Generation and annotation of lodgepole pine and oleoresin-induced expressed sequences from the blue-stain fungus *Ophiostoma clavigerum*, a Mountain Pine Beetle-associated pathogen

Scott DiGuistini[1], Steven G. Ralph[2], Young W. Lim[1], Robert Holt[3], Steven Jones[3], Jörg Bohlmann[2] & Colette Breuil[1]

[1]Department of Wood Science, University of British Columbia, Vancouver, British Columbia, Canada; [2]Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada; and [3]BC Cancer Agency Genome Sciences Centre, Vancouver, British Columbia, Canada

Correspondence: Colette Breuil, Department of Wood Science, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1Z4. Tel.: +1 604 822 9738; fax: +1 604 822 9104; e-mail: colette.breuil@ubc.ca

Received 20 July 2006; revised 29 October 2006; accepted 11 November 2006.
First published online 20 December 2006.

DOI:10.1111/j.1574-6968.2006.00565.x

Editor: Holger Deising

Keywords
ophiostomatoid; conifer; pathogen; lodgepole pine; *Pinus contorta*; fungi.

Abstract

Ophiostoma clavigerum is a destructive pathogen of lodgepole pine (*Pinus contorta*) forests in western North America. It is therefore a relevant system for a genomics analysis of fungi vectored by bark beetles. To begin characterizing molecular interactions between the pathogen and its conifer host, we created an expressed sequence tag (EST) collection for *O. clavigerum*. Lodgepole pine sawdust and oleoresin media were selected to stimulate gene expression that would be specific to this host interaction. Over 6500 cDNA clones, derived from four normalized cDNA libraries, were single-pass sequenced from the 3′ end. After quality screening, we identified 5975 high-quality reads with an average PHRED 20 of greater than 750 bp. Clustering and assembly of this high-quality EST set resulted in the identification of 2620 unique putative transcripts. BLASTX analysis revealed that only 67% of these unique transcripts could be matched to known or predicted protein sequences in public databases. Functional classification of these sequences provided initial insights into the transcriptome of *O. clavigerum*. Of particular interest, our ESTs represent an extensive collection of cytochrome P450 s, ATP-binding-cassette-type transporters and genes involved in 1,8-dihydroxynaphthalene-melanin biosynthesis. These results are discussed in the context of detoxification of conifer oleoresins and fungal pathogenesis.

Introduction

Ophiostoma clavigerum (Robinson-Jeffrey & Davidson) Harrington is a pathogenic fungus of the order *Ophiostomatales*. It is exclusively associated with the closely related bark beetles *Dendroctonus ponderosae* Hopkins [Mountain pine beetle (MPB)] and *Dendroctonus jeffreyii* Hopkins (Kim *et al*., 2005). When MPBs attack their primary conifer host, healthy lodgepole pine (*Pinus contorta* Douglas var. *latifolia* Engelmann), they introduce *O. clavigerum* and other associated fungi. *Ophiostoma clavigerum* colonizes the sapwood rapidly (Robinson, 1962). During colonization, a blue/black melanin pigment is produced, transpiration is blocked and tree death occurs; however, the strength of the wood is not affected. MPB attack and the development of staining fungi result in reduced timber value and an accumulation of unsalvageable trees. In the current MPB epidemic in British Columbia, Canada, tree and wood losses are estimated to be worth billions of dollars (http://www.for.gov.bc.ca/hfp/mountain_pine_beetle/).

The major component of lodgepole pine defense against *O. clavigerum* is oleoresin (Raffa & Berryman, 1982). Oleoresin is composed mainly of 10-carbon monoterpenes, 15-carbon sesquiterpenes and 20-carbon diterpenes (Keeling & Bohlmann, 2006). It is toxic towards some fungi and bark beetles (Delorme & Lieutier, 1990; Himejima *et al*., 1992). Because *O. clavigerum* survives and grows in a saturated oleoresin environment, it likely possesses mechanisms for modifying or degrading the antimicrobial oleoresin components. While we anticipate that detoxifying enzymes such as P450 s and broad-specificity transport proteins such as ATP-binding-cassette (ABC) transporters will likely be important, additional components remain to be discovered. A prerequisite for fully characterizing *O. clavigerum*'s

growth, development and tolerance towards lodgepole pine oleoresin is its genomic sequence. As this is not currently available, we used cDNA libraries to generate large numbers of expressed sequence tags (ESTs), with which we identified genes expressed on media that simulated the conditions that *O. clavigerum* would encounter while growing in lodgepole pine. To facilitate identification of rarer transcripts, we used cDNA library normalization techniques to reduce the frequency of highly expressed genes (Bonaldo *et al.*, 1996). This paper describes an EST collection derived from four normalized cDNA libraries that were generated from filamentous fungal growth of *O. clavigerum* under three different conditions: (1) lodgepole pine sawdust, (2) malt extract agar (MEA) and (3) MEA supplemented with lodgepole pine monoterpene and diterpene metabolites. We have sequenced more than 6500 ESTs representing 2600 unique putative transcripts (UPTs).

