Biological Research
versión impresa ISSN 0716-9760

Biol. Res. v.39 n.2 Santiago 2006

http://dx.doi.org/10.4067/S0716-97602006000200018
Biol Res 39: 367-376, 2006

ARTICLE

Generation and analysis of expressed sequence tags from *Botrytis cinerea*

EVELYN SILVA[1, 3], JORGE VALDÉS[1, 3], DAVID HOLMES[1, 3], AMIR SHMARYAHU[1, 3] and PABLO D T VALENZUELA[1, 2, 3]

[1] Fundación Ciencia para la Vida and Instituto MIFAB, Santiago, Chile.
[2] Pontificia Universidad Católica de Chile, Santiago, Chile.
[3] Universidad Andrés Bello, Santiago, Chile.

Dirección para Correspondencia

ABSTRACT

Botrytis cinerea is a filamentous plant pathogen of a wide range of plant species, and its infection may cause enormous damage both during plant growth and in the post-harvest phase. We have constructed a cDNA library from an isolate of *B. cinerea* and have sequenced 11,482 expressed sequence tags that were assembled into 1,003 contigs sequences and 3,032 singletons. Approximately 81% of the unigenes showed significant similarity to genes coding for proteins with known functions: more than 50% of the sequences code for genes involved in cellular metabolism, 12% for transport of metabolites, and approximately 10% for cellular organization. Other functional categories include responses to biotic and abiotic stimuli, cell communication, cell homeostasis, and cell development. We carried out pair-wise comparisons with fungal databases to determine the *B. cinerea* unisequence set with relevant similarity to genes in other fungal pathogenic counterparts. Among the 4,035 non-redundant *B. cinerea* unigenes, 1,338 (23%) have significant homology with *Fusarium verticillioides* unigenes. Similar values were obtained for *Saccharomyces cerevisiae* and *Aspergillus nidulans* (22% and 24%, respectively). The lower percentages of homology were with *Magnaporthe grisae* and *Neurospora crassa* (13% and 19%, respectively). Several genes involved in putative and known fungal virulence and general pathogenicity were identified. The results provide important information for future research on this fungal pathogen.

Key terms: *Botrytis cinerea*, cDNA sequencing, fungal pathogenicity.

INTRODUCTION

Botrytis cinerea, also known as "gray mold fungus," causes serious pre- and post-harvest diseases in at least two hundred plant species (Jarvis, 1977), including agriculturally important crops and harvested commodities, such as grapes, tomatoes, strawberries, cucumbers, bulb flowers, cut flowers and ornamental plants. The broad host range of *B. cinerea* results in great economic losses not only during growth but also during storage and transport of the products (Elad et al., 2004). In addition, *B. cinerea* strains have been shown to be highly variable genetically as well physiologically (Muñoz et al., 2002; Cotoras and Silva, 2005), and several strains have developed resistance to most of the fungicides used to control them (Latorre et al., 1994).

Genome-wide analysis provides a new and powerful way of investigating fungal pathogens. The application of molecular genetic analysis to the study of other phytopathogenic fungi like as *Magnaporthe grisae*, *Aspergillus nidulans* and *Fusarium oxysporum*, has led to the identification of some of the genes involved in fungal pathogenicity (Idnurm and Howlett, 2001; Sweigard et al., 1998). These include genes involved in detoxification of antifungal compounds produced by plants (Straney and Van Etten, 1994), biosynthesis of phytotoxic compounds (Panaccione et al., 1992), breakdown of the host plant cuticle (Tonukari et al., 2000), conidiogenesis(Hamer and Givan, 1990), appressorium formation and function (Balhadère and Talbot 2001; Talbot et al., 1993), and amino acid metabolism (Balhadère et al., 1999), as well as those involved in conserved signaling pathways (Xu and Hamer, 1996).

Several studies have dealt with the infection mechanisms of *B. cinerea*, with emphasis on microscopic and biochemical studies (Staples and Mayer, 1995) or molecular genetics (Prins et al., 2000). The application of molecular-genetic tools such as transformation (Hamada et al., 1994), differential gene expression analysis (Benito et al., 1996), gene cloning (van der Vlugt-Bergmans et al., 1997), and targeted mutagenesis (van Kan et al., 1997) have led to novel insights about the *B. cinerea* genes involved in the infection process. However, very little is known about spore adhesion and germination.

