augment cell penetration and appressorium formation by *P. pannosa*. Glycoside hydrolase family proteins were previously reported as occurring during the initial infection stages of *M. oryzae* on rice, where these enzymes were found to be involved in host cell-wall degradation and appressorium

formation (Kawahara et al., 2012). In addition, the same work found that glycoside hydrolase-encoding transcripts were upregulated in both host and pathogen, suggesting that these cell-wall-degrading enzymes play roles in both host defence and pathogen attack. Similarly, glycoside hydrolases were also previously characterized in the biotrophic rust fungus, *Uromyces fabae* (Murphy, Powlowski, Wu, Butler, & Tsang, 2011).

Aquaporins are specialized pore proteins that help channel water and enhance the permeability of cell membranes. Genes encoding aquaporins are found in mammals, plants and some microorganisms including the powdery mildew pathogen, *B. graminis* (Tanghe, Dijck, & Thevelein, 2006). Although the function of these aquaporins is not well established for fungi, a high level of aquaporin transcription was observed during mycelial growth of the ectomycorrhizal fungus *Laccaria bicolor*. Such findings suggest a demand for elevated water permeability in the plasma membrane when hyphae are growing (Nehls & Dietz, 2014). As expected, genes encoding products that contribute to establishment, development and growth of *P. pannosa* were highly expressed during eucalypt infection.

Secreted proteins play an important role in communication between the fungal pathogen and the host plant; some of these secreted proteins may be effectors, which are defined as proteins and other compounds that enhance disease development by targeting host processes, but are redundant to basal growth processes in the pathogen (Godfrey et al., 2010). In this work, we predicted the set of secreted proteins using the machine learning program DeepLoc, which relies on the sequence information to provide results that outperform current state-of-the-art algorithms based on homology (Almagro Armenteros et al., 2017).

The transcript c11058_g1_i1, which encoded a protein orthologue of the effector protein GoEC2 of *G. orontii*, was one of the most abundant transcripts during *P. pannosa* infection of eucalypt; the encoded protein is predicted to be extracellular. However, this protein did not appear in the search results against PHI-base (see below). GoEC2 functionality during *G. orontii* infection of *Arabidopsis thaliana* was explored by Schmidt et al. (2014), who reported the potential of GoEC2 to enhance host susceptibility by promoting fungal penetration. In *B. graminis*, high transcript levels of the orthologous gene, BEC2, were observed at the time of appressorium formation, after which transcript levels decreased (Schmidt et al., 2014). In addition to being found among *P. pannosa* transcripts in this study, and the orthologous genes in *G. orontii* and *B. graminis*, orthologous genes have also been described in the genome of *Erysiphe pisi*, causal agent of pea powdery mildew (Spanu et al., 2010), indicating gene conservation among powdery mildew pathogen species.

Only six of the putative effectors found after the BLASTP were comprised in the final set of 144 PHI-base proteins. This may be because the PHI database consists only of entries that have been tested experimentally, whereas many of the effectors in the NCBI nr database are only predicted based on significant similarity. In addition, when experimental evidence is available it is not necessarily uploaded to the PHI database. Therefore, the list of putative effectors is much larger in the nr database compared to the PHI database.

AVR genes encode proteins that may potentially have dual roles as an AVR protein and as an effector, and are also known to increase the pathogenicity of *B. graminis* f. sp. *hordei* in barley plants (Amselem et al., 2015). In *P. pannosa*, eight transcripts were found homologous to genes encoding putative AVR10-like proteins including five *B. graminis* and three *E. necator* proteins. Resistant barley encodes Mla10 resistance proteins that recognize the presence of AVR_{A10}, thereby eliciting the hypersensitive response (localized cell death) typical of gene-for-gene interactions (Ridout et al., 2006). These transcripts were not selected as putative effectors by the workflow used in this study because unlike other known effectors, the AVR_{A10} genes encode proteins that do not contain a secretory signal peptide (Ridout et al., 2006) and the corresponding proteins were not predicted as extracellular by DeepLoc. AVR_{A10} genes belong to a large gene family along with their homologous AVRk1 genes and the EKA family, with more than 1,350 homologues in the *B. graminis* genome (Spanu et al., 2010). It has recently been hypothesized that this enormous number of homologues in the genome could act as an reservoirs from which new effector genes may quickly evolve to overcome the host resistance (Amselem et al., 2015). Our *P. pannosa* transcriptome encoded 102 EKA-like proteins, 40 of which were predicted to be secreted. Out of the six effectors found by a BLASTP search against the Erysiphales protein database that are homologous in the final PHI-base set, four were EKA-like proteins (Table 3).