Materials and methods

Fungal strain and culture conditions

The *O. clavigerum* (strain SLKw1407) used in this study was isolated from a *D. ponderosae*-infested lodgepole pine tree from an MPB epidemic region near Kamloops, BC. Mycelium for library construction were collected from solid media inoculated with a suspension containing 5×10^5 spores and were incubated for 7 days under ambient conditions. Mycelia were generated on three media: (1) ~10 g of fresh lodgepole pine sawdust that was electron beam sterilized (Iontron Sterilization, Coquitlam, BC) and mixed with 1.5% (w/v) autoclaved agar, (2) 1% MEA (Difco Laboratories, Detroit, MI; w/v) and (3) 1% MEA supplemented with 75 μL of a diterpene blend (abietic, isopimaric and pimaric acids) spread onto the surface of 12.5 mL of media. The diterpene-supplemented cultures were then placed inside a 2 L gas-tight glass bell with a headspace saturated with 250 μL of a mixture of monoterpenes [(±)α-pinene, (−)β-pinene, 3-carene, β-phellandrene, (±) limonene, α-terpinolene and γ-terpinene]. The selected metabolites were blended in a ratio similar to that described by Shrimpton (1973) based on extractive analysis of MPB-attacked trees and is more fully described in Appendix S1.

RNA isolation, cDNA library normalization

For each of the three media conditions, total RNA was isolated from *O. clavigerum* mycelia pooled from 12 to 20 cultures. mRNA was isolated and 5 μg of poly(A^+) mRNA was reverse transcribed and directionally cloned into a 5′EcoRI and 3′XhoI predigested phagemid vector pBS II SK(+) (Stratagene, La Jolla, CA). The libraries were normalized using the method described by Bonaldo *et al.* (1996). The libraries were normalized to $C_o t = 5$, except for the sawdust library, which was split into two fractions, one normalized to $C_o t = 10$ while the other was normalized to $C_o t = 5$. See Appendix S1 for an explanation of $C_o t$.

DNA sequencing and analysis

Randomly selected clones from all libraries were partially sequenced from the 3′ end using the –21 M13 forward primer (Table 1). Chromatograms were processed and quality trimmed using PHRED v. 0.020425.c (Ewing & Green, 1998), and vector sequences were removed using CROSS_MATCH (http://www.phrap.org/). Finally, sequences with PHRED 20 of at least 100 bp were assembled using CAP3 (Huang & Madan, 1999) with default settings, except for a minimum overlap of 40 bp and a minimum percent identity of 95.

Annotation of ESTs

ESTs were annotated using the results from BLASTX and BLASTN (Altschul *et al.*, 1990) analysis unless otherwise noted. The sequence databases used for *O. clavigerum* EST annotation were as follows: the nonredundant division of GenBank (downloaded on 2005-03-07 http://www.ncbi.nlm.nih.gov/BLAST/), Swiss-prot (downloaded on 2005-09-27 http://

Table 1. Primers used in this study

Primer name	Application	Sequence 5′–3′	Primer length
Anchored oligo d(T)	cDNA library construction	GAGAGAGAGAGAGAGAGAACTAGT**CTCGAG**T(T_{16})VN*	51 bp
−21M13 F	cDNA library construction	TGTAAAACGACGGCCAGT	18 bp
M13R	cDNA library construction	CAGGAAACAGCTATGAC	17 bp
T7	Normalization	GTAATACGACTCACTATAGGGC	22 bp
SKpBS	Normalization	CGCTCTAGAACTAGTGGATCC	21 bp
CFEM1-F	qRT-PCR	AGCCACCGGGCTCAGCCAGACA	22 bp
CFEM1-R	qRT-PCR	GGCAACTGCGGCACCGATCC	20 bp
1407btub-F	qRT-PCR	TCTCGACAGCAATGGAGT	18 bp
1407btub-R	qRT-PCR	CCCGAGGCCTCGTTGAAGTA	20 bp

*XhoI (bold) restriction site added.

www.expasy.org/sprot/), the gene ontology (GO) v. 2006-01 (Ashburner *et al.*, 2000) and PHI-base v. 2.1 (Winnenburg *et al.*, 2006). Further analysis using GO annotations was performed with GoMiner (Zeeberg *et al.*, 2003), and the results from this analysis were visualized using VennMaster (Kestler *et al.*, 2005). Further description can be found in Appendix S1.