The objective of this investigation is to generate an inventory of potential pathogenicity determinants via a ground-wide analysis of expressed sequence tags (ESTs) in *B. cinerea*. Such an inventory is indispensable for deepening our understanding of host-pathogen interactions and for suggesting candidate determinants that could aid in the design of novel antifungicides.

Here, we describe the construction and analysis of an EST library from *B. cinerea* growing in culture. The data represents a substantial increase in available sequence information on *B. cinerea* expressed genes. Through the analysis of the ESTs, we also identified genes shown by others to be preferentially expressed during

filamentous growth, putative virulence factors, signal transduction components, and secondary metabolism enzymes providing a solid framework for investigating the cellular events associated with pathogenicity.

MATERIALS AND METHODS

Isolation and properties of strains

The *Botrytis cinerea* strain (A1) used was isolated from blueberry fruits and was maintained on malt-yeast extract agar medium (2% malt extract, 0.2% yeast extract, and 1.5% agar) at 23°C for five days. For longer periods, samples were stored in the form of spore suspensions in 20% glycerol at -80 °C. For this purpose, mycelia were removed from the culture plates by gently brushing the surfaces with a sterile platinum loop, suspended in distilled water, and filtered through two layers of gauze. The concentration of spores was determined using a Neubauer counter and a light microscope.

RNA extraction and RT-PCR

B. cinerea A1 was grown in Erlenmeyer flasks containing 100 ml of malt-yeast extract liquid medium with $2x10^5$ conidia/ml for four days at 23°C without shaking. Mycelia mats were harvested, washed, dried between sterile paper towels, and frozen. RNA was extracted from the frozen biomass by grinding with a pestle and mortar in liquid nitrogen. The resulting powder was suspended in lysis buffer and extracted following the method of Chang et al. (1993). The RNA was quantified spectrophotometrically and its integrity was evaluated by RT-PCR.

Previous to PCR, cDNA synthesis was carried out by incubation of 3 µg of total RNA with random primers and MMLV reverse transcriptase (Promega) at 37ºC for 60 min. All reactions were conducted in a final volume of 25 ml containing 4 mM primers, 200 mM of each dNTP, 1.5 mM $MgCl_2$, 10 ng of cDNA, and 1 unit of Taq DNA polymerase (Promega).

PCR was performed using 1 ml cDNA template with 10 pmol of each PCR primer and 2 units of Taq polymerase (Promega). Reactions were performed in a PJ Research thermal cycler, and the amplified products (30 cycles) were analyzed on a 1.5% agarose gel in TAE buffer.

Preparation of a cDNA library from mycelia

Poly A^+ mRNA was isolated using the Stratagene Poly A Quick mRNA isolation kit (Stratagene, La Jolla, California, USA). Unidirectional double-stranded cDNA was prepared from poly A^+ RNA and cloned using the vector pExpress I (Agencourt Bioscience Corp.), exploiting the *Eco*RI and *Xho*I restriction sites. The cDNA libraries were not normalized, and no attempt was made to reduce the redundancy of highly expressed transcripts. The EST sequences are registered in the GenBank database under accession numbers EB801722-EB812785.

Computational methods

The DNA sequence chromatograms were analyzed using the Phred software (Ewing

et al., 1998) for base call and for estimation of error probability. All sequences were subjected to trimming process to remove ribosomal RNA, polyA tails, low-quality sequences, and vector and adapter regions.

Sequences with 94% identity over 40 or more bases were assembled using CAP3 (Huang and Madan, 1999). The resulting contigs and singletons (unigenes) were compared to the protein non-redundant database using BlastX (Altschul et al., 1997) and analyzed with the InterProScan program (Zdobnov and Apweiler, 2001) to assign putative function. Gene ontology (GO) categories (Ashburner et al., 2000) were extracted from the best hits obtained from the BlastX comparison against SwissProt-Trembl database (e-value < e-15 and >70% of alignment coverage) and compared to the InterProScan GO catalog. All GO assignments were curated manually.