Studies of population biology rely on the use of molecular markers that can exhibit polymorphism among different populations. Microsatellite or simple sequence repeats (SSRs) markers have the advantage over other DNA fingerprinting because they are co-dominant, highly polymorphic, species specific and require minimal DNA. The search for SSRs within the 12,106 transcripts of *P. pannosa* yielded 392 SSR motifs, of which 309 sequences allow for putative primer design. Searches for repeat motifs within nucleotide sequences from EST libraries, whole genomes and transcriptomes have been successfully used for several fungi and plants (Frenkel et al., 2012; Parchman, Geist, Grahnen, Benkman, & Buerkle, 2010; Salgado et al., 2014; Tucker et al., 2015). Although SSR markers derived from expressed sequences or transcriptome data can be considered less informative due to DNA sequence conservation compared to genomic sequences (Bouck & Vision, 2007), such markers are cost and time effective, and provide useful polymorphic markers from putatively mRNA-derived sequences (Frenkel et al., 2012; Tucker et al., 2015). The SSR sequences identified here provide substantial resources to design of SSR markers specifically for examining the population structure of *P. pannosa*.

This study describes the generation of transcriptome sequences of *P. pannosa* using next-generation sequencing technology and de novo assembly for protein identification and secretome prediction that allow further study of *P. pannosa*. To our knowledge, this is the first study of *P. pannosa* transcriptome during infection of *E. urophylla* using RNA-Seq. The results generated in this work increase general knowledge about powdery mildew pathogens and provide useful information for new advances in research and disease management.

Furthermore, this study provides a basis for better understanding the *P. pannosa*-eucalypt pathosystem, which could parallel studies of the eucalypt transcriptome to help determine host resistance mechanisms.

ACKNOWLEDGEMENTS

The authors are very thankful to the USDA Forest Service for partial funding of this project, and CAPES for financial support. The authors are also grateful to Raul de P. Pires for laboratory assistance.