Quantitative real-time PCR

Trizol (Invitrogen, Mississauga, ON) extractions were used to purify RNA for quantitative real-time PCR (qPCR) with the following modifications: (1) centrifugations were performed at 4 °C and (2) 1-bromo-3-chloro-propane (BCP) was substituted for chloroform. DNaseI (Fisher Scientific, Ottawa, ON) treatment of the Trizol extracted RNA ensured adequate removal of all genomic DNA contamination. cDNA was produced from 5 μg of total RNA using Superscript II (Invitrogen, Mississauga, ON) and oligo $(dT)_{12–18}$ following the manufacturer's protocol. qPCR was performed on a Stratagene M3000P (La Jolla, CA) and data analysis was performed within SAS (Appendix S1; Statistical Analysis Systems, Cary, NC). Primers for the analysis of the transcript of interest were CFEM1-F and CFEM1-R (Table 1); the amplicon size was 90 bp. Primers for the reference transcript were 1407btub-F and 1407btub-R (Table 1); the amplicon size was 135 bp. Primer specificity (a single product of expected length) was confirmed by analysis on a 2% agarose gel and by melting curve analysis. PCR reactions were composed of forward and reverse primers, each at 300 or 600 nM (optimum primer concentration was determined using a dilution curve), 1 × iQ supermix premix (Bio-Rad, Mississauga, ON) and 50 ng of *O. clavigerum* cDNA in a total volume of 25 μL. Cycling parameters for qPCR were 95 °C for 10 min, followed by 40 cycles of 95 °C for 10 s, 62 °C for 30 s, 72 °C for 30 s and an observation step of 82 °C for 18 s, followed by a melting point analysis. Each analysis was replicated three times with biologically and technically independent samples.

Results

Generation and assembly of EST collections

We generated a total of 6528 3′-EST sequences from four normalized, uni-directional, cDNA libraries (MPB01-04) based on mycelial growth under three different culture conditions: (1) lodgepole pine sawdust, (2) MEA and (3) MEA supplemented with monoterpenes and diterpenes (Table 2). After quality filtering, sequences < 100 bp were excluded, and 5975 high-quality sequences remained, with an average PHRED 20 for these reads of 786 bp. The ratio of high-quality reads to total reads was 92% and this did not differ substantially between libraries. CAP3 analysis of the high-quality ESTs identified 2620 UPTs: 4497 sequences clustered into 1142 contigs (the average number of contig members was 3.94) and 1478 singletons (Table 2). Although most UPTs were represented by a single member, 965 contigs contained two to five members, 170 contigs contained six to 20 members and seven contigs contained 21 or more members. The largest contig contained 72 ESTs, and was annotated as a gene encoding a 40S ribosomal protein (Table 3). Libraries MPB01-04 contained 365, 618, 378 and 400 unique transcripts, respectively; only 52 UPTs were represented with ESTs from all four libraries. All sequences have been deposited into GenBank's dbEST database (accession nos. EE724403–EE730376).

Table 2. *Ophiostoma clavigerum* EST summary

Total 3′ End sequences	
MPB01 ($C_o t = 10$, Sawdust)	1536
MPB02 ($C_o t = 5$, MEA)	1920
MPB03 ($C_o t = 5$, MEA+Terpenes)	1536
MPB04 ($C_o t = 5$, Sawdust)	1536
Total	6528*
Complete assembly	
High quality reads	5975
Reads in contigs†	4497
Contigs	1142
Singletons	1478
Unique sequences	2620

*Average PHRED 20 for all reads is 786; PHRED 20 should be interpreted as a 1 in 100 probability that a base has been called incorrectly; in other words, sequences are considered 99% accurate.

†CAP3 assembly used default settings with the minimum overlap set to 40 bp and the minimum percent identity set to 95.

EST analysis

Of the 2620 UPTs, BLASTX of the nonredundant database identified 67% (1755 transcripts) that were similar to previously deposited protein sequences [score ≥ 100, Expected (E)-value $\leq 10^{-15}$]. Ninety-two percent of the transcripts with 'best matches' were similar to known or predicted protein sequences from *Magnaporthe grisea* (35%) *Neurospora crassa* (34%), or *Gibberella zeae* (23%). Another 7% (183) of these transcripts could be matched with proteins from other fungi, and only 1% had 'best matches' that were not fungi.