To evaluate the number of orthologous genes present in the library, a BlastX was performed to compare the *B. cinerea* unigene set against selected fungal species. Candidate orthologs were defined as those with hits with e-values lower than e^{-15}. The results were subsequently sorted and analyzed with specific Perl scripts.

RESULTS AND DISCUSSION

Characteristics of the *Botryis cinerea* cDNA library

The titer of the cDNA library was $2.2x10^6$ colony forming units/ml. This value was considered to be representative for the analysis of the genes expressed in mycelium (Liu and Yang, 2005; Trial et al., 2003). Blue/white plaque identification following plating of an aliquot of the library revealed 96% recombinant plaques. The quality of the library was assessed by examining the insert size of a few randomly selected recombinant plaques. The average insert size was 1.25 kb, and this representative value is similar to that of other libraries such as *Trichoderma harzarium* (Liu and Yang, 2005) and *Gibberella zeae* (Trial et al., 2003).

Generation and annotation of expressed sequence tags

EST sequences were produced from the cDNA library and scanned visually to confirm overall quality of peak shape and correspondence with base identification. After the cleaning process, the average length per EST of the remaining sequences (8,821) was 906 base pairs and the Phred quality value was larger than 20. The sequences were assembled into 1,003 contigs and 3,032 singletons (Table 1).

TABLE 1

ESTs sequence analysis

Total ESTs sequenced	11,482
Average length per EST	906 pb
Number of contigs	1,003
Number of singletons	3,032
Total of unigenes	4,035
Unigenes with database hits	81%
Unigenes with a known function	25%

Contigs were composed of multiple ESTs ranging from 2 to 229. The percentage of unigene sequences with similarity to GenBank database was 81%, a value similar to other fungal libraries reported in the literature (Nugent et al., 2004; Liu and Yang, 2005). The total number of unigenes was 4,035 corresponding to less than half of the total sequences obtained (Table 1). The ratio of sequenced ESTs to the number of unigenes is similar to that reported for other fungal EST libraries, such as *Ustilago maydis* and *T. harzarium* (Nugent et al., 2004; Lui and Yang, 2005).

The sequences identified represent 4,035 putative genes expressed during mycelia growth, and if it is assumed that *B. cinerea* has a gene density similar to *Neurospora crassa* and *M. grisae*, this number of genes might represent more than one-third of all *B. cinerea* genes.

The similarity of sequences present in the *B. cinerea* strain T4 library (Viaud et al., 2005) and ours was determined by assembling both libraries with the same parameters mentioned in the Materials and Methods section. From this analysis, a total of 1,841 contigs were obtained, from which 447 sequences were found only in the Viaud et al. library, 574 only from our library and 820 contigs produced by both sources. This analysis also produced 2,233 singletons from the library of Viaud et al. and 4,786 from ours. These results show the higher level of novel sequences provided by our work accounting for approximately 30% of the total unigenes set. The sum of the unique genes represented by both our library and that of Viaud et al. corresponds to more than half the gene content available for this microorganism.

To determine the frequency of ESTs, a file containing contig identifiers was constructed indicating the number of ESTs per contig. High redundancy of a specific cDNA sequence among ESTs is likely to be correlated with a higher expression level of the corresponding gene. After cleaning for contaminating and low quality sequences, a total of 8,821 sequences were assembled. From the 5,789 ESTs contained in the contig sequences, we estimate a 65.6% redundancy in the library. The contigs containing the highest number of ESTs are listed in Table 2, and are considered highly expressed genes.

