# Putative Pathogenicity genes of *Phytophthora cinnamomi* identified via RNA-Seq analysis of pre-infection structures

**Anandi Reitmann**, **Dave K. Berger**, **and Noëlani van den Berg**


A. Reitmann

University of Pretoria, Department of Genetics, Forestry and Agricultural Biotechnology Institute, Pretoria, 0002, South Africa


D.K. Berger

University of Pretoria, Department of Plant Science, Forestry and Agricultural Biotechnology Institute, Pretoria, 0002, South Africa


N. van den Berg (Corresponding author)

University of Pretoria, Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute, Pretoria, 0002, South Africa

E-mail: noelani.vdberg@fabi.up.ac.za

Tel. +27 12 420 3856

## Abstract

*Phytophthora cinnamomi* is an economically important oomycete that infects more than 3,000 plant species. We aimed to identify the repertoire of genes expressed during pre-infection stages by analysing an RNA-Seq library of cysts and germinating cysts of a *P. cinnamomi* isolate, originating from *Persea americana*. Over 70,000 transcripts were identified from 225,049 contigs, assembled from 13 million Illumina paired-end reads. Contaminant sequences were eliminated, resulting in 37,534 transcripts used in further analysis. A total of 1,394 transcripts had a putative role in pathogenesis. Genes aiding in detoxification and metabolite transport (cytochrome P450 and ABC transporters) and protection against oxidative stress were most abundant, followed by the genes coding cell wall degrading enzymes. The transcript set included 44 putative RXLR effector genes and genes encoding elicitin and necrosis-inducing proteins. Expression patterns of seven putative pathogenicity genes (encoding RXLR-, necrosis-inducing *Phytophthora* protein 1 (NPP1), elicitin, polygalacturonase, cellulose binding and elicitor lectin (CBEL), mucin, and adhesion proteins) were assessed across four *in vitro* developmental stages of *P. cinnamomi*. High expression of these genes in zoospores suggests their functional importance in the subsequent developmental stage, germination of cysts, implying a role in pre-infection. This work is the first step towards understanding the molecular basis of infection strategies employed by *P. cinnamomi*.

Key words: *Phytophthora cinnamomi*, Pathogenicity, Expression analysis, RXLR, Oomycete

## Introduction

*Phytophthora cinnamomi* Rands is an oomycete belonging to the *Stramenopiles* which also include golden-brown algae and diatoms (Cavalier-Smith 1993). This hemi-biotrophic pathogen causes root rot of woody plants and in some species, stem cankers and dieback of young shoots are also observed (Hardham 2005). As a broad host range pathogen, with a nearly worldwide distribution (OEPP/EPPO 2004), it is a threat to natural ecosystems, biodiversity, forestry and agriculture. The economic impact due to *P. cinnamomi* infestation is evident in the forest and food industry, affecting eucalyptus, pine, oak (Zentmyer 1980), avocado, pineapple, peach, chestnut and macadamia trees. Losses include not only decreases in crop yield and product value, but also large amounts of money spent annually on control measures.

As a soil-borne pathogen there are no effective means to eradicate *P. cinnamomi* from infested areas as it survives in moist soil or dead plant material as chlamydospores for up to six years (Zentmyer and Mircetich 1966). Phosphite injections have proved helpful in combating *P. cinnamomi* in *Persea americana* (Darvas et al. 1984), *Eucalyptus marginata* and *Banksia* species (Shearer and Fairman 2007). Some industries, such as avocado, heavily rely on tolerant rootstocks to produce sufficient fruit in the presence of *P. cinnamomi.*

*P. cinnamomi* predominantly reproduces asexually even in the presence of both A1 and A2 mating types (Hardham 2005). The main infection agent is uni-nucleate, bi-flagellate, motile zoospores that are released from sporangia (Hardham 2005). They are spread to suitable hosts by water present in soil, through rain splash and soil movement. Zoospores locate living plant material by responding to chemo- and electrotactic signals (Zentmyer 1961). Contact with plant cell wall carbohydrates, together with chemical and physical influences, induces zoospore encystment, where spores attach to the host cell surface and deposit a cell wall that allows the build up of turgor pressure (Hardham 2005). Within two to three

days, sporangia appear on the plant surface (Hardham 2005) and the asexual reproduction cycle may repeat several times, rapidly increasing the amount of inocula.

Disease results from a compatible plant-pathogen interaction, where the pathogen successfully evades host detection and invades and colonizes host tissue.  In an incompatible interaction  strong defence responses at the site of infection, such as localized cell death (hypersensitive response) (Jones and Dangl 2006) and callose deposition (Benhamou and Nicole 1999) restrict pathogen growth.  This is the result of recognition of pathogen-associated molecular patterns (PAMPs) by host pathogen recognition receptors (PRRs) (so called PAMP-triggered immunity (PTI)) and detection of avirulence factors by host resistance proteins (effector-triggered immunity (ETI)) (Jones and Dangl 2006; Tsuda and Katagiri 2010).  To understand this system, insight into both plant defence pathways, as well as knowledge of pathogenicity genes (especially the pathogen effector repertoire) is critical.

Pathogenicity genes encode proteins (or molecules) that enable infection of a host directly or indirectly.  Diverse pathogenicity factors may act at different stages of infection and at different host locations (eg. apoplast, cytoplasm, nucleus).  For example, RXLR effector Avrblb2 from *Phytophthora infestans* acts at the haustorial interface, preventing the secretion of host papain-like cysteine protease C14 into the apoplast and thereby enhancing potato susceptibility to *P. infestans* (Bozkurt et al. 2011).  Identification and functional characterization of pathogenicity genes gives insight into the mechanisms employed by pathogens to establish infection.

Little is known about the molecular events that govern a successful *P. cinnamomi* infection. Transcriptomic analysis proved a useful tool for identifying infection related genes in other oomycete species (Panabières et al. 2005; Randall et al. 2005; Chen et al. 2007; Krajaejun et al. 2014).  Recently, RNA-Seq emerged as the preferred technology for transcriptome

4

analysis and gene expression profiling. Ye et al. (2011) profiled expression patterns of 30 pathogenicity gene families from *Phytophthora sojae*, during *in vitro* growth as well as *in planta* infection (Ye et al. 2011). Currently the genomes of thirteen oomycete species have been published (Tyler et al. 2006; Haas et al. 2009; Baxter et al. 2010; Lévesque et al. 2010; Lamour et al. 2012; Jiang et al. 2013; Studholme et al. 2016). Sequence data for oomycete host species are rapidly accumulating. Together, these genomic resources have been used to identify host and pathogen transcripts in large-scale transcriptome analysis, enabling the investigation of gene expression changes in both the host and the pathogen simultaneously. Even in the absence of reference genome sequences, RNA-Seq technology can be applied via *de novo* assembly (Ward et al. 2012).

In this study, we generated an RNA-Seq library of the cyst and germinating cyst stages of *P. cinnamomi* from a high sporulating isolate recovered from avocado. Paired-end reads were *de novo* assembled and annotated based on homology to *P. infestans* protein models. More than 1,300 putative pathogenicity-related transcripts were identified and the expression pattern of seven pathogenicity genes was determined over four *in vitro* life stages (mycelia, sporulating mycelia, zoospores, and cysts and germinating cysts) using reverse transcriptase quantitative PCR (RT-qPCR).

**Methods**

**Isolate information**

*Phytophthora cinnamomi* isolate, CNVDB29 (mating type A2), is maintained by the Fruit Tree Biotechnology Programme at the University of Pretoria, South Africa and originates from roots of *P. americana* collected in Tzaneen, South Africa. The culture was purified from long term storage stock by isolating a single hyphal tip from water agar onto ½ potato dextrose agar (PDA) and then maintained on 5% V8 (vegetable juice) agar (cleared V8 juice,

0.1 g/L CaCO$_3$, 20 mg/L *B*-sitosterol, 17 g/L agar and distilled water). Isolate identity was confirmed by a species specific PCR (Online Resource 1), using primers directed against the *LPV3* gene (Kong et al. 2003).


**Enrichment of the *in vitro* developmental structures of *P. cinnamomi***

Mycelia, sporulating mycelia, zoospores and cysts and germinating cysts were produced using an adjusted protocol from Chen and Zentmeyer (1970). Multiple 10% V8 agar blocks (approximately 2 mm x 2 mm in size) containing a 3-day-old culture, were incubated in 5% V8 broth. After two days incubation under fluorescent light at 23 °C, mycelial plugs were washed with three changes of distilled water and harvested with a nylon sieve. Sporangia formation was induced by incubating mycelial plugs under fluorescent light at 23 °C in river water (filtered twice through chromatography paper MN260 (Macherey-Nagel, Düren, Germany)) until sporangia developed (36 hours). Both mycelia and sporulating mycelia were harvested with a nylon sieve, frozen in liquid nitrogen, ground to a fine powder and stored at -70 °C. Zoospore release from sporangia was induced by washing sporulating mycelia with two changes of distilled water at 4 °C, followed by incubation at 4 °C for 12 min. Zoospores released approximately 1 hour later, while being incubated at 23 °C and were separated from sporulating mycelia through 3 layers of Mira cloth (pore size 22-25 µm) placed in a nylon sieve. A sample of the zoospore suspension was collected and analysed under the microscope to assess the composition of the collected material. To enrich for germinating cysts, zoospores were vigorously shaken for 2 min and incubated in the dark at 19 °C for approximately 90 min. This resulted in a germination rate of 20-40%. Both zoospores and germinating cysts were harvested via centrifugation (10000 rpm at 4 °C for 5 min) in the Centrifuge 5810 R (Eppendorf, Hamburg, Germany). We assume that the mRNA population was not altered during this time. Pellets were immediately frozen in liquid nitrogen and stored at -70 °C.