Following this preliminary annotation, UPTs were ranked by frequency analysis. Frequency was calculated as the number of ESTs contributing to a UPT. The frequency analysis revealed that a small number of UPTs appeared to be over-represented within individual libraries. For example, 20 of the 22 MPB0586 EST sequences originated from the MPB03 library (Table 3). The BLASTX 'best match' for

MPB0586 was *Aspergillus fumigatus* EAL88523, which was annotated as a cysteine-containing domain present in fungal extracellular membrane proteins (CFEM). qPCR indicated that the expression of the CFEM-domain-containing gene, MPB0586, was induced by fungal growth in the presence of the selected oleoresin metabolites. Monoterpenes stimulated the expression of this gene by *c.* 2.75-fold ($P < 0.0001$, critical$_\alpha = 0.0167$) compared with expression on MEA alone, while the solvent DMSO caused some increase in expression (1.64, $P = 0.0004$, critical$_\alpha = 0.0167$) compared with MEA. In relation to the DMSO treatment, the diterpene abietic acid did not have an effect on MPB0586 expression ($P = 0.0121$, critical$_\alpha = 0.0167$). We did not test the combined effects of diterpene acids with monoterpenes.

As no information is available for the genes involved in *O. clavigerum*'s pathogenicity, UPTs were compared against PHI-base, a curated sequence collection of fungal-verified virulence and pathogenicity genes. To determine the EST collection's comprehensiveness, we analyzed the cytochrome P450 and ABC-type transporter gene families, as these are likely involved in host defense chemical detoxification. Because we previously identified genes for most of the Ophiostomatoid enzymes in the 1,8-dihydroxynaphthalene-melanin biosynthetic pathway (Loppnau *et al.*, 2004), we used this pathway as a reference for assessing the EST collection's completeness. BLASTX was used to compare the *O. clavigerum* UPTs with PHI base (score ≥ 100, *E* value $\leq 10^{-50}$), and UPTs were ranked by frequency;

Table 3. UPTs ranked by EST frequency

		Library MPB0x					
Uniseq ID	EST frequency	1	2	3	4	Best match*; accession no.	*E* value
MPB0888	72	3	34	33	2	40S ribosomal protein S3; XP_322575	6.00E-38
MPB0492	58	7	19	27	5	Hypothetical protein; EAA54505	7.00E-20
MPB0391	37	3	16	16	2	Cytochrome c reductase iron–sulfur subunit; CAA26308	1.00E-93
MPB0564	33	6	9	14	4	Hypothetical protein; XP_331746.1	7.00E-60
MPB01049	30	4	15	9	2	No significant similarity	NA
MPB0410	25	11	1	0	13	No significant similarity	NA
MPB0586	22	0	2	20	0	CFEM domain protein; EAL88523	2.00E-09
MPB0924	19	1	8	7	3	Hypothetical protein; EAA50425.1	3.00E-75
MPB0732	18	1	7	10	0	No significant similarity	NA
MPB0205	18	7	5	3	3	No significant similarity	NA
MPB0556	17	4	4	8	1	Vacuolar ATP synthase subunit E; XP_327732	3.00E-55
MPB0939	17	1	9	5	2	Short chain dehydrogenase/reductase family; EAL84601	4.00E-36
MPB0278	17	6	3	0	8	Malate dehydrogenase-like protein; AAX07691	1.00E-116
MPB0545	16	0	9	7	0	Microtubule associated protein EB1; EAL92397	8.00E-70
MPB0652	16	3	5	7	1	Dihydrolipoyllysine-residue acetyltransferase-like protein; AAX07694	1.00E-105
MPB0413	16	0	7	9	0	Aspartate-semialdehyde dehydrogenase; EAL92885	6.00E-87
MPB0107	16	10	1	1	4	Related to protein-tyrosine phosphatase; CAD70824	3.00E-30
MPB0365	16	6	2	0	8	Hypothetical protein; XP_328441	2.00E-47
MPB0464	15	5	1	2	7	Cytochrome c oxidase subunit VIa family; EAL92949	3.00E-31

*'Best Match' determined using results of a BLASTX analysis against the nonredundant database.
Frequency has not been normalized for differences in the number of reads per library.