TABLE 2

Genes highly represented

Contig ID	Number of ESTs	e-value	Description *(organism)*
44	229	$3e^{-28}$	Hypothetical protein MG09875.4 *(Magnaporthe grisea)*
80	90	0	Translation elongation factor eEF-1 alpha chain *(Aureobasidium pullula*
609	72	e^{-68}	Woronin body major protein *(Neurospora crassa)*
187	66	$4e^{-14}$	Hypothetical protein *(Neurospora crassa)*
201	58	$8e^{-30}$	Mitochondrial electron transport *(Neurospora crassa)*
49	59		Unknown protein
200	52	$1e^{-147}$	ATP carrier protein *(Gibberella zeae)*
553	47	$1e^{-109}$	Aldo / keto reductase *(Gibberella zeae)*
272	43	$2e^{-43}$	Phosphatidylserine decarboxylase proenzyme *(Aspergillus nidulans)*
181	42	e^{-42}	NAD-dependent alcohol dehydrogenase *(Aspergillus nidulans)*
362	41		Unknown protein
208	37	$1\ e^{-20}$	CG32602PA *(Drosophila melanogaster)*
999	35	e^{-100}	Endopolygalacturonase 1 *(Botryotinia fuckeliana)*
916	33	$2\ e^{-62}$	Chain A, aspartic proteinase *(Irpex lacteus)*
824	33	e^{-116}	Putative mitochondrial cyclophilin 1 *(Botryotinia fuckeliana)*
4	32		Unknown protein
198	32	$9e^{-76}$	Hypothetical dyoxigenase protein *(Neurospora crassa)*
865	32	0	Aspartic proteinase precursor *(Botryotinia fuckeliana)*
893	29	$6e^{-61}$	Dehydrogenase/reductaseoxidoreductase *(Magnaporthe grisea)*
894	29	$6e^{-68}$	Histone H3 *(Hypocrea jecorina)*
3	25	$1e^{-103}$	Cinnamyl-alcohol dehydrogenase *(Aspergillus nidulans)*

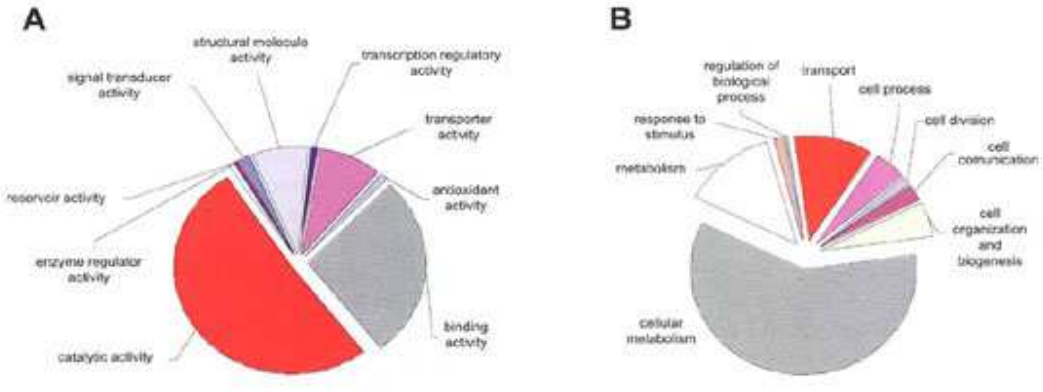
The most frequently represented EST shows a 40.8% sequence identity to a hypothetical protein from *M. grisae*. The next most frequently represented ESTs show significant similarity to genes involved in protein synthesis and energy metabolism, such as mitochondrial electron transport proteins, alcohol dehydrogenases, and Woronin body major protein (Table 2). Those genes are also highly represented in other fungal libraries (Liu and Yang, 2005; Ebbole et al., 2004). Another potentially highly expressed gene is polygalactouronase. This enzyme, involved in host cell wall degradation, has been shown to be essential for virulence (ten Have et al., 1998).

Another highly expressed sequence is the elongation factor 1-alpha, which is related to protein synthesis in general. This activity is a critical function, allowing differentiation in response to a physiological requirement. The Woronin body major protein precursor has been proposed to be a fundamental factor in the hyphal growth and pathogenicity process (Soundararajan et al., 2004), thus its presence in the library and expression at high levels is not a surprise. The ADP-ATP translocase gene is highly expressed. This gene also was found highly expressed in *T. harzarium* and *M. grisae* (Liu and Yang, 2005; Ebbole et al., 2004).

Functional identification of the genes

ESTs were assigned putative functions according to GO categories. All identified

ESTs were categorized into molecular function (Figure 1A) and physiological processes (Figure 1B), respectively. In terms of molecular functions, the sequences exhibit similarity to a broad diversity of proteins involved in binding activity (27.3%), catalytic activity (51.3%), transporter (8.6%), and structural activity (7.9%). In terms of physiological processes, the sequences obtained were classified in several groups having significant matches to proteins with known functions. The most represented ESTs correspond to cellular metabolism (60.8%), transport (11.8%), and cell organization and biogenesis (5.2%). Using this information, preliminary models for glycolisis, pyruvate, sucrose, and starch metabolism were constructed. We also found all the genes necessary for mitochondrial electron chain and ATP synthesis. Other pathways related to secondary metabolism and signal transduction were reconstructed only in part (results not shown).