ORCID

Jane E. Stewart  <https://orcid.org/0000-0001-9496-6540>

Acelino C. Alfenas  <https://orcid.org/0000-0001-7776-3362>

Ned B. Klopfenstein  <https://orcid.org/0000-0002-9776-3973>

REFERENCES

- Alfenas, A. C., Zauza, E. A. V., Mafia, R. G., & de Assis, T. F. (2009). *Clonagem e doenças do eucalipto*. Viçosa: Editora UFV.
- 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, 3387–3395.
- Amano, K. (1986). *Host range and geographical distribution of the powdery mildew fungi* (p. 741). Tokyo: Japan Scientific Societies Press.
- Amselem, J., Vigouroux, M., Oberhaensli, S., Brown, J. K., Bindschedler, L. V., Skamnioti, P., ... Sacristán, S. (2015). Evolution of the EKA family of powdery mildew avirulence-effector genes from the ORF 1 of a LINE retrotransposon. *BMC Genomics*, 16, 917. <https://doi.org/10.1186/s12864-015-2185-x>
- Bindschedler, L. V., Panstruga, R., & Spanu, P. D. (2016). Mildew-Omics: How global analyses aid the understanding of life and evolution of powdery mildews. *Frontiers in Plant Science*, 7, 1–11. <https://doi.org/10.3389/fpls.2016.00123>
- Bouck, A., & Vision, T. (2007). The molecular ecologist's guide to expressed sequence tags. *Molecular Ecology*, 16, 907–924. <https://doi.org/10.1111/j.1365-294X.2006.03195.x>
- Braun, U., & Cook, R. T. A. (2012). *Taxonomic manual of the Erysiphales (Powdery mildew)*, CBS Biodiv. Utrecht: CBS-KNAW Fungal Biodiversity Centre.
- Bruce, M., Neugebauer, K. A., Joly, D. L., Migeon, P., Cuomo, C. A., Wang, S., ... Fellers, J. P. (2014). Using transcription of six *Puccinia triticina* races to identify the effective secretome during infection of wheat. *Frontiers in Plant Science*, 4, 520. <https://doi.org/10.3389/fpls.2013.00520>
- Castro, J. C., Maddox, J. D., Cobos, M., Requena, D., Zimic, M., Bombarely, A., ... Medina, A. E. (2015). *De novo* assembly and functional annotation of *Myrciaria dubia* fruit transcriptome reveals multiple metabolic pathways for L-ascorbic acid biosynthesis. *BMC Genomics*, 16, 997. <https://doi.org/10.1186/s12864-015-2225-6>
- Cho, S. E., Park, J. H., Lee, S. H., Lee, C. K., & Shin, H. D. (2016). Occurrence of powdery mildew caused by *Podosphaera pannosa* on lemon eucalyptus. *Forest Pathology*, 46, 264–266. <https://doi.org/10.1111/efp.12277>
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21, 3674–3676. <https://doi.org/10.1093/bioinformatics/bti610>
- Fonseca, N. R., Guimaraes, L. M. S., Pires, R. P., Klopfenstein, N. B., & Alfenas, A. C. (2017). Eucalypt powdery mildew caused by *Podosphaera pannosa* in Brazil. *Tropical Plant Pathology*, 42, 261–272.
- Frenkel, O., Portillo, I., Brewer, M. T., Péros, J. P., Cadle-Davidson, L., & Milgroom, M. G. (2012). Development of microsatellite markers from the transcriptome of *Erysiphe necator* for analysing population structure in North America and Europe. *Plant Pathology*, 61, 106–119. <https://doi.org/10.1111/j.1365-3059.2011.02502.x>
- Godfrey, D., Böhlenius, H., Pedersen, C., Zhang, Z., Emmersen, J., & Thordal-Cristensen, H. (2010). Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *BMC Genomics*, 11, 317. <https://doi.org/10.1186/1471-2164-11-317>