Table 4. Comparison of *Ophiostoma clavigerum* UPTs with the PHI-base database*

Uniseq ID	EST frequency	Accession; best match; pathogen	*E* value
MPB0197	9	AAK98783; putative vacuolar ATPase MVP1; *Magnaporthe grisea*	2.00E-62
MPB0489	7	BAB85760; putative mitochondrial carrier protein; *Fusarium oxysporum*	4.00E-86
MPB01268	5	AAQ16572; putative mitochondrial cyclophilin 1; *Botryotinia fuckeliana*	2.00E-67
MPB0945	4	AAP68994; thiol-specific antioxidant protein 1; *Cryptococcus neoformans* var. *grubii*	1.00E-62
MPB0610	3	CAC17748; trehalose-6-phosphate phosphatase; *Candida albicans*	5.00E-72
MPB0132	3	CAA67930; putative mannosyl transferase; *Candida albicans*	2.00E-91
MPB01050	3	AAB86583; manganese-superoxide dismutase precursor; *Candida albicans*	1.00E-63
MPB0293	2	AAD47837; alanine racemase; *Cochliobolus carbonum*	2.00E-77
MPB0893	2	CAB56523; ornithine decarboxylase; *Phaeosphaeria nodorum*	2.00E-81
MPB0212	2	AAF09475; osmotic sensitivity MAP kinase; *Magnaporthe grisea*	1.00E-100
MPB0720	1	CAD88591; superoxide dismutase; *Botryotinia fuckeliana*	2.00E-69

*Comparison was performed using BLASTX (score ≥ 100, *E* value $\leq 10^{-50}$).

Table 4 highlights the top ten sequences. Fourteen cytochrome P450s containing the heme-binding site motif: Phe-X-X-Gly-X-Arg-X-Cys-X-Gly (Werck-Reichhart & Feyereisen, 2000) were identified and manually confirmed within the *O. clavigerum* EST collection (Table 5). We also examined ABC-transporter proteins, and found six sequences whose 3′ ends contained elements of the ABC transporter signature motif, the linker peptide region (CDD: cd00267; http://www.ncbi.nlm.nih.gov/Structure/cdd/; Table 5). The *O. clavigerum* EST collection contained a number of genes from the 1,8-dihydroxynaphthalene–melanin biosynthesis pathway. Sequences with similarity to genes from *Ophiostoma* and *Ceratocystis* species were found and included a polyketide synthase and scytalone dehydratase as well as tetra- and tri- hydroxynaphthalene reductases. This analysis also identified a sequence present in *A. fumigatus*, AYG1, not previously shown in *Ophiostoma* and *Ceratocystis* species (Table 6).

Comparing *O. clavigerum* UPTs against the GO database extended the functional annotation. GO annotations were assigned using the single 'best-match' BLASTX hit against the GO sequence database (score ≥ 100, *E* value $\leq 10^{-15}$), and 905 (34%) of *O. clavigerum's* unique sequences could be associated with GO terms. Following the assignment of GO terms to *O. clavigerum's* UPTs, GoMiner was used to compare the distributions of ESTs among the biological process and molecular function categories and between the cDNA libraries (data not shown). ESTs with GO associations for biological processes like actin cytoskeleton binding and organization, as well as transcriptional and cell cycle regulation were overrepresented in libraries MPB01 and MPB04, which originated from *O. clavigerum* growth on sawdust

Table 5. UPTs belonging to the P450 and ABC gene families

Uniseq ID	EST Frequency	Swiss-prot ID; description*	*E* value
P450s			
MPB0658	8	P79084; O-methylsterigmatocystin oxidoreductase	6.00E-21
MPB0465	4	Q8K4D6; cytochrome P450 4X1	6.00E-24
MPB0943	3	P17549; benzoate 4-monooxygenase	1.00E-141
MPB0897	2	P17177; cytochrome P450 27, mitochondrial precursor	0.003
MPB1123	2	O13317; isotrichodermin C-15 hydroxylase	4.00E-29
MPB2073	1	Q12645; pisatin demethylase	3.00E-06
MPB1836	1	P17549; benzoate 4-monooxygenase	7.00E-50
MPB1190	1	Q12612; trichodiene oxygenase	3.00E-14
MPB1208	1	Q9VE01; probable cytochrome P450 12a5, mitochondrial precursor	4.00E-06
MPB1396	1	P15540; cytochrome P450 steroid 21-hydroxylase	6.00E-07
MPB1828	1	No significant hit	NA
MPB1650	1	Q92088; cytochrome P450 2M1 lauric acid omega-6-hydroxylase	6.00E-25
MPB1907	1	Q9LTM7; cytochrome P450 71B16	8.00E-14
MPB2140	1	P54781; cytochrome P450 61 C-22 sterol desaturase	7.00E-55
ABCs			
MPB00211	4	P32386; ATP-dependent bile acid permease	1.00E-15
MPB00109	2	P51533; ATP-dependent permease PDR10	5.00E-43
MPB01201	1	P43569; probable ATP-dependent transporter YFL028C	2.00E-07
MPB02272	1	P53049; oligomycin resistance ATP-dependent permease YOR1	9.00E-65
MPB02382	1	P36619; leptomycin B resistance protein pmd1	8.00E-46
MPB01806	1	P43569; probable ATP-dependent transporter YFL028C	4.00E-52