**RNA extraction and cDNA synthesis**

Total RNA was extracted from mycelia and sporulating mycelia using the CTAB Lithium-chloride method (Chang et al. 1993). Tri-Reagent (Ambion, Life technologies, Carlsbad, CA, USA) was used, to obtain total RNA from zoospores and cysts and germinating cysts. All RNA samples were treated with DNase I enzyme (Fermentas, Life Sciences, Hanover, USA), purified with the Qiagen RNeasy clean-up kit (Qiagen, Valencia, California, USA) and concentrations determined with the NanoDrop 2000 Spectrophotometer (Thermo Scientific, Wilmington, USA). Prior to sequencing, RNA quality was assessed with the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). cDNA was synthesized with the ImProm-II Reverse Transcriptase system (Promega, Madison, Wisconsin, USA) with oligo poly-dT primers (Roche Diagnostics) for RT-qPCR.

**Sequencing and RNA-Seq library assembly**

A cDNA library of cysts and germinating cysts of *P. cinnamomi* isolate CNVDB29 was prepared by the Beijing Genomics Institute (BGI) through fragmentation of mRNA enriched RNA (oligodT-primers) and random hexamer primed cDNA synthesis. The cDN Afragments (100bp in length) were sequenced on the HiSeq 2000 Illumina platform (deposited in the European Nucleotide Archive run accession ERR863673). Reads were trimmed to 91bp and overlapping single-end reads assembled into contigs using Trinity (Grabherr et al. 2011) at default parameters. Transcript sequences were obtained by assembling paired end reads (with gap size 150bp) into contigs, and in cases where paired ends spanned two contigs, gaps were filled with unassigned nucleotides (N). This was achieved by aligning paired reads with @blastdb via blastx (E-value $<1\times10^{-5}$). The orientation and coding regions of transcripts were determined according to the best blast hit in the National Centre for Biotechnology Information (NCBI) non-redundant protein database, the Swiss protein database, Clusters of orthologous groups (COG) database and KEGG classes (in that order). EST Scan (Iseli et al. 1999) (http://estscan.sourceforge.net/) was used to predict the

coding region and direction of transcripts in cases where they did not significantly align to any protein sequences in the mentioned databases. To assess the robustness of the assembly a select few transcripts were PCR-amplified with primers designed from the gene models from our transcriptome, sequenced and alligned to the full length sequences in the genome of *P. cinnamomi* var. *cinnamomi* v1.0 (isolate CBS 144.22) (http://genome.jgi.doe.gov/Phyci1/Phyci1.home.html, downloaded March, 2013) via blastx analysis (E-value $<1 \times 10^5$). Furthermore, predicted protein sequences were also compared and aligned with Mafft Version 7 (Katoh and Standley 2013), trimmed in MEGA 4 (Tamura et al. 2007) and visual outputs created with GeneDoc Version 2.7.000 (Nicholas et al. 1997). All sequence data generated in this study was deposited in the European Nucleotide Archive (http://www.ebi.ac.uk/ena) (project PRJEB8451).

**Functional annotation**

Transcripts were annotated using the blastx algorithm (E-value $<1 \times 10^5$) against the COG database (www.ncbi.nlm.nih.gov/COG/), the NCBI non-redundant protein database (www.ncbi.nlm.nih.gov/), as well as protein models predicted in the genomes of *P. infestans* (*Phytophthora infestans* Sequencing Project, Broad Institute of Harvard and MIT (http://www.broadinstitute.org/ downloaded February, 2013)) and *P. cinnamomi* var. *cinnamomi* v1.0 (isolate CBS 144.22) (http://genome.jgi.doe.gov/Phyci1/Phyci1.home.html, downloaded March, 2013).

Signal peptides were predicted in translated amino acid sequences starting with methionine, using neural networks and Hidden Markov Models (HMM) (Nielsen and Krogh 1998) in SignalPv3.0 (Nielsen and Krogh 1998; Bendtsen et al. 2004; Emanuelsson et al. 2007) employing a Sprob cut-off: 0.7 (Kebdani et al. 2010). Sequences containing predicted transmembrane helices (TMHs) downstream of the signal peptide as predicted by TMHMM

v2.0 default parameters (Krogh et al. 2001) were discarded to restrict analysis to putative proteins that are secreted into the apoplast.

The method described in Whisson et al. (2007) was used to identify putative RXLR sequences using HMM. Briefly, the parameters entailed the presence of a signal peptide within the first 10-40 amino acids from the start codon (HMM score>0.9), the presence of an RXLR motif within the first 100 amino acids of the protein, with the EER motif following no more than 40 amino acids after the RXLR motif (Whisson et al. 2007). Contigs were used (instead of transcripts) to identify the longest open reading frame starting with a methionine residue (Cxxxx), as contigs do not contain any unassigned nucleotides. The filtered contig dataset was analysed with Reg-ex (http://www.pythonregex.com/) for the physical presence of an RXLR sequence (Arg-any amino acid-Leu-Arg) within the candidate sequences.

**RXLR protein comparison**

Putative RXLR sequences were identified from the transcriptomic library of *P. cinnamomi* isolate CNVDB29, as described above. Full length copies of putative RXLR proteins were obtained from the genome of *P. cinnamomi* var. *cinnamomi* v1.0 (isolate CBS 144.22) (identity numbers Phyci1-xxxx as denoted in http://genome.jgi.doe.gov/, downloaded March, 2013) with blastp analysis (to ensure the comparison of the reading frame that contains the RXLR motif) (E-value <$1 \times 10^{-5}$). The full length putative RXLR proteins then were compared via blastp analysis (E-value <$1 \times 10^{-8}$, identity >30%) to predicted proteins of *P. infestans*, *Phytophthora sojae*, *Phytophthora ramorum*, *Phytophthora capsici*, *Hyaloperonospora arabidopsidis* and *Pythium ultimum* as in the FungiDB database version 2.3 (Stajich et al. 2012) (March, 2014) to identify putative orthologs. Protein sequences were aligned with Mafft Version 7 (Katoh and Standley 2013), trimmed in MEGA 4 (Tamura et al. 2007) and visual outputs created with GeneDoc Version 2.7.000 (Nicholas et al. 1997).

**Expression analysis via RT-qPCR**

The annotations of transcripts selected for expression analysis (C83210, U15350, U4987, U6659, U6179, U8949 and U6809) were confirmed through blastx analysis (E-value $<1\times10^{-5}$) to *Phytophthora* proteins in the NCBI database (http://www.ncbi.nlm.nih.gov/), FungiDB database version 2.2 (Stajich et al. 2012), and the *P. cinnamomi* genome v1.0 (http://genome.jgi.doe.gov/, downloaded March, 2013) (Online Resource 2).

Primers were designed using Primer3Plus version 2.3.5 (Untergasser et al. 2012) and NetPrimer (PREMIER Biosoft, Palo Alto, CA, USA) (primer sequences are provided under Online Resource 3). Primer sequences for candidate endogenous control genes ubiquitin-conjugated enzyme (Ubc) and Beta-tubulin (β-tubulin) were obtained from literature (King et al. 2010) and the WS041 primer pair from Narayan et al. (2010). All primers were synthesized by Integrated DNA Technologies (Coralville, United States). Primer specificity was evaluated with conventional PCR and melting peak analysis via RT-qPCR using the Bio-Rad CFX Manager Software Version 1.5.534.0511 (Bio Rad laboratories, Hercules, California, USA). Primer efficiency was assessed with the standard curve method ($R^2>0.98$, Efficiency 90 – 110%).

RT-qPCR reactions were performed on the C1000 cycler CFX96 Real-Time System (Bio Rad laboratories). Each 20 µl reaction contained 10µl 2x SensiMix SYBR No-ROX Master mix (Bioline, London, UK), 2 µl diluted cDNA (1:5), 1µl of each forward and reverse gene specific primer (10 µM) and Sabax water. All reactions were performed in triplicate for four biological replicates of each life stage (mycelia, sporulating mycelia, zoospores, cysts and germinating cysts). Cycling conditions were 95 °C for 10 min, 42 cycles of 95 °C for 15 sec, Ta for 15 sec and 72 °C for 15 sec. The expression stability of potential endogenous control genes was calculated with the Bio-Rad CFX Manager Software Version 1.5.534.0511 (Bio Rad laboratories), employing the geNorm method (Vandesompele et al. 2002). Gene

expression levels of each target gene were calculated by the same software, taking the determined primer efficiency into account and using three endogenous control genes (WS041, Ubc, β-tubulin with M-value < 0.5) to normalize the data. Differences in gene expression levels between life stages were statistically analysed using analysis of variance (ANOVA) followed by a Least Square Means Differences Student's t-test using the JMP '9 software (p<0.05) (SAS Institute, Cary, North Carolina).

## Results

### RNA-Seq library construction and assembly

A cDNA library of the cyst and germinating cyst phase of *P. cinnamomi,* isolate CNVDB29, was sequenced on the Illumina HiSeq 2000 platform, generating 13,032,666 paired-end reads (Q20 = 95.3%, N = 0.01%, GC content = 46.3%). Overlapping reads were *de novo* assembled into 225,049 contigs, with a mean contig length of 213 nt (N50 = 257 nt). Transcript depth of coverage ranged from 0.05 to 3,962 (calculated as number of reads * the read length / transcript sequence length) with 51 % of transcripts having a coverage of 5x or more. Transcript assembly was validated by sequencing and aligning select transcripts with the predicted full length sequence from the *P. cinnamomi* genome as well as comparing predicted protein sequences (Online Resource 4a,b and 5).

### Cyst and germinating cyst cDNA library annotation

#### *Functional annotation based on protein sequence similarity*

A total of 37,534 transcripts  were annotated based on similarity to protein sequences in the genomes of *P. cinnamomi* (CBS 144.22) and *P. infestans* (E-value <1x10$^{-5}$) (Haas et al. 2009) (Table 1) using blastx analysis. This included 36,500 transcripts with significant homology to 11,674 genes predicted in the genome of *P. cinnamomi* and 34,280 transcripts with significant homology to 10,209 *P. infestans* proteins. The remaining transcripts, with no

significant similarity to these *Phytophthora* proteins, were annotated to the NCBI non redundant protein database (E-value >1x10$^{-5}$).  Investigation of the origin of the 21,881 transcripts showing significant similarity revealed that approximately 42 % of the genera belonged to the domain bacteria and 8.7 % to the domain fungi (Online Resource 6).  These transcripts (likely originating from other organisms present in the river water used to induce sporulation) were regarded as contaminant sequences and were excluded.