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., ... Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29, 644–652. <https://doi.org/10.1038/nbt.1883>
- Guyon, K., Balagué, C., Roby, D., & Raffaele, S. (2014). Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. *BMC Genomics*, 15, 336. <https://doi.org/10.1186/1471-2164-15-336>
- Huynh, M. D., Page, J. T., Richardson, B. A., & Udall, J. A. (2015). Insights into transcriptomes of Big and Low sagebrush. *PLoS One*, 10, 1–11. <https://doi.org/10.1371/journal.pone.0127593>
- Jeon, J., Choi, J., Lee, G.-W., Park, S. Y., Huh, A., Dean, R. A., & Lee, Y. H. (2015). Genome-wide profiling of DNA methylation provides insights into epigenetic regulation of fungal development in a plant pathogenic fungus *Magnaporthe oryzae*. *Scientific Reports*, 5, 8567. <https://doi.org/10.1038/srep08567>
- Kawahara, Y., Oono, Y., Kanamori, H., Matsumoto, T., Itoh, T., & Minami, E. (2012). Simultaneous RNA-seq analysis of a mixed transcriptome of rice and blast fungus interaction. *PLoS One*, 7, e49423. <https://doi.org/10.1371/journal.pone.0049423>
- Keane, P. J., Kile, G. A., Podger, F. D., & Brown, B. N. (2000). *Diseases and pathogens of eucalyptus*. Collingwood, VIC: CSIRO Publishing.
- Kemen, A. C., Agler, M. T., & Kemen, E. (2015). Host-microbe and microbe-microbe interactions in the evolution of obligate plant parasitism. *New Phytology*, 206, 1207–1228. <https://doi.org/10.1111/nph.13284>
- Kim, H. A., Lim, C. J., Kim, S., Choe, J. K., Jo, S.-H., Baek, N., & Kwon, S.-Y. (2014). High-throughput sequencing and *de novo* assembly of *Brassica oleracea* var. *capitata* L. for transcriptome analysis. *PLoS One*, 9, 1–10. <https://doi.org/10.1371/journal.pone.0092087>
- Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12, 323. <https://doi.org/10.1186/1471-2105-12-323>
- Li, D., Deng, Z., Qin, B., Lui, X., & Men, Z. (2012). *De novo* assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). *BMC Genomics*, 13, 192. <https://doi.org/10.1186/1471-2164-13-192>
- Liu, J.-J., Sturrock, R. N., Snieszko, R. A., Williams, H., Benton, R., & Zamany, A. (2015). Transcriptome analysis of the white pine blister rust pathogen *Cronartium ribicola*: *de novo* assembly, expression profiling, and identification of candidate effectors. *BMC Genomics*, 16, 678. <https://doi.org/10.1186/s12864-015-1861-1>
- Lorrain, C., Hecker, A., & Duplessis, S. (2015). Effector-mining in the poplar rust fungus *Melampsora larici-populina* Secretome. *Frontier Plant Science*, 6, 1051. <https://doi.org/10.3389/fpls.2015.01051>
- Meinhardt, L. W., Costa, G. G., Thomazella, D. P., Teixeira, P. J., Carazzolle, M. F., Schuster, S. C., ... Bailey, B. A. (2014). Genome and secretome analysis of the hemibiotrophic fungal pathogen, *Moniliophthora roreri*, which causes frosty pod rot disease of cacao: mechanisms of the biotrophic and necrotrophic phases. *BMC Genomics*, 15, 1–25. <https://doi.org/10.1186/1471-2164-15-164>