*Descriptions have been generated from Swiss-prot and belong to the ID of the 'Best Match'.
'No significant hit' indicates that the sequence contains a heme-binding motif but could not be matched to a sequence from the public databases with significant similarity to suggest a relationship.

Table 6. Analysis of UPTs identified to be involved in DHN-melanin biosynthesis*

Uniseq ID	EST frequency	Accession; description; fungus	*E* value
MPB0329	2	BAA18956; polyketide synthase; *Colletotrichum lagenarium*	3.00E-100
MPB0166	8	AAF03354; serine protease-like hydrolase; *Aspergillus fumigatus*	2.00E-66
MPB0789	10	AAK07185; tetrahydroxynaphthalene reductase; *Ophiostoma floccosum*	1.00E-123
MPB0755	8	AAK11296; scytalone dehydratase; *Ophiostoma floccosum*	5.00E-86
MPB0851	9	AAK60499; trihydroxynaphthalene reductase; *Ophiostoma floccosum*	1.00E-113

*UPTs were identified by comparing known *A. fumigatus* pathway members against the *O. clavigerum* EST collection using TBLASTX.

media. Libraries derived from fungal growth on MEA possessed roughly the same proportion of unique sequences as libraries generated from fungal growth on sawdust. However, libraries derived from fungal growth on MEA contained a more narrow GO distribution, with more unique sequences assigned to fewer biological process and molecular function categories. For the library derived from *O. clavigerum* growth on MEA supplemented with oleoresin terpenoids, ESTs with GO associations for the biological processes of cytoplasm organization and biogenesis were overrepresented, along with several shared terms related to vitamin metabolism; under molecular function, cytoskeleton-binding proteins were overrepresented, in addition to a large group of proteins putatively identified as having oxidoreductase activity.

Discussion

Ophiostoma clavigerum is an MPB-associated pathogen that causes economic losses by discoloring wood and killing MPB-infested lodgepole pine trees. Very limited gene sequence information is available for this fungus, despite its economic and ecological importance. In this work, we generated and described a normalized EST collection for a strain of *O. clavigerum* isolated from a lodgepole pine tree in one of British Columbia's MPB epidemic regions. We sequenced more than 6500 ESTs, which represented *c.* 2620 UPTs.

We used GoMiner with VennMaster to compare GO terms associated with unassembled ESTs with terms for assembled UPTs, and identified GO categories overrepresented within biological process and molecular function hierarchies. For cDNA libraries derived from fungus grown on sawdust, overrepresented biological processes included actin cytoskeleton reorganization as well as transcription and cell-cycle regulation. Fungal growth in conifer stems and on sawdust-derived media is nutrient-limited (Meerts, 2002). Consistent with this, cultures grown on lodgepole pine sawdust had overrepresented GO categories with similarity to those for nutrient-limited yeast cells (Gasch & Werner-Washburne, 2002). Because specific nutrient limitations are often developmental cues (Cullen & Sprague, 2000) and can induce virulence-associated gene expression (Snoeijers *et al.*, 2000), we anticipate that the ESTs derived from *O. clavigerum* grown on sawdust will contain sequences useful for characterizing fungal pathogenicity.

Oleoresin is a major component of conifer defense systems (Keeling & Bohlmann, 2006). However, *O. clavigerum* cultured with oleoresin shows only slightly decreased growth rates (Shrimpton & Whitney, 1968); further, some oleoresin components stimulate its growth (Paine & Hanlon, 1994). GO analysis of ESTs from fungi grown on oleoresin indicated that oxidoreductase enzymes were overrepresented. Many plant pathogens utilize oxidoreductases for antimicrobial natural product detoxification and for host chemical perception (Idnurm & Howlett, 2001; Palmer *et al.*, 2004). As little is known about the molecular mechanisms utilized by *O. clavigerum* to colonize lodgepole pine, these observations represent good leads for dissecting the molecular genetics of oleoresin detoxification and pathogenicity.