**Table 1.  Number of transcripts with significant amino acid sequence similarity to proteins predicted in the genomes of *Phytophthora infestans* and *Phytophthora cinnamomi* (E-value <1x10$^{-5}$)**

| | Total | *P. infestans* proteins[a] | *P. cinnamomi* proteins (CBS 144.22)[b] | Similarity to both *Pc* and *Pi* proteins |
|---|---|---|---|---|
| **Number of transcripts** | 72,220 | 34,280 | 36,500 | 33,246 |
| **Percentage transcripts with significant blastx hit (%)** | - | 47.5 | 50.5 | 46 |

[a]*Phytophthora infestans* Sequencing Project, Broad Institute of Harvard and MIT; http://www.broadinstitute.org/ downloaded February, 2013

[b]*P. cinnamomi* var. *cinnamomi* v1.0; http://genome.jgi.doe.gov/Phyci1/Phyci1.home.html, downloaded March, 2013

Over 80% of the transcripts in the remaining dataset were 100-500 nt in length (mean length = 444 nt, N50 = 348 nt) (Fig. 1).  The functional annotations based on homology to the *P. infestans* gene models (representing 53% of the transcripts) were used to infer function.  A large percentage of the 9,583 core *Phytophthora* orthologous genes (80.4%), predicted by Haas et al. (2009) was represented in this *P. cinnamomi* cyst and germinating cyst dataset. Approximately 36% of the dataset were annotated as hypothetical proteins.
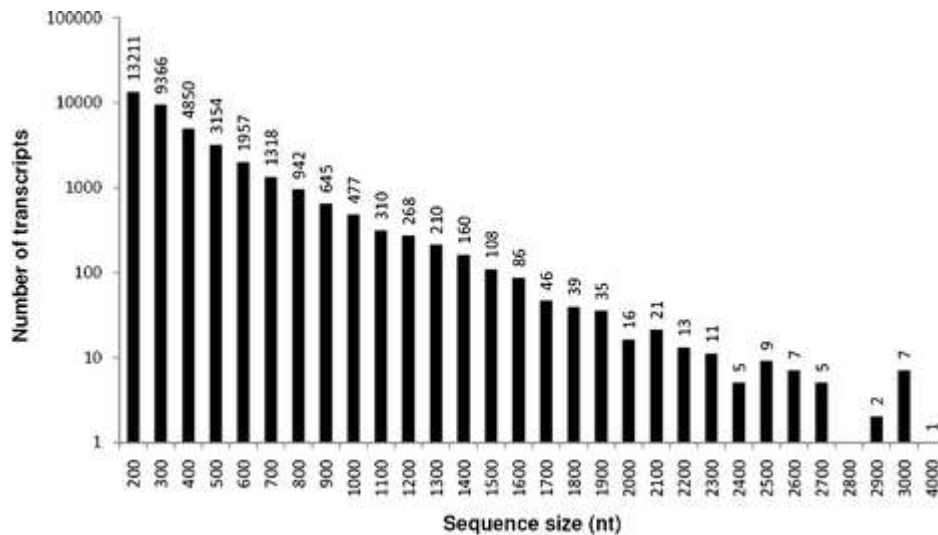
**Fig. 1    Length distribution of transcripts from the *Phytophthora cinnamomi* CNVDB29 cyst and germinating cyst library**

A total of 8.7% (3,254 transcripts) of annotated transcripts did not show significant similarity to *P. infestans* protein models.  Putative functional annotations were obtained for 1,965 of the 3,254 transcripts by using the EuKaryotic Orthologous Groups (KOG) annotations assigned to the proteins predicted in the *P. cinnamomi* genome (Joint Genome Institute database (http://www.jgi.doe.gov/,downloaded March, 2013)).

**Functional enrichment analysis**

The COG database provided functional annotations for 48.4% of the 37,534 transcripts that showed similarity to proteins in the genomes of *P. infestans* and *P. cinnamomi* (CBS 144.22).  Enrichment analysis divided the transcripts into 26 functional categories (Fig. 2). The function of 771 transcripts with homology to proteins in the COG database was unknown.  Highly represented classes included transcripts involved in translation (21.7%), post translational modification (10.2%) and amino acid transport and metabolism (9.7%).  A total of 30.8% of transcripts were dedicated to transport and metabolizing compounds (including amino acids, carbohydrates, lipids, nucleotides, coenzymes and inorganic ions). Transcripts that play a role in replication and cell wall biogenesis (10.3%) are consistent with

the process of cyst germination. Functional categories that could potentially contribute to pathogenesis (1,457 transcripts) include signal transduction mechanisms (4.3%), secretion (2%) and defence mechanisms (1.3%). The defence mechanism category was mainly constituted of transcripts involved in transport (215 transcripts).
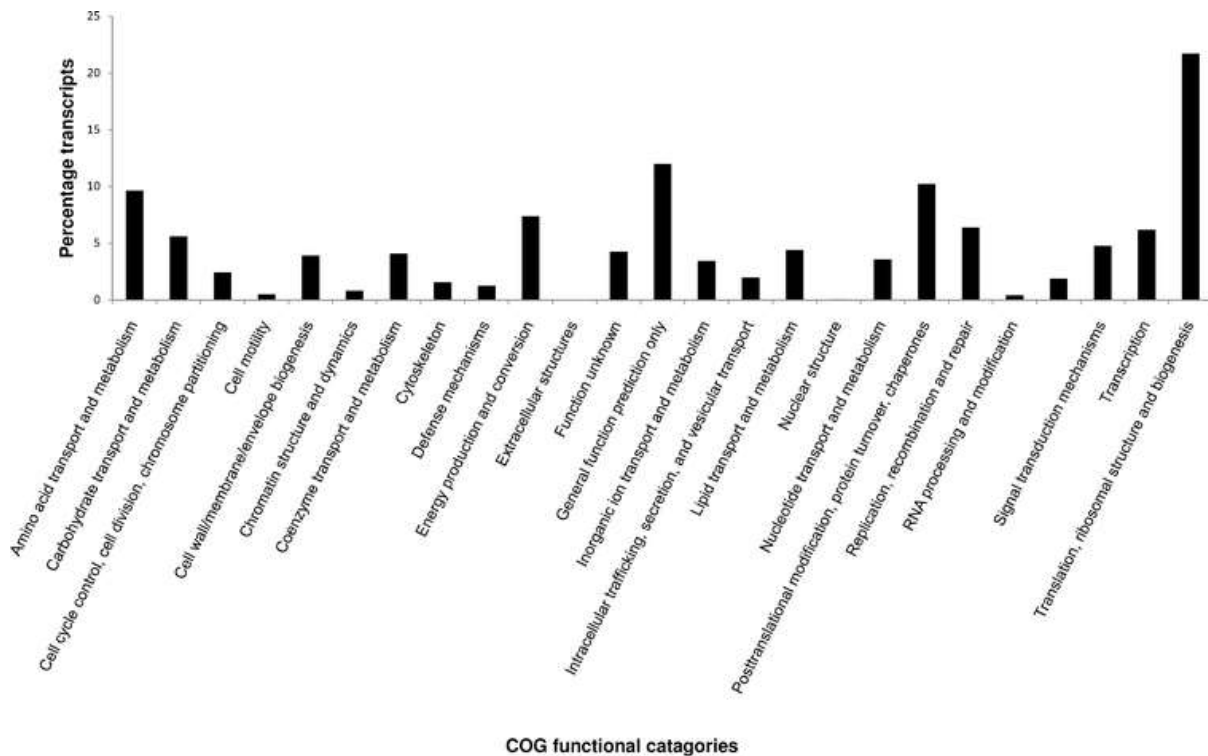


**Fig. 2 Functional classification of *Phytophthora* specific transcripts according to the Clusters of orthologous groups (COG) database** Each bar represents the percentage transcripts that correspond to each of the 26 functional classes of the COG database. A single transcript may be represented in more than one functional class.

**Putative secreted proteins**

The start codon (methionine) could be identified in the predicted open reading frame of 2,245 of the 37,534 transcripts. A total of 110 transcripts contained a predicted signal peptide (and lacked TMHs) (listed in Online Resource 7). Ninety nine transcripts had an assigned function based on homology to the *P. infestans* protein models and included nine conserved hypothetical proteins and 14 transcripts that were identified as pathogenicity

14

genes in this study (discussed later in Table 2). Some transcripts with predicted signal peptides did not show homology to *P. infestans* but only to *P. cinnamomi* proteins (U12323, U15167, U19168, U24069, U32988, U37984, U40431, U5155, U54017, U55798 and U58247).