* Cell processes included cell communication, cell cycle, cell death, cell differentiation and cell growth

Figure 1: Distribution of unigenes into GO categories.
The percentages are calculated from the number of unigenes in each functional category divided by the total number of unigenes in the cDNA library. A) represents molecular functions and B) represents physiological processes.

Comparison to other fungal species

Pair-wise comparisons with fungal databases were carried out to determine the *B. cinerea* unique set of ESTs with sequences similar to genes in other fungi including *M. grisae, A. nidulans, N. crassa, Saccharomyces cerevisiae,* and *Fusarium verticillioides* (Figure 2). Among 4,035 non-redundant *B. cinerea* unigenes, 1,338 (23%) sequences had similarity with *F. verticillioides* and 22% and 24% with *S. cerevisiae* and *A. nidulans,* respectively. Lower percentages of homology were to *M. grisae* and *N. crassa* (13% and 19%, respectively).

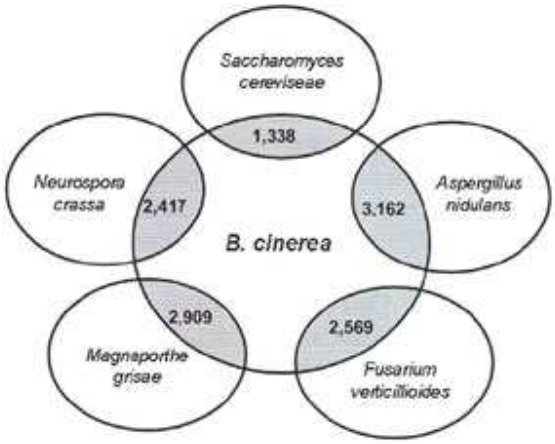


Figure 2: Unisequences shared among six fungal species. *B. cinerea*, *F. verticilloides*, *M. grisae*, and *A. nidulans* correspond to EST libraries. *S. cerevisiae* and *N. crassa* correspond to genomic data. Numbers in gray areas indicate the shared unisequences. The data was obtained from TIGR and NCBI databases.

The most frequent GO categories present in the shared sequences are those related to cell components. Among the biological process category, the more frequently shared sequences correspond to GTP binding proteins, ABC transporters, serine/threonine protein kinases, and heat-shock proteins. This result is similar to that reported previously (Austin et al., 2004).

Candidate genes with significant similarity to known fungal virulence or pathogenicity factors

ESTs were examined for genes with similarity to known fungal virulence or pathogenicity factors. Twenty virulence/pathogenicity genes were identified (Table 3). Among those, three are significantly similar to known virulence factor genes in the human pathogen *Candida albicans*, including *phr1* and *cpy1,* which are involved in yeast-to-hyphae or yeast-to-pseudohyphae transition (Mukhtar et al., 1992; Saporito-Irwin et al., 1995). *Phr1* encodes glycosylphosphatidylinositol-anchored proteins, whereas the third gene, *sap2,* is related to penetration and also is involved in adherence to epithelial cells (Zaugg et al., 2001).

In addition, sequences with similarity to pathogenicity genes from *M. grisae,* such as *gas1, pmk1,* and *pth2,* were detected. The first two are related to appressorium formation and penetration in this fungus. Deletion mutants of *gas1* had normal growth and conidiation and formed normal appressoria but were reduced in appressorial penetration and lesion formation (Xue et al., 2002). PMK1 exhibits similarity to several MAP kinases that have been described in *M. grisea,* such as OSM1 (*Hog1*) and MPS1 (*Slt2*) (Xu, 2000). Only PMK1 appears to be involved in pathogenicity of *M. grisea,* since *pmk1* mutants are impaired in appressorial formation, penetration, and invasive growth. PMK1 orthologs have been identified in many fungi and have been shown to be essential for the infection process in several

foliar pathogens, such as *Cochliobolus heterostrophus, Colletotrichum lagenarium,* and *Pyrenophora teres* (Lev et al., 1999; Takano et al., 2000; Ruiz-Roldan et al., 2001). Another gene, *path2*, that encodes an ortholog of a gene encoding carnitine acetyltranferase, an enzyme required for activated fatty acids to transverse the mitochondrial membrane for oxidation (Sweigard et al., 1998).