P450-mediated phytoalexin detoxification is important for some plant–fungal interactions (Maloney & VanEtten, 1994). P450s have been shown in both trees and fungi to participate in terpenoid biosynthesis (Parker & Scott, 2005; Ro *et al.*, 2005). To identify genes contributing to *O. clavigerum* pathogenesis, we screened the EST collection for members of the P450 protein family. Fourteen P450 UPTs were identified in *O. clavigerum* ESTs that contained the heme-binding site motif. The majority of P450 clones were sequenced once; none were sequenced more than eight times. Alignment and phylogenetic analyses indicated no relationship between the four cDNA libraries and P450 CYP classification or P450 EST frequency (data not shown). Similarly, ABC transporters have been implicated in antibiotic resistance for a number of agriculturally important fungal pathogens (Urban *et al.*, 1999) and are also involved in plant antifungal terpenoid secretions (Jasinski *et al.*, 2001). We found six UPTs whose 3′ ends were similar to known ABC genes. Because lodgepole pine is rich in antimicrobial oleoresin components, both P450s and ABC transporters represent interesting candidates for functional characterization in *O. clavigerum* interactions with lodgepole pine defense chemicals.

Because Ophiostomatoid fungi cause substantial economic losses to the forest industry by discoloring sapwood, 1,8-dihydroxynaphthalene-melanin biosynthesis is their most extensively studied pathway (Loppnau *et al.*, 2004). As expected, the *O. clavigerum* EST collection contained 1,8-dihydroxynaphthalene-melanin biosynthesis pathway genes previously characterized in Ophiostomatoids. In addition, we identified the AYG1 gene, a polyketide-shortening hydrolase (Fujii *et al.*, 2004) not previously shown in Ophiostomatoid fungi. 1,8-Dihydroxynaphthalene-melanin has been shown to play an important role in many fungal pathosystems by strengthening cell walls and providing environmental protection against reactive oxygen species and UV. However, its role in *O. clavigerum* biology is uncertain. ESTs identified in this project will allow for future studies testing the roles of 1,8-dihydroxynaphthalene-melanin in *O. clavigerum* interactions with lodgepole pine.

In summary, we generated the first genome-scale expressed sequence data for *O. clavigerum* and provided preliminary annotations for many of the UPTs. In addition, we identified candidate genes for further studies to test the role of oleoresin tolerance in the pathogenicity of

O. clavigerum towards lodgepole pine. These results initiate the characterization of the molecular interactions between this fungal pathogen and its host. The ESTs will serve as reagents for developing additional genomics tools for characterizing *O. clavigerum* gene expression, and will be broadly applicable to studies of Ophiostomatoid fungi vectored by bark beetles.

Acknowledgements

We thank Claire Oddy for technical assistance with library development, and, from the Genome Sciences Centre, Dr Gordon Robertson for critically reviewing the manuscript and Rob Kirkpatrick for bioinformatics assistance. The work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (grant to J.B, C.B).

References

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Ashburner M, Ball CA, Blake JA *et al.* (2000) Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet* **25**: 25–29.

Bonaldo MF, Lennon G & Soares MB (1996) Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res* **6**: 791–806.

Cullen PJ & Sprague GF Jr (2000) Glucose depletion causes haploid invasive growth in yeast. *Proc Natl Acad Sci USA* **97**: 13619–13624.

Delorme L & Lieutier F (1990) Monoterpene composition of the preformed and induced resins of Scots pine, and their effect on bark beetles and associated fungi. *Eur J For Path* **20**: 304–316.

Ewing B & Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**: 186–194.

Fujii I, Yasuoka Y, Tsai HF, Chang YC, Kwon-Chung KJ & Ebizuka Y (2004) Hydrolytic polyketide shortening by ayg1p, a novel enzyme involved in fungal melanin biosynthesis. *J Biol Chem* **279**: 44613–44620.

Gasch AP & Werner-Washburne M (2002) The genomics of yeast responses to environmental stress and starvation. *Funct Integr Genomics* **2**: 181–192.

Himejima M, Hobson KR, Otsuka T, Wood DL & Kubo I (1992) Antimicrobial terpenes from oleoresin of Ponderosa pine tree *Pinus Ponderosa* – a defense mechanism against microbial invasion. *J Chem Ecol* **18**: 1809–1818.