**Table 2. Transcripts with a putative role in *Phytophthora cinnamomi* pathogenicity[a]**

| Functional categories of putative pathogenicity transcripts[b] | Number of transcripts | |
|---|---|---|
| | With similarity to *Pi* genome protein models[c] | With similarity to *Pc* (CBS 144.22) but not *Pi* genome protein models[d] |
| **Adhesion** | **39** | **16** |
| CBEL protein | 7 | - |
| Adhesins | 9 | - |
| Mucins | 23 | - |
| C-type lectin | - | 16 |
| **Hydrolases** | **123** | **28** |
| Cutinases | 3 | - |
| Glucosidases | 51 | - |
| Glucanases | 55 | - |
| Polygalacturonases | 11 | - |
| Cellulases | - | - |
| Chitinases | 3 | 21 |
| Ubiquitin proteases | - | 7 |
| **Inhibitors of proteases** | **18** | **9** |
| **Effectors** | **57** | **-** |
| RXLR effector genes[e] | - | - |
| Elicitins and elicitin-like | 20 | - |
| Crinkling and necrosis-inducing proteins | 28 | - |
| Necrosis-inducing *Phytophthora* protein 1 | 9 | - |
| **Small cysteine-rich proteins** | **4** | **-** |
| **Detoxification, drug resistance, and metabolite transport** | **538** | **7** |
| Cytochrome P450 | 25 | - |
| Putative ABC transporters | 513 | 7 |
| **Protection against oxidative stress** | **499** | **23** |
| Glutathione s-transferases | 25 | 3 |
| Peroxidases | 68 | 4 |
| Reductases | 360 | 13 |
| Superoxide dismutases | 34 | - |
| Glutaredoxins | 12 | - |
| Glutathione synthetases | 0 | 3 |
| **Signal transduction and regulation** | **114** | **9** |
| Mitogen-activated protein kinases | 72 | - |
| Myb-like proteins | 42 | 7 |

| | | |
|---|---|---|
| Cell surface glycoproteins | - | 2 |
| **Virulence genes** | **2** | **-** |
| Virulence | 2 | - |
| **TOTAL** | **1,394** | **92** |

[a]Pathogenicity gene categories modified from Torto-Alalibo et al. (2007)(Torto-Alalibo et al. 2007)

[b]Functional annotations are based on protein sequence homology and not on the presence of conserved motifs.

[c]Functional annotations are based on protein sequence homology to *P. infestans* protein models (blastx analysis (E-value <1x10$^{-5}$)) and not on the presence of conserved motifs.

[d]Transcripts lacking significant similarity to *P. infestans* proteins, but showing significant similarity to models of the *P. cinnamomi* (CBS 144.22) genome (blastx analysis (E-value <1x10$^{-5}$)) were considered. KOG annotations assigned to protein models of the *P. cinnamomi* (CBS 144.22) genome were used to infer function.

[e]Functional annotation based on HMM, discussed separately elsewhere.


**Putative pathogenicity related genes identified from the cyst and germinating cyst cDNA library of *P. cinnamomi***

The annotated *P. cinnamomi* CNVDB29 dataset was manually screened for specific pathogenicity candidate genes as reported in the literature (Table 2). A total of 1,394 putative pathogenicity related transcripts were identified (a full list of the transcripts are provided under Online Resource 8). The most represented gene categories were genes involved in detoxification and the protection against oxidative stress, followed by hydrolases. We speculate that the pathogen induces these genes in response to being under stress during the *in vitro* production of cysts. Effector genes were also highly abundant, representing 57 transcripts. Transcripts annotated as cellulases were absent.


A total of 44 putative RXLR effector transcripts were identified using computational methods, with 19 sequences harbouring a predicted signal peptide (Online Resource 9). Five of the 44 sequences, although predicted by the HMM, did not contain an actual RXLR sequence. The full length amino acid sequences were determined for 25 of the 44 putative RXLR transcripts, by performing a blastp analysis (E-value <1x10$^{-5}$) (Online Resource 9) on the translated putative RXLR sequences (*P. cinnamomi* CNVDB29) to the protein models

predicted in the genome of *P. cinnamomi* var. *cinnamomi* v1.0 (isolate CBS 144.22) http://genome.jgi.doe.gov/, downloaded March, 2013).

**Putative RXLR gene homologues**

Full length copies of 25 predicted RXLR proteins of *P. cinnamomi* (CBS 144.22), with expressed homologues in *P. cinnamomi* CNVDB29 from this study, were compared to predicted RXLR proteins of four *Phytophthora* species from four different clades (Blair et al. 2008) as well as another oomycete (*H. arabidopsidis*) using blastp analysis. The alignment statistics for the six full length *P. cinnamomi* putative RXLR proteins that met the requirements (E-value $<1\times10^{-8}$, identity >30%) are listed in Table 3. *P. cinnamomi* predicted RXLR protein Phyci1-16230 has 66% amino acids identical to *P. sojae* suppressor of necrosis 1-like protein 2 (SNEL2) (Fig. 3a). PsSNEL1, an ortholog of SNEL2 is also included in the figure. Phyci1-30885 shares 50 and 33% identical amino acids with *P. sojae* and *P. ramorum* proteins PsAvh42 and PrAvh77, respectively (Fig. 3b). PrAvh77 does not contain an RXLR motif. The protein models predicted in the *P. cinnamomi* genome (CBS 144.22) were not always identical to the proteins predicted from the transcriptomic data (CNVDB29) as in the case of Phyci1-172110, Phyci1-292291 and Phyci1-246465. This may be due to incorrect gene models in the genome of *P. cinnamomi* and incomplete transcripts from the transcriptomic data. Phyci1-97937 shares the RXLR-EER motif with *P. sojae* Avh20 (36% amino acid identity), but lacks a SP (Fig. 3d). The contig from which the RXLR transcript was originally predicted in *P. cinnamomi* CNVDB29 (C112596), however does contain a SP.

**Table 3.** **Protein sequence similarity between full length copies of putative RXLR proteins of** *Phytophthora cinnamomi* **(CBS 144.22) and other** *Phytophthora* **species as determined through blastp analysis (E-value <1x10[-8])[a]**

| Query ID[a] | Hit ID[bc] | Hit annotation[c] | E-value | Score (Bit) | Query length (aa) | Hit protein length (aa) | % Identity | % Positive | % Gaps | Hit species |
|---|---|---|---|---|---|---|---|---|---|---|
| Phyci1-16230 | Ps298800 | PsSNEL2 | 2.22E-10 | 57 | 158 | 189 | 66 | 71 | 16 | *P. sojae* |
| Phyci1-30885 | Ps286041 | PsAvh42 | 1.04E-21 | 95 | 160 | 153 | 50 | 67 | 11 | *P. sojae* |
| | Pr97272 | PrAvh77 | 7.74E-10 | 55 | 160 | 154 | 33 | 56 | 13 | *P. ramorum* |
| C15073[d] (Phyci1-172110) | PITG14203 | RXLR effector | 1.84E-20 | 91 | 155 | 153 | 42 | 56 | 8 | *P. infestans* |
| | Pr97517 | PrAvh_322 | 1.03E-13 | 68 | 155 | 151 | 36 | 56 | 7 | *P. ramorum* |
| Phyci1-97937 | Ps286380 | PsAvh20 | 2.56E-21 | 93 | 150 | 162 | 36 | 53 | 22 | *P. sojae* |
| Phyci1-292291 | Ps540011 | PsRxL508 | 0 | 654 | 840 | 389 | 92 | 94 | 54 | *P. sojae* |
| Phyci1-246465 | Ps531198 | PsAvh448 | 4.05E-28 | 120 | 2,041 | 115 | 58 | 78 | 94 | *P. sojae* |
| | Pr97565 | PrAvh387 | 1.47E-11 | 65 | 2,041 | 139 | 40 | 57 | 93 | *P. ramorum* |

% Identity – number of aa of query identical to hit, divided by the total number of aa of the hit, % Positive – number of conserved aa substitutions of query, divided by the total number of aa of the hit, % Gaps – percentage of gaps in the query and the hit sequence

[a]Full length copies of RXLR proteins predicted in *Phytophthora cinnamomi* CNVDB29, were obtained via blastp analysis (E-value <1x10[-8]) of translated putative RXLR transcripts of CNVDB29 to predicted proteins in the *Phytophthora cinnamomi* genome CBS 144.22, ID numbers Phyci1-xxxx.

[b]Only the single top scoring hit (according to E-value) for each species was considered, although multiple proteins from a single species may have shown significant similarity to the query sequence.

[c]Protein ID numbers and annotation from FungiDB version 2.3 (March, 2014)

[d]Cxxxx refers to translated RXLR-motif containing contigs as identified from *P. cinnamomi* CNVDB29 transcripts.

## a Phyci1-16230

```
P. cinnamomi : MKLSYLIAFAAVVVASTAVPASASTGLTTNNLAEDFQVAPELAALRSLRGANQLS---------------S :  56
PsSNEL2      : MKLSYLIAFAAVVVASTAAPASASTGLTTTNLAEDVQLAPELGALRSLRGASQDG---------------S :  56
PsSNEL1      : MKLSYIITFAAVVVASTTAPATAATGLTASNLAEKLQVGRQLGAGGGLRGASQECDQSQNQNQNKPSTGNTS :  72

P. cinnamomi : KGDDKKPKGEHGDKKKGG-----DKKKDGDKKKDG----DKKKK-EKKEKKEKRKEEEREKKHGKKDGDKKHGK : 120
PsSNEL2      : KGDEKKPTEHGDKKKDD-----EKKKGSDNKKDKDDKER-EKKEKKEKKDDKEKEKKEKKEKKEKKKKKQE : 123
PsSNEL1      : QDDGKQQGQDDGDKKKDDGDKKEKKDDGDHKQEEKKDDKKDDGDKKDNKADKEKKQQEKKKEKQEKKEQQKKE : 144

P. cinnamomi : KDGDKKHGKKD-------------------------------GEKKHGDKKHEQKDGKKDHKKTDGSKK : 158
PsSNEL2      : KEDKKKDGKKDKHGDTKDGKKHDDKKEKKHDDKKEKKHDDKKHGDKKHEQKEGKKDHKKTDGSSK : 189
PsSNEL1      : KEDKKKSDNKDKDC--FKASAEYLVGCRSVSIHRRPVARGLLGHLSRQFRRS--------------- : 194
```