- Murphy, C., Powlowski, J., Wu, M., Butler, G., & Tsang, A. (2011). Curation of characterized glycoside hydrolases of Fungal origin. *Database*, 2011, bar020–bar020. <https://doi.org/10.1093/database/bar020>
- Nehls, U., & Dietz, S. (2014). Fungal aquaporins: cellular functions and ecophysiological perspectives. *Applied Microbiology Biotechnology*, 98, 8835–8851. <https://doi.org/10.1007/s00253-014-6049-0>
- Old, K. M., Wingfield, M. J., & Qing Yuan, Z. (2003). A manual of diseases of eucalypts in southeast asia. *Cifor*, 106. <https://doi.org/10.17528/cifor/001428>
- Páez, U. A. H., Romero, I. A. G., Restrepo, S. R., Gutiérrez, F. A. A., & Castaño, D. M. (2015). Assembly and analysis of differential transcriptome responses of *Hevea brasiliensis* on interaction with *Microcyclus ulei*. *PLoS One*, 10, 1–21. <https://doi.org/10.1371/journal.pone.0134837>
- Parchman, T. L., Geist, K. S., Grahnen, J. A., Benkman, C. W., & Buerkle, C. A. (2010). Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics*, 11, 180. <https://doi.org/10.1186/1471-2164-11-180>
- Qin, Y., Ortiz-Urquiza, A., & Keyhani, N. O. (2014). A putative methyltransferase, *mtrA*, contributes to development, spore viability, protein secretion and virulence in the entomopathogenic fungus *Beauveria bassiana*. *Microbiology*, 160, 2526–2537. <https://doi.org/10.1099/mic.0.078469-0>
- Ridout, C. J., Skamnioti, P., Porritt, O., Sacristan, S., Jones, J. D. G., & Brown, J. K. M. (2006). Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. *The Plant Cell*, 18, 2402–2414. <https://doi.org/10.1105/tpc.106.043307>
- Ross-Davis, A. L., Stewart, J. E., Hanna, J. W., Kim, M.-S., Knaus, B. J., Cronn, R., ... Klopfenstein, N. B. (2013). Transcriptome of an *Armillaria* root disease pathogen reveals candidate genes involved in host substrate utilization at the host-pathogen interface. *Forest Pathology*, 43, 468–477. <https://doi.org/10.1111/efp.12056>
- Salgado, L. R., Koop, D. M., Pinheiro, D. G., Rivallan, R., Le Guen, V., Nicolás, M. F., ... García, D. (2014). *De novo* transcriptome analysis of *Hevea brasiliensis* tissues by RNA-seq and screening for molecular markers. *BMC Genomics*, 15, 236. <https://doi.org/10.1186/1471-2164-15-236>
- Schmidt, S. M., Kuhn, H., Micali, C., Liller, C., Kwaaitaal, M., & Ranstruga, R. (2014). Interaction of a *Blumeria graminis* f. sp. *hordei* effector candidate with a barley ARF-GAP suggests that host vesicle trafficking is a fungal pathogenicity target. *Molecular Plant Pathology*, 15, 535–549. <https://doi.org/10.1111/mpp.12110>
- Schnell, J. R., Dyson, H. J., & Wright, P. E. (2004). Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annual Review Biophysical Biomolecular Structure*, 33, 119–140. <https://doi.org/10.1146/annurev.biophys.33.110502.133613>
- Spanu, P. D., Abbott, J. C., Amselem, J., Burgis, T. A., Soanes, D. M., Stüber, K., ... Panstruga, R. (2010). Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science*, 30, 1543–1546. <https://doi.org/10.1126/science.1194573>
- Staton, E. (2018) Sestaton/Hmmer2Go: Hmmer2Go Version 0.17.8. Zenodo. <https://doi.org/10.5281/zenodo.592436>
- Tanghe, A., Van Dijck, P., & Thevelein, J. M. (2006). Why do microorganisms have aquaporins? *Trends Microbiology*, 14, 78–85.
- Thakur, K., Chawla, V., Bhatti, S., Swarnkar, M. K., Kaur, J., Shankar, R., & Jha, G. (2013). *De novo* transcriptome sequencing and analysis for *Venturia inaequalis*, the devastating apple scab pathogen. *PLoS One*, 8, e53937. <https://doi.org/10.1371/journal.pone.0053937>
- Tollenaere, C., Susi, H., Nokso-Koivisto, J., Koskinen, P., Tack, A., Auvinen, P., ... Laine, A.-L. (2012). SNP design from 454 sequencing of *Podosphaera plantaginis* transcriptome reveals a genetically diverse pathogen metapopulation with high levels of mixed-genotype infection. *PLoS One*, 7, 1–10. <https://doi.org/10.1371/journal.pone.0052492>
- Tucker, M. A., Moffat, C. S., Ellwood, S. R., Tan, K. C., Jayasena, K., & Oliver, R. P. (2015). Development of genetic SSR markers in *Blumeria graminis* f. sp. *hordei* and application to isolates from Australia. *Plant Pathology*, 64, 337–343. <https://doi.org/10.1111/ppa.12258>
- Wang, X., Shi, W., & Rinehart, T. (2015). Transcriptomes that confer to plant defense against powdery mildew disease in *Lagerstroemia indica*. *International Journal of Genomics*, 2015, 528395. <https://doi.org/10.1155/2015/528395>
- Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Review Genetics*, 10, 57–63. <https://doi.org/10.1038/nrg2484>
- Weßling, R., Schmidt, S. M., Micali, C. O., Knaust, F., Reinhardt, R., Neumann, U., ... Panstruga, R. (2012). Transcriptome analysis of enriched *Golovinomyces orontii* haustoria by deep 454 pyrosequencing. *Fungal Genetics Biology*, 49, 470–482. <https://doi.org/10.1016/j.fgb.2012.04.001>
- Winnenburg, R., Urban, M., Beacham, A., Baldwin, T. K., Holland, S., Lindeberg, M., ... Köhler, J. (2008). PHI-base update: Additions to the pathogen-host interaction database. *Nucleic Acids Research*, 36, 572–576. <https://doi.org/10.1093/nar/gkm858>
- Yazawa, T., Kawahigashi, H., Matsumoto, T., & Mizuno, H. (2013). Simultaneous transcriptome analysis of sorghum and *Bipolaris sorghicola* by using RNA-seq in combination with *de novo* transcriptome assembly. *PLoS One*, 8, e62460. <https://doi.org/10.1371/journal.pone.0062460>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Fonseca NR, Ibarra Caballero J, Kim M-S, et al. Transcriptome analysis of a powdery mildew pathogen (*Podosphaera pannosa*) infecting *Eucalyptus urophylla*: *De novo* assembly, expression profiling and secretome prediction. *For Path*. 2019;49:e12508. <https://doi.org/10.1111/efp.12508>