TABLE 3

ESTs from *Botrytis cinerea* with similarity to genes associated with fungal pathogenic and virulence

Description	Gene	Similarity to GenBank, *Organism*	Blastx Score	Blastx e value	Reference
Appresorium formation and penetration					
Pathogenicity MAP kinase 1	PMK1	U70134, *Magnaporthe grisea*	118	$1e^{-53}$	Xu and Hamer, (1996)
CAP22 expression during appressorium formation	CAP20	U18061, *Glomerella cingulata*	163	$7e^{-41}$	Hwang et al., (1995)
Secretory aspartic protease	SAP2	AF115320, *Candida albicans*	43	$9e^{-16}$	Zaugg et al., (2001)
Glycosylphosphati-dylinositol-anchored protein	GAS1	XM362679, *Magnaporthe grisae*	199	$1e^{-132}$	Xue et al., (2002)
Endopolygalactu-ronase 1	PG1	AY665552.1, *Botryotinia fuckeliana*	592	$1e^{-180}$	ten Have et al., (1998)
Cyclophilin 1	BCP1	AY277722.1, *Botryotinia fuckeliana*	258	0.0	Viaud et al., (2003)
Pectin methyl esterase	PME1	AJ309701.1, *Botryotinia fuckeliana*	491	$1e^{-140}$	Valette-Collet et al., (2(
Colonization of host tissue					
Putative mitochondrial carrier protein	FOW1	AB078975, *Fusarium oxysporum*	165	$3e^{-85}$	Inoue et al., (2002)
Mitogen-activated protein kinase	CHK1	AF178977, *Cochliobolus heterostrophus*	111	$2e^{-51}$	Lev et al., (1999)
Mitogen-activated protein kinase	UBc3	AF170532, *Ustilago maydis*	118	$3e^{-52}$	Mayorga & Gold, (199
Putative zinc finger protein	PAC1	AY005467.1, *Sclerotinia sclerotiorum*	181	$1e^{-102}$	Rollins & Dickman, (20
Defense and detoxification					
Pisatin demethylase	PDA1	AY487143, *Nectria haematococca*	61	$1e^{-22}$	Straney & Van Etten, (1
Proline transport	PHR1	AF247189, *Candida albicans*	397	$1e^{-126}$	Saporito-Irwin et al., (1
Carboxypeptidase Y	CPY1	M95182, *Candida albicans*	377	$1e^{-159}$	Mukhtar et al., (1992)
Carnitine acetyl transferase	PTH2	AF027979, *Magnaporthe grisea*	213	$1e^{-74}$	Sweigard et al., (1998)
Trichodiene oxygenase (Cytochrome P450)	TRI4	U22462.1, *Fusarium sporotrichioides*	112	$9e^{-45}$	Hohn et al., (1995)
Trichodiene oxygenase (Cytochrome P450)	CND11	XM_958484.1, *Neurospora crassa*	453	$1e^{-141}$	Galagan et al., (2003)
Terpene synthase	CND15	U05213.1, *Streptomyces sp.*	68	$3e^{-15}$	Cane et al., (1994)
General pathogenicity factors					
Aminotransferasa	Got2	NM_013177.1, *Rattus norvegicus*	149	$1e^{-107}$	Kwitek et al., (2004)
Peptidyl-propyl isomerase	PIC1	DQ140395.1, *Botryotinia fuckeliana*	130	$1e^{-54}$	Gioti et al., (2006)

Another pathogenicity gene present in the library is *fow1*, which potentially encodes a mitochondrial carrier protein in *F. oxysporum* (Inoue et al., 2002). Mutants of *fow1* have normal growth and conidiation in culture but are defective in colonizing

Sociedad de Biología de Chile

Canadá 253, piso 3º, Dpto. F.

PO Box 16164

Santiago - Chile

Tel.: (56-2) 22093503

Fax: (56-2) 22258427

socbiol@biologiachile.cl