## b Phyci1- 30885

```
P. cinnamomi : MRVLLWVLLVALVTLLSSTDALSTNNSDKKQVVQPNS------EEVATRMLAANYENN-NDKRMLRGESKMT :  65
P. sojae     : MRLLLWVLLATLVTLLSSADALSITNSDKKQVSVLDS-----GKTATSRMLTADYELT-NVKRSLRGESEKT :  66
P. ramorum   : MRLIFWALIVTLISLLSSGEAAST---DKKQMAQLDSDIGFLSRDSERDLREGTKKIADDDDDELAADDELA :  69

P. cinnamomi : YATNAENNEERA---VFIKFQGSVQRLREKLRDINPESTKNIEKRFQK--LADKGHTPDYYLETYQIGTFNSR : 133
P. sojae     : -----KQDEERMGNFLVKIKGSIQKLKDAYR-------AHIEKQFVLDLIDK-KTPDDLYKEYRRAPWGVS : 125
P. ramorum   : -----ADDEERA--IMVKIRGSLSKLRAKYLK---NKMNRLTPRFEK--LVKDGNSLDDVKKMYRIGGWGTP : 129

P. cinnamomi : HWNRRFYKKYEEWYKRTHPDWVSEITK- : 160
P. sojae     : KLKKKAYKGYAAYYKREHPNWVSSIPKK : 153
P. ramorum   : SRNKRFYKLFEEWVKTNYPN---GLPKQ : 154
```

## c Phyci1-172110

```
P. cinnamomi : ------------------------------------------------------NGADDEERA-W :  10
C15073       : MRLCLVLLAVTATLLSSCDAAPSSTN---LLQTKTQTEKLRTGSDRDSRLFRTQETSTANTNGADDEERA-W :  68
P. infestans : MRLSCVLLIALVALLSSFDQVSTKTSDDALTMKALRQGNAVATVTNNKRLLQTELPGSSA-ILEDNEERASW :  71
P. ramorum   : YQRSCILVIMVAMMITCCSAMDPKMK----TQNMVQTDSVPVNDRNGYRFLRSYKTEEAG-DATDNEDRA-L :  66

P. cinnamomi : GLKTLLSLVK-TPKPNDFAKVDLDRMLVDGVFKLEMFERWNK-YSTEKILATLDLSKRKNIPYGSILQEFLN :  80
C15073       : GLKTLLSLVK-TPKPNDFAKVDLDRMLVDGVFKLEMFERWNK-YSTEKILATLDLSKRKNIPYGSILQEFLN : 138
P. infestans : GLKKLVNLFMFKGASSNFSTAKLTKMLSDPAFKSEMFKTWNAKYSTEKVIAILDLSKKKNIPYGPMLQDFRN : 143
P. ramorum   : DLKALKSFVRIKKTENDFSKRYLERMMIDPKFKKEMIEKWGKVYTQEKVISKL----KRKD---PAMLLDFLN : 132

P. cinnamomi : -TYRLRQGAANLRRGN--- :  95
C15073       : -TYRLRQGAANLRRGNQ-A : 155
P. infestans : -TYGRRPIKP--------Q : 153
P. ramorum   : GNTALRPLIKVERPQNLGA : 151
```

## d Phyci1-97937

```
P. cinnamomi : MHLPSTLVIVAVALFASGHAVTDLSQTKFSQSL--GKLHPEVRRVTGDPGDGNRLLRSATTATEDTDEERA :  70
P. sojae     : MRLSATLLVVAAAGLFATGHAASDVSQPTIPQSLDVGQLDPGAETVT-----PKRLLRA---VCDDNEEERG :  64

P. cinnamomi : LWTELKARFPGTIENMMKNTKKERERIKEQALVFAKFLKQEELPYNVVVFYMRQGYSNSKAFKMGLRYETYK : 142
P. sojae     : AWSALLSRFPGTAKNLAREAEKQTAKIKLYADDFATYRANGDVPYEVTVAYMVKGYSDSRAVN------TLR : 130

P. cinnamomi : ENP-----------------------WLYH-Q : 150
P. sojae     : DLPGKPISVSLIQVGRKAKCCELETCLYQHLA : 162
```

**Fig. 3 Alignment of putative RXLR proteins from *Phytophthora cinnamomi* to putative homologous RXLR proteins from *Phytophthora* species** Shaded amino acids indicate amino acids with 100% conservation

(black) or only 80% conservation between two or three of the sequences (grey)(GeneDoc Version 2.7.000, conservation mode with similarity groups enabled). RXLR-motifs (Arg-any amino acid-Leu-Arg) are indicated with blocks. a) Phyci1-16230 (CBS 144.22) aligned to SNEL2 and SNEL1 from *P. sojae* b) Phyci1-30885 (CBS 144.22) aligned to PsAvh42 and PrAvh77 c) Phyci1-172110 (CBS 144.22) aligned to contig C15073 (CNVDB29), PITG14203 and Phyra-97517 d) Phyci1-97937 (CBS 144.22) aligned to PsAvh20 e) The shortened Phyci1-292291 (CBS 144.22) sequence aligned to contig C17530 (CNVDB29) and Physo-540011. Asterisks indicate the suggested start sites for the predicted RXLR transcripts from *P. cinnamomi* isolates CBS 144.22 and CNVDB29 and the putative homologue from *P. sojae.* f) Phyci1-246465 (CBS 144.22) aligned to contig C83210 (CNVDB29), PsAvh448 and PrAvh387.

**Gene expression profiling of target putative pathogenicity genes**

Reverse transcriptase qPCR was used to assess the expression pattern of seven putative pathogenicity genes of *P. cinnamomi*. This included one RXLR effector gene (C83210, homologue of Phyci1-246465 discussed above), one necrosis-inducing *Phytophthora* protein 1 (NPP1) (U15350), an elicitin-like gene (U4987), a polygalactorunase gene (U6659) and three genes involved in pathogen adhesion (U6179, U6809 and U8949) (Fig. 4).

Six of the seven genes assessed, were significantly upregulated in the zoospore developmental stage in comparison to the mycelial phase (Fig. 4). This included the effector genes (RXLR, NPP1 and elicitin-like), the polygalactorunase gene, one cellulose binding and elicitor lectin (CBEL) gene and an adhesin-like gene. The RXLR gene, the elicitin-like gene and the adhesion-like gene remained significantly induced in the cyst and germinating cyst phase relative to mycelia, although their expression levels were significantly lower than in the zoospore developmental stage.
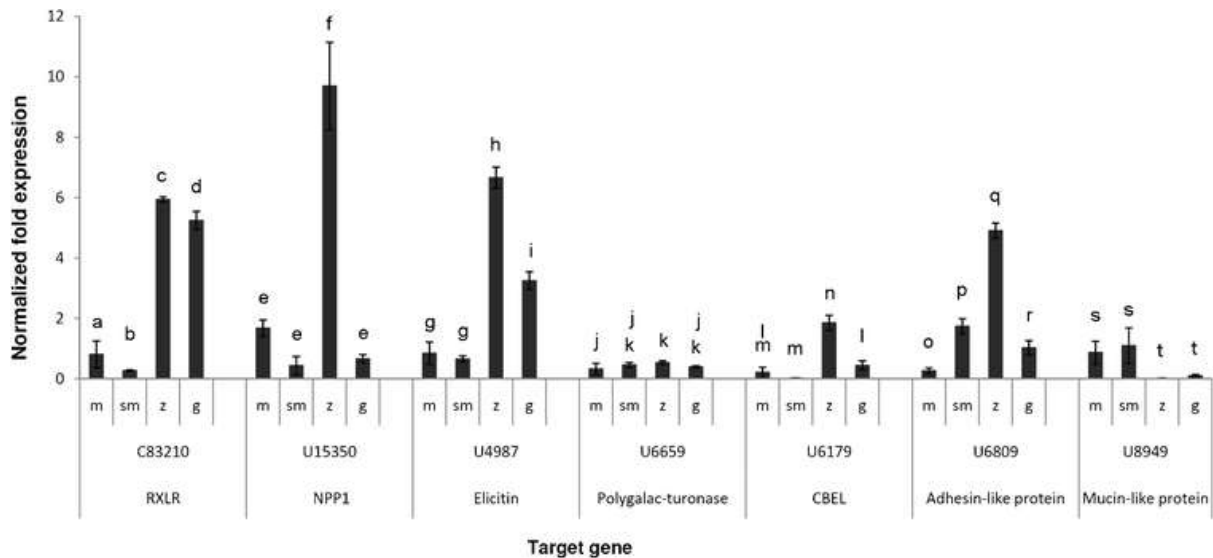
**Fig. 4 Expression analysis of target putative pathogenicity genes across *in vitro* developmental stages of *Phytophthora cinnamomi*** Gene expression levels of each target gene were calculated using three endogenous control genes (WS041, Ubc, β-tubulin) to normalize the data. Error bars represent standard deviation and are indicated for each data point. Analysis of variance (ANOVA) followed by a Least Square Means Differences Student's t-test ($p > 0.05$) was performed considering each target gene individually and does not allow comparison of gene expression levels between genes. For each individual target gene, bars containing the same letters do not differ significantly ($p < 0.05$). m – mycelia, sm – sporulating mycelia, z – zoospore, g – cysts and germinating cyst, NPP1 – necrosis-inducing *Phytophthora* protein 1, CBEL – cellulose binding elicitor lectin

The expression level of C83210 (RXLR-motif containing transcript) showed a 7.3 and 6.5 fold increase in zoospores and germinating cysts respectively, relative to expression levels in the mycelia, while expression was significantly lower in sporulating mycelia (Fig. 4). The NPP1 transcript accumulated to similar levels in mycelia, sporulating mycelia and cysts and germinating cysts, with a highly significant 5.8 fold induction (relative to mycelia) in the zoospore developmental stage. U4987 (elicitin-like gene) was significantly induced in the zoospore and cysts and germinating cyst phase (7.9 and 3.8 fold increase relative to the mycelia phase), with no noteworthy difference between expression levels in the mycelia and sporulating mycelia. The relative expression of the cell wall degrading enzyme, polygalacturonase (U6659) did not vary significantly between the developmental stages, except for a significant 1.6 fold induction observed in the zoospore phase, relative to

mycelia. The CBEL transcript was significantly upregulated (8.8 fold) in zoospores relative to mycelia, with cysts and germinating cysts having similar expression levels than observed in mycelia. An 18.7 fold increase was observed in the relative expression level of U6809 (the adhesin-like gene) in zoospores and a 6.6 fold increase in sporulating mycelia, relative to mycelia. U6809 transcript abundance was significantly higher in germinating cysts (3.9 fold induction) than in mycelia. Lastly, the mucin-like transcript (U8949) was significantly less abundant in zoospores and cyst and germinating cyst phases relative to mycelia.