Huang X & Madan A (1999) CAP3: a DNA sequence assembly program. *Genome Res* **9**: 868–877.

Idnurm A & Howlett BJ (2001) Pathogenicity genes of phytopathogenic fungi. *Mol Plant Pathol* **2**: 241–255.

Jasinski M, Stukkens Y, Degand H, Purnelle B, Marchand-Brynaert J & Boutry M (2001) A plant plasma membrane ATP binding cassette-type transporter is involved in antifungal terpenoid secretion. *Plant Cell* **13**: 1095–1107.

Keeling CI & Bohlmann J (2006) Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. *New Phytol* **170**: 657–675.

Kestler HA, Muller A, Gress TM & Buchholz M (2005) Generalized Venn diagrams: a new method of visualizing complex genetic set relations. *Bioinformatics* **21**: 1592–1595.

Kim JJ, Allen EA, Humble LM & Breuil C (2005) Ophiostomatoid and basidiomycetous fungi associated with green, red, and grey lodgepole pines after mountain pine beetle (*Dendroctonus ponderosae*) infestation. *Can J For Res* **35**: 274–284.

Loppnau P, Tanguay P & Breuil C (2004) Isolation and disruption of the melanin pathway polyketide synthase gene of the softwood deep stain fungus *Ceratocystis resinifera*. *Fungal Genet Biol* **41**: 33–41.

Maloney AP & VanEtten HD (1994) A gene from the fungal plant pathogen *Nectria haematococca* that encodes the phytoalexin-detoxifying enzyme pisatin demethylase defines a new cytochrome P450 family. *Mol Gen Genet* **243**: 506–514.

Meerts P (2002) Mineral nutrient concentrations in sapwood and heartwood: a literature review. *Ann For Sci* **59**: 713–722.

Paine TD & Hanlon CC (1994) Influence of oleoresin constituents from *Pinus ponderosa* and *Pinus jeffreyi* on growth of Mycangial fungi from *Dendroctonus ponderosae* and *Dendroctonus jeffreyi*. *J Chem Ecol* **20**: 2551–2562.

Palmer AG, Gao R, Maresh J, Erbil WK & Lynn DG (2004) Chemical biology of multi-host/pathogen interactions: chemical perception and metabolic complementation. *Annu Rev Phytopathol* **42**: 439–64.

Parker JE & Scott DB (2005) Indole diterpene biosynthesis in ascomycetes fungi. *Handbook of Mycology*, *Vol. 22*. (An Z, ed), pp. 405–426. CRC Press, New York.

Raffa KF & Berryman AA (1982) Physiological differences between lodgepole pines resistant and susceptible to the Mountain Pine Beetle (Coleoptera, Scolytidae) and associated microorganisms. *Environ Ent* **11**: 486–492.

Ro DK, Arimura G, Lau SY, Piers E & Bohlmann J (2005) Loblolly pine abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. *Proc Natl Acad Sci USA* **102**: 8060–8065.

Robinson RC (1962) Blue stain fungi in lodgepole pine (*Pinus contorta* Dougl. var. latifolia Engelm.) infested by the Mountain Pine Beetle (*Dendroctonus monticolae* Hopk.). *Can J Bot* **40**: 609–614.

Shrimpton DM (1973) Extractives associated with wound response of lodgepole pine attacked by the mountain pine beetle and associated microorganisms. *Can J Bot* **51**: 527–534.

Shrimpton DM & Whitney HS (1968) Inhibition of growth of blue stain fungi by wood extractives. *Can J Bot* **46**: 757–761.

Snoeijers SS, Pérez-Garcia A, Joosten MHAJ & De Wit PJGM (2000) The effect of nitrogen on disease development and gene

expression in bacterial and fungal plant pathogens. *Eur J Plant Path* **106**: 493–506.

Urban M, Bhargava T & Hamer JE (1999) An ATP-driven efflux pump is a novel pathogenicity factor in rice blast disease. *EMBO J* **18**: 512–521.

Werck-Reichhart D & Feyereisen R (2000) Cytochromes P450: a success story. *Genome Biol* **1**: R3003.

Winnenburg R, Baldwin TK, Urban M, Rawlings C, Kohler J & Hammond-Kosack KE (2006) PHI-base: a new database for pathogen host interactions. *Nucleic Acids Res* **34**: D459–D464.

Zeeberg BR, Feng W, Wang G *et al.* (2003) GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol* **4**: R28.

Supplementary material

The following supplementary material is available for this article:

Appendix S1.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1574-6968.2006.00565.x (This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.