## Discussion

To our knowledge, this is the first investigation of pathogenicity genes in the ubiquitous pathogen *P. cinnamomi*. Expression analysis of *Phytophthora* pathogenicity-related genes *in planta*, demonstrated that many pathogenicity genes expressed during infection, could also be detected in the pre-infection stages (Ye et al. 2011; Chen et al. 2013; Chen et al. 2014). Genes upregulated in germinating cysts of several other *Phytophthora* spp. has shed light on the molecular mechanisms underpinning infection (Shan et al. 2004; Chen et al. 2011; Ye et al. 2011; Chen et al. 2013; Chen et al. 2014; Hosseini et al. 2015). Accordingly, we investigated the transcriptome of the cyst and germinating cyst stage of *P. cinnamomi* and identified 1,534 putative pathogenicity transcripts. This included transcripts with homology to gene families implicated in pathogenesis in the literature (1,394 transcripts), 96 putative secreted transcripts (excluding those also listed in Table 2) and 44 transcripts containing an RXLR-motif.

Based on protein homology, various apoplastic effectors, fulfilling different functions in infection, were identified from the germinating cyst phase of *P. cinnamomi*. Protease inhibitors, representing 27 transcripts in *P. cinnamomi* (Table 2), are employed by many pathogens as a counter defence to host enzyme attack (Misas-Villamil and van der Hoorn 2008). Elicitin proteins (representing 20 transcripts) are suggested to function in sterol

acquisition in *Phytophthora* species (Vauthrin et al. 1999; Jiang et al. 2006) and recognition of elicitin proteins by the host defence system induces a hypersensitive response (HR) in some plant species (Kamoun et al. 1993), associating this family with pathogenesis. The exact role of small cysteine-rich proteins (representing four transcripts) during infection remains to be determined, though some small cysteine-rich proteins from *Phytophthora cactorum* have been shown to trigger host cell death in the Solanaceae, including *Nicotiana benthamiana* and tomato and are implicated to function in pathogenicity and oxidative stress tolerance (Chen et al. 2015). Genes that aid in pathogen adhesion include CBEL proteins (seven transcripts), adhesins (nine transcripts) and mucin transcripts (23 transcripts). CBEL proteins, found in the cell walls of oomycete species (Séjalon-Delmas et al. 1997), bind to cellulose components of the host cell wall (Gaulin et al. 2002) and act as PAMPs, as recognition of their cellulose binding domains lead to activation of the HR in tobacco plants (Gaulin et al. 2006).

Cytoplasmic effectors present in the dataset include members of the Crinkling and necrosis-inducing (CRN) protein family and the RXLR effector family (characterized by a conserved amino acid sequence Arg-any amino acid-Leu-Arg). The CRN effector family (representing 28 transcripts) is expanded in both hemi-biotrophic and necrotrophic oomycete pathogens (Stam et al. 2013). Characterization of *P. capsici* CRN effectors indicated that most CRN domains do not elicit cell death in plants, but their regulated expression patterns and localization to the host nucleus suggest that they are functionally relevant during infection even if not directly in pathogenesis (Stam et al. 2013). The importance of CRN effectors are further supported by the specific phosphorylation patterns observed in conserved regions of CRN members (Resjö et al. 2014) suggesting post translational regulation of this family.

Functional characterization of oomycete RXLR proteins through gene silencing, have previously illustrated that specific RXLR proteins are essential for virulence (Bos et al. 2010). From a single developmental stage 44 putative RXLR effectors were identified. As RXLR

effector genes were identified in the genomes of all plant pathogenic oomycete species with a biotrophic phase (Jiang et al. 2008; Haas et al. 2009; Baxter et al. 2010;Lamour et al. 2012), their presence in *P. cinnamomi* was not unexpected. Functional characterization of RXLR genes that are shared by various pathogenic oomycetes may suggest common mechanisms that enable these oomycetes to infect living hosts. Except for an N-terminal signal peptide and the RXLR-motif, little sequence conservation remains between orthologs and paralogs of the RXLR family (Jiang et al. 2008; Haas et al. 2009). This lack of similarity accommodates their diverse functions (dependent on the C-terminal amino acid sequence (Win et al. 2007; Jiang et al. 2008)) and most likely results from positive selection to remain undetected by the host surveillance system (Tyler et al. 2006; Win et al. 2007; Jiang et al. 2008). Putative homologous proteins were identified for six of the 44 predicted RXLR transcripts of *P. cinnamomi*, because they shared significant protein sequence similarity to RXLR proteins from other oomycete species. Phyci1-16230 shares 66% identical amino acids with the SNEL2 from *P. sojae.* SNEL2 is a paralog of PsSNEL1, which in turn is an ortholog of suppressor of necrosis 1 protein (SNE1) of *P. infestans* (Wang et al. 2011). SNE1 was shown to be localized in the host nucleus (tomato) and has the ability to suppress cell death induced by effector recognition or pathogen induced necrosis during the necrotrophic phase of *P. infestans* (Kelley et al. 2010). It will be interesting to see if the protein produced from Phyci1-16230 is indeed an ortholog of SNE1 and functionally similar. Cluster analysis (including construction of phylogenetic gene trees) and evaluation of synteny between genomic regions (where the genes are located) should be completed before the RXLR genes identified in *P. cinnamomi* are confirmed as orthologs.

Except for effector molecules, transcripts representing other infection related gene families were also investigated. Necrosis-inducing proteins (NIPs) (represented by nine necrosis-inducing *Phytophthora* protein 1 (NPP1) transcripts in this study) are suggested to facilitate the switch from biotrophy to necrotrophy in hemi-biotrophic pathogens. Confusion exists regarding the trophic lifestyle of *P. cinnamomi* as it has been described as a necrotroph, a

hemi-biotroph (Cahill et al. 2008) and suggested to have alternating biotrophic and necrotrophic phases (depending on environmental conditions) (Shearer and Crane 2012). *In planta* expression analysis of *P. cinnamomi* NIPs, together with investigation of changes in pathogen metabolism, may shed light on the switch between trophic phases. Hydrolases (representing over 900 transcripts) are secreted by the pathogen to break down components of the plant cell wall (a physical barrier for pathogen entry) (Have et al. 2002) and included polygalactorunases, glucanases and glucosidases.

Putative members of the *P. cinnamomi* secretome were extracted from the dataset (110 transcripts) (Online Resource 5), as any secreted protein has the potential to directly interact with host proteins and act as a pathogenicity factor. Several of the transcripts encoding putative secreted proteins were also identified as potential pathogenicity factors according to their functional annotation as listed in Table 2. Secreted transcripts specific to *P. cinnamomi* (11) and transcripts with unknown function (9) are potential interesting candidates for further investigation. Note that transcripts do not represent full length mRNA sequences and the complete protein may contain TMH in the region that was not sequenced that would retain the protein in membranes.

The expression patterns of seven putative pathogenicity genes were evaluated across the *in vitro* developmental stages of *P. cinnamomi*. In theory genes are upregulated directly prior to the stage where their functional products are required (Judelson et al. 2008). This implies a functional role for the RXLR-, elicitin-like and adhesion-like proteins (upregulated in cysts and germinating cysts) during pre-infection and early infection and the NPP1, polygalactorunase and CBEL in the cysts phase (as they are induced in zoospores).

Transcript levels of putative RXLR transcript C83210, homologous to *P. cinnamomi* (CBS 144.22) Phyci1-246465, *P. sojae* Avh448 and Avh387 of *P. ramorum*, accumulated to higher levels in zoospores and germinating cysts than in *P. cinnamomi* mycelia, consistent with the expression pattern observed for most RXLR genes from oomycete species (Judelson et al.

2008; Chen et al. 2011; Ye et al. 2011). It is interesting that the NPP1-like (U15350) transcript is upregulated during the zoospore phase of *P. cinnamomi*, as NIPs are thought to facilitate the switch between trophic phases in hemi-biotrophic pathogens. An antagonistic interaction, between NIPs and suppressor of necrosis protein 1 (SNE1) is suggested to regulate the switch between trophic phases in *P. infestans* (Kelley et al. 2010) and may explain the early induction of NIP observed in pre-infection structures in other oomycetes (Ye et al. 2011). U4987, an elicitin-like gene that shows homology to a *P. infestans* elicitin protein from clade 3a (INL3a-like protein) was more abundant in zoospore and germinating cyst phases than in mycelia, supporting the pattern observed for clade 3 elicitins from *P. infestans*, *P. sojae* and *P. capsici* (Qutob et al. 2003; Jiang et al. 2006; Chen et al. 2013). The *P. cinnamomi* polygalacturonase (U6659) expression profile showed a significant increase in transcript abundance in zoospores relative to mycelia, suggesting that the product may be functionally active in cysts, aiding in plant cell wall hydrolysis once the pathogen is in contact with the host surface. The CBEL protein transcript (U6179) from *P. cinnamomi* accumulated to highest levels in the zoospore phase, supporting the role of CBEL proteins recognizing and binding to cellulose in plant cell walls (Gaulin et al. 2002) prior to zoospore encystment. Another gene involved in pathogen adhesion is the adhesin-like gene (U6809) that was more abundant in sporulating mycelia than in germinating cysts. In contrast to this, the relative expression level of transcript U8949 (mucin-like protein) was higher in mycelia and sporulating mycelia than in zoospores and germinating cysts, suggesting that it may rather act as a protective agent than in facilitating pathogen adhesion (Perez et al. 1999).

Although the same infection strategy is used by most *Phytophthora* species (release of motile zoospores that locate suitable infection sites, zoospore encystment, germination and penetration of host tissue), some differences exist that would impact on pathogenicity gene expression patterns. At higher temperatures caducous sporangia produced by some *Phytophthora* species germinate directly, while indirect germination only occurs at lower

temperatures by the formation and release of zoospores (Rao 1970). Judelson and Tani (2007) reported zoospore release from only 58% of *P. infestans* detached sporangia under laboratory conditions (Judelson and Tani 2007). This would imply that genes necessary for adhesion, germ tube formation and tissue penetration (CWDE) would be expressed in germinating sporangia rather than in cysts, as is expected in non-caducous species (such as *P. cinnamomi*).

Six of the seven pathogenicity gene groups whose expression patterns were assessed in *P. cinnamomi* in this study (excluding the adhesion-like gene), were found to be expressed during infection of soybean by *P. sojae* (Ye et al. 2011) and all three effector gene groups (RXLR, NPP1 and elicitin-like genes) during infection of *N. benthamiana* by *P. cactorum* (Chen et al. 2013; Chen et al. 2014). Although gene expression was only assessed during the *in vitro* developmental stages of *P. cinnamomi*, the presence of these transcripts during *in planta* infection in other *Phytophthora* model systems, increases the confidence that they may act as pathogenicity factors.

The *P. cinnamomi* transcriptomic library will make a significant contribution to the available molecular toolbox to study *P. cinnamomi* in the future. The experimental approach used in this study, however, had some shortcomings. Due to the nature of the sample constitution in this study specifically (pooling total RNA extracted from germinating cysts from multiple experiments), we could not utilize this data to capture gene expression levels, which is one of the main advantages offered by RNA-Seq analysis. With the *P. cinnamomi* genome recently published (Studholme et al. 2016), differential gene expression analysis *in planta* is currently underway in our laboratory and will enable a better understanding of pathogenesis. The *P. cinnamomi* dataset was annotated with a homology based approach which did not allow the identification of novel transcripts. This dataset, however still holds the potential to identify novel transcripts by identifying conserved motifs within transcripts lacking significant sequence similarity to known proteins.

In summary, using RNA-Seq and a *de novo* assembly approach, we successfully identified over 1,300 putative pathogenicity unigenes from the cyst and germinating cyst phase of *P. cinnamomi* that included genes, similar to those implicated in pathogenicity in other *Phytophthora* species. Among the transcripts identified, were a range of effector genes (including RXLR effectors) that serve as candidate genes for further investigation. The expression patterns of seven putative pathogenicity related genes were evaluated across four *in vitro* developmental stages. The majority of the genes assessed were highly induced in the zoospore developmental stage, indicating that they are functionally important in the subsequent phase, namely cysts and may be involved in the initial infection process. Analysis of individual gene expression profiles revealed similar expression patterns as indicated for other oomycete species. The transcriptomic data generated in this study, together with *P. cinnamomi* genomic data, is a valuable tool that can be utilized to investigate *P. cinnamomi* genes *in planta* and to study alternative splicing within transcripts. This information serves as a basis to start unravelling *P. cinnamomi* phytopathology.

**Author contributions statement**

NvdB and A.R conceived and designed the experiments. A.R conducted the experiments, analysed the data and wrote the manuscript. NvdB and D.K.B provided guidance with data analysis and revised the manuscript. All authors reviewed the manuscript.

**Online Resources**

Supplementary information accompanies this paper at http://link.springer.com/.

Competing financial interests: The authors declare no competing financial interests or conflict of interest.

# References

Baxter, L., Tripathy, S., Ishaque, N., Boot, N., Cabral, A., Kemen, E., et al. (2010). Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science, 330*(6010), 1549-1551.

Bendtsen, J. D., Nielsen, H., von Heijne, G., & Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology, 340*(4), 783-795.

Benhamou, N., & Nicole, M. (1999). Cell biology of plant immunization against microbial infection: The potential of induced resistance in controlling plant diseases. *Plant Physiology and Biochemistry, 37*(10), 703-719.

Blair, J. E., Coffey, M. D., Park, S., Geiser, D. M., & Kang, S. (2008). A multi-locus phylogeny for *Phytophthora* utilizing markers derived from complete genome sequences. *Fungal Genetics and Biology, 45*(3), 266-277.

Bos, J. I. B., Armstrong, M. R., Gilroy, E. M., Boevink, P. C., Hein, I., Taylor, R. M., et al. (2010). *Phytophthora infestans* effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. *Proceedings of the National Academy of Sciences of the United States of America, 107*(21), 9909-9914.

Bozkurt, T. O., Schornack, S., Win, J., Shindo, T., Ilyas, M., Oliva, R., et al. (2011). *Phytophthora infestans* effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface. *Proceedings of the National Academy of Sciences of the United States of America, 108*(51), 20832-20837.

Cahill, D. M., Rookes, J. E., Wilson, B. A., Gibson, L., & McDougall, K. L. (2008). *Phytophthora cinnamomi* and Australia's biodiversity: Impacts, predictions and progress towards control. *Australian Journal of Botany, 56*(4), 279-310.

Cavalier-Smith, T. (1993). Kingdom protozoa and its 18 phyla. *Microbiological Reviews, 57*(4), 953-994.

Chang, S., Puryear, J., & Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter, 11*(2), 113-116.

Chen, D., & Zentmyer, G. A. (1970). Production of sporangia by *Phytophthora cinnamomi* in axenic culture. *Mycologia, 62*(2), 397-402.

Chen, X., Klemsdal, S. S., & Brurberg, M. B. (2011). Identification and analysis of *Phytophthora cactorum* genes up-regulated during cyst germination and strawberry infection. *Current Genetics, 57*, 297-315.

Chen, X., Shen, G., Wang, Y., Zheng, X., & Wang, Y. (2007). Identification of *Phytophthora sojae* genes upregulated during the early stage of soybean infection. *Federation of European Microbiological Societies Microbiology Letters, 269*(2), 280-288.

Chen, X. R., Xing, Y. P., Li, Y. P., Tong, Y. H., & Xu, J. Y. (2013). RNA-Seq reveals infection-related gene expression changes in *Phytophthora capsici. PLoS ONE, 8*(9), e74588.

Chen, X. R., Zhang, B.-Y., Xing, Y.P., Li, Q.Y., Li, Y.P., Tong, Y.H., et al. (2014).Transcriptomic analysis of the phytopathogenic oomycete *Phytophthora cactorum* provides insights into infection-related effectors. *BMC Genomics, 15*(1), 1-24.

Chen, X. R., Li, Y. P., Li, Q. Y., Xing, Y. P., Liu, B. B., Tong, Y. H., et al. (2015). SCR96, a small cysteine-rich secretory protein of *Phytophthora cactorum*, can trigger cell death in the Solanaceae and is important for pathogenicity and oxidative stress tolerance. *Molecular Plant Pathology,*doi:10.1111/mpp.12303

Darvas, J., Toerien, J., & Milne, D. (1984). Control of avocado root rot by trunk injection with fosetyl-Al. *Plant Disease, 68*(8), 691-693.

Emanuelsson, O., Brunak, S., von Heijne, G., & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP, and related tools. *Nature Protocols, 2*(4), 953-971.

Gaulin, E., Dramé, N., Lafitte, C., Torto-Alalibo, T., Martinez, Y., Ameline-Torregrosa, C., et al. (2006). Cellulose binding domains of a *Phytophthora* cell wall protein are novel pathogen-associated molecular patterns. *Plant Cell, 18*(7), 1766-1777.

Gaulin, E., Jauneau, A., Villalba, F., Rickauer, M., Esquerré-Tugayé, M., & Bottin, A. (2002). The CBEL glycoprotein of *Phytophthora parasitica* var. *nicotianae* is involved in cell

wall deposition and adhesion to cellulosic substrates. *Journal of Cell Science, 115*(23), 4565-4575.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology, 29*(7), 644-652.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., et al. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans. Nature, 461*(7262), 393-398.

Hardham, A. R. (2005). *Phytophthora cinnamomi. Molecular Plant Pathology, 6*(6), 589-604.

Have, A., Tenberge, K., Benen, J. E., Tudzynski, P., Visser, J., & Kan, J. L. (2002). The contribution of cell wall degrading enzymes to pathogenesis of fungal plant pathogens. In F. Kempken (Ed.), *Agricultural Applications* (Vol. 11, pp. 341-358, The Mycota). Berlin Heidelberg: Springer.

Hosseini, S., Resjö, S., Yongfeng Liu, Durling, M., Heyman, M., Levander, F., et al. (2015). Comparative proteomic analysis of hyphae and germinating cysts of *Phytophthora pisi* and *Phytophthora sojae. Journal of Proteomics*, doi: 10.1016/j.jprot.2015.01.006

Iseli, C., Jongeneel, C. V., & Bucher, P. (1999). ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proceedings of the Seventh International Conference on Intelligent Systems for Molecular Biology, 99*, 138-148.

Jiang, R. H. Y., de Bruijn, I., Haas, B. J., Belmonte, R., Löbach, L., Christie, J., et al. (2013). Distinctive expansion of potential virulence genes in the genome of the oomycete fish pathogen *Saprolegnia parasitica. PLoS Genetics,* doi:10.1371/journal.pgen.1003272

Jiang, R. H. Y., Tripathy, S., Govers, F., & Tyler, B. M. (2008). RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proceedings of the National Academy of Sciences,*doi:10.1073/pnas.0709303105

Jiang, R. H. Y., Tyler, B. M., Whisson, S. C., Hardham, A. R., & Govers, F. (2006). Ancient origin of elicitin gene clusters in *Phytophthora* genomes. *Molecular Biology and Evolution, 23*(2), 338-351.

Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature, 444*(7117), 323-329.

Judelson, H. S., Ah-Fong, A. M. V., Aux, G., Avrova, A. O., Bruce, C., Cakir, C., et al. (2008). Gene expression profiling during asexual development of the late blight pathogen *Phytophthora infestans* reveals a highly dynamic transcriptome. *Molecular Plant-Microbe Interactions, 21*(4), 433-447.

Judelson, H. S., & Tani, S. (2007). Transgene-induced silencing of the zoosporogenesis-specific NIFC gene cluster of *Phytophthora infestans* involves chromatin alterations. *Eukaryotic Cell, 6*(7), 1200-1209.

Kamoun, S., Young, M., Glascock, C., & Tyler, B. M. (1993). Extracellular protein elicitors from *Phytophthora*: Host-specificity and induction of resistance to fungal and bacterial phytopathogens. *Molecular Plant-Microbe Interactions, 6*(1), 15-25.

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution,* doi:10.1093/molbev/mst010.

Kebdani, N., Pieuchot, L., Deleury, E., Panabières, F., Le Berre, J. Y., & Gourgues, M. (2010). Cellular and molecular characterization of *Phytophthora parasitica* appressorium-mediated penetration. *New Phytologist, 185*(1), 248-257.

Kelley, B. S., Lee, S. J., Damasceno, C. M. B., Chakravarthy, S., Kim, B. D., Martin, G. B., et al. (2010). A secreted effector protein (SNE1) from *Phytophthora infestans* is a broadly acting suppressor of programmed cell death. *Plant Journal, 62*(3), 357-366.

King, M., Reeve, W., Van der Hoek, M. B., Williams, N., McComb, J., O'Brien, P. A., et al. (2010). Defining the phosphite-regulated transcriptome of the plant pathogen *Phytophthora cinnamomi. Molecular Genetics and Genomics,* doi:10.1007/s00438-010-0579-7

Kong, P., Hong, C. X., & Richardson, P. A. (2003). Rapid detection of *Phytophthora cinnamomi* using PCR with primers derived from the *Lpv* putative storage protein genes. *Plant Pathology, 52*(6), 681.

Krajaejun, T., Lerksuthirat, T., Garg, G., Lowhnoo, T., Yingyong, W., Khositnithikul, R., et al. (2014). Transcriptome analysis reveals pathogenicity and evolutionary history of the pathogenic oomycete *Pythium insidiosum*. *Fungal Biology, 118*(7), 640-653.

Krogh, A., Larsson, B., von Heijne, G., & Sonnhamme, E. (2001). Predicting transmembrane protein topology with a Hidden Markov Model: application to complete genomes. *Journal of Molecular Biology, 305*, 567–580.

Lamour, K. H., Mudge, J., Gobena, D., Hurtado-Gonzales, O. P., Schmutz, J., Kuo, A., et al. (2012). Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. *Molecular Plant-Microbe Interactions, 25*, 1350–1360.

Lévesque, C. A., Brouwer, H., Cano, L., Hamilton, J. P., Holt, C., Huitema, E., et al. (2010). Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biology, 11*, R37.

Misas-Villamil, J. C., & van der Hoorn, R. A. L. (2008). Enzyme–inhibitor interactions at the plant–pathogen interface. *Current Opinion in Plant Biology, 11*(4), 380-388.

Narayan, R. D., Blackman, L. M., Shan, W., & Hardham, A. R. (2010). *Phytophthora nicotianae* transformants lacking dynein light chain 1 produce non-flagellate zoospores. *Fungal Genetics and Biology, 47*(8), 663-671.

Nicholas, K. B., Nicolas, H. B. Jr., & Deerfield, D. W. II. (1997). GeneDoc: Analysis and visualization of genetic variation. *EMBNEW.NEWS, 4*, 14.

Nielsen, H., & Krogh, A. (1998). Prediction of signal peptides and signal anchors by a Hidden Markov Model. *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology,6*, 122-130.

OEPP/EPPO (2004). Diagnostic protocols for regulated pests, *Phytophthora cinnamomi*. *Bulletin OEPP/EPPO Bulletin, 34*, 155-157.

Panabières, F., Amselem, J., Galiana, E., & Le Berre, J. Y. (2005). Gene identification in the oomycete pathogen *Phytophthora parasitica* during *in vitro* vegetative growth through expressed sequence tags. *Fungal Genetics and Biology, 42*(7), 611-623.

Perez, V., Huet, J. C., O'Donohue, M., Nespoulous, C., & Pernollet, J. C. (1999). A novel elicitin necrotic site revealed by α-cinnamomin sequence and site-directed mutagenesis. *Phytochemistry, 50*(6), 961-966.

Qutob, D., Huitema, E., Gijzen, M., & Kamoun, S. (2003). Variation in structure and activity among elicitins from *Phytophthora sojae.Molecular Plant Pathology, 4*(2), 119-124.

Randall, T. A., Dwyer, R. A., Huitema, E., Beyer, K., Cvitanich, C., Kelkar, H., et al. (2005). Large-scale gene discovery in the oomycete *Phytophthora infestans* reveals likely components of phytopathogenicity shared with true fungi. *Molecular Plant-Microbe Interactions, 18*(3), 229-243.

Rao, V. (1970). Influence of temperature upon growth and sporulation in two species of *Phytophthora*. *Mycopathologia et mycologia applicata, 42*(1-2), 39-48.

Resjö, S., Ali, A., Meijer, H. J., Seidl, M. F., Snel, B., Sandin, M., et al. (2014). Quantitative label-free phosphoproteomics of six different life stages of the late blight pathogen *Phytophthora infestans* reveals abundant phosphorylation of members of the CRN effector family. *J Proteome Res, 13*, doi:10.1021/pr4009095

Séjalon-Delmas, N., Villalba Mateos, F., Bottin, A., Rickauer, M., Dargent, R., & Esquerré-Tugayé, M. T. (1997). Purification, elicitor activity, and cell wall localization of a glycoprotein from *Phytophthora parasitica* var. *nicotianae*, a fungal pathogen of tobacco. *Phytopathology, 87*(9), 899-909.

Shan, W., Marshall, J. S., & Hardham, A. R. (2004). Gene expression in germinated cysts of *Phytophthora nicotianae*. *Molecular Plant Pathology, 5*(4), 317-330.

Shearer, B. L., & Crane, C. E. (2012). *Phytophthora cinnamomi* visible necrotic lesion-colonisation relationships in native flora. *Australasian Plant Pathology, 41*(6), 633-644.

Shearer, B. L., & Fairman, R. G. (2007). A stem injection of phosphite protects *Banksia* species and *Eucalyptus marginata* from *Phytophthora cinnamomi* for at least four years. *Australasian Plant Pathology, 36*, 78–86.

Stajich, J., Harris, T., Brunk, B., Brestelli, J., Fischer, S., Harb, O., et al. (2012). FungiDB: an integrated functional genomics database for fungi. *Nucleic Acids Research, 1*(40), D675-D681.

Stam, R., Jupe, J., Howden, A. J. M., Morris, J. A., Boevink, P. C., Hedley, P. E., et al. (2013). Identification and characterisation CRN effectors in *Phytophthora capsici* shows modularity and functional diversity. *PLoS ONE,* doi:10.1371/journal.pone.0059517

Studholme, D. J., McDougal, R. L., Sambles, C., Hansen, E., Hardy, G., Grant, M., et al. (2016). Genome sequences of six *Phytophthora* species associated with forests in New Zealand. *Genomics Data,* doi:http://dx.doi.org/10.1016/j.gdata.2015.11.015

Tamura, K., Dudley, J., Nei, M., & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution, 24*(8), 1596-1599.

Torto-Alalibo, T. A., Tripathy, S., Smith, B. M., Arredondo, F. D., Zhou, L., Li, H., et al. (2007). Expressed sequence tags from *Phytophthora sojae* reveal genes specific to development and infection. *Molecular Plant-Microbe Interactions, 20*(7), 781-793.

Tsuda, K., & Katagiri, F. (2010). Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. *Current Opinion in Plant Biology, 13*(4), 459-465.

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H. Y., Aerts, A., et al. (2006). *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science, 313*(5791).

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3 - new capabilities and interfaces. *Nucleic Acids Research,* doi:10.1093/nar/gks596

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology, 3*(7), 1-11.

Vauthrin, S., Mikes, V., Milat, M. L., Ponchet, M., Maume, B., Osman, H., et al. (1999). Elicitins trap and transfer sterols from micelles, liposomes and plant plasma membranes. *Biochimica et biophysica acta Biomembranes, 1419*(2), 335-342.

Wang, Q., Han, C., Ferreira, A. O., Yu, X., Ye, W., Tripathy, S., et al. (2011). Transcriptional programming and functional interactions within the *Phytophthora sojae* RXLR effector repertoire. *Plant Cell, 23*(6), 2064-2086.

Ward, J. A., Ponnala, L., & Weber, C. A. (2012). Strategies for transcriptome analysis in nonmodel plants. *American Journal of Botany, 99*(2), 267-276.

Whisson, S. C., Boevink, P. C., Moleleki, L., Avrova, A. O., Morales, J. G., Gilroy, E. M., et al. (2007). A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature, 450*(7166), 115-118.

Win, J., Morgan, W., Bos, J., Krasileva, K. V., Cano, L. M., Chaparro-Garcia, A., et al. (2007). Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Signaling & Behavior, 3*(4), 251-253.

Ye, W., Wang, X., Tao, K., Lu, Y., Dai, T., Dong, S., et al. (2011). Digital gene expression profiling of the *Phytophthora sojae* transcriptome. *Molecular Plant-Microbe Interactions, 24*(12), 1530-1539.

Zentmyer, G. A. (1961). Attraction of zoospores of *Phytophthora cinnamomi* to avocado roots. *California Avocado Society 1961 Yearbook, 45*, 93-95.

Zentmyer, G. A. (1980). *Phytophthora cinnamomi and the diseases it causes*. St. Paul, Minn.: American Phytopathological Society.

Zentmyer, G. A., & Mircetich, S. M. (1966). Saprophytism and persistence in soil by *Phytophthora cinnamomi. Phytopathology, 56*, 710-712.