Fungal Genetics and Biology **28,** 94–106 (1999)
Article ID fgbi.1999.1166, available online at http://www.idealibrary.com on IDEAL®

Initial Assessment of Gene Diversity for the Oomycete Pathogen *Phytophthora infestans* Based on Expressed Sequences

Sophien Kamoun,*,[1] Peter Hraber,†,[1] Bruno Sobral,† Donald Nuss,‡ and Francine Govers§

**Department of Plant Pathology, The Ohio State University, Ohio Agricultural Research and Development Center, 1680 Madison Avenue, Wooster, Ohio 44691;* †*National Center for Genomic Resources, 1800 Old Pecos Trail, Santa Fe, New Mexico 87505;* ‡*Center for Agricultural Biotechnology, University of Maryland Biotechnology Institute, Plant Sciences Building, College Park, Maryland 20742-4450; and* §*Laboratory of Phytopathology, Wageningen Agricultural University, and Graduate School Experimental Plant Sciences, Binnenhaven 9, 6709 PD Wageningen, The Netherlands*

Accepted for publication August 4, 1999

Kamoun, S., Hraber, P., Sobral, B., Nuss, D., Govers, F. 1999. Initial Assessment of Gene Diversity for the Oomycete Pathogen *Phytophthora infestans* Based on Expressed Sequences. *Fungal Genetics and Biology* 28, 94–106. A total of 1000 expressed sequence tags (ESTs) corresponding to 760 unique sequence sets were identified using random sequencing of clones from a cDNA library constructed from mycelial RNA of *Phytophthora infestans.* A number of software programs, represented by a relational database and an analysis pipeline, were developed for the automated analysis and storage of the EST sequence data. A set of 419 nonredundant sequences, which correspond to a total of 632 ESTs (63.2%), were identified as showing significant matches to sequences deposited in public databases. A putative cellular identity and role was assigned to all 419 sequences. All major functional categories were represented by at least several ESTs. Four novel cDNAs containing sequences related to elicitins, a family of structurally related proteins that induce the hypersensitive response and condition avirulence of *P. infestans* on *Nicotiana* plants, were among the most notable genes identified. Two of these elicitin-like cDNAs were among the most abundant cDNAs examined. The set also contained several ESTs with high sequence similarity to unique plant genes. © 1999 Academic Press

Index Descriptors: oomycetes; *Phytophthora infestans;* genomics; cDNA sequencing; expressed sequence tags; functional catalog; gene expression; elicitins; elicitin-like proteins.

The oomycetes comprise a diverse group of organisms that includes pathogens of plants, insects, and animals. As oomycetes form a unique group of eukaryotes with an independent evolutionary history, oomycete plant pathogens may have distinct genetic and biochemical mechanisms for interacting with plants (Judelson, 1996; Kumar and Rzhetsky, 1996; Paquin *et al.*, 1995; Van de Peer and De Wachter, 1997). Among the notable oomycete pathogens, members of the genus *Phytophthora* cause destructive diseases on thousands of plant species (Erwin and Ribeiro, 1996). *Phytophthora infestans*, the Irish potato famine fungus, causes late blight, a worldwide devastating disease of potato and tomato (Fry and Goodwin, 1997a; Fry and Goodwin, 1997b; Judelson, 1997). In recent years, severe late blight epidemics occurred following the migration to Europe and North America of aggressive A2 mating type strains (Fry and Goodwin, 1997a; Fry and Goodwin, 1997b). With the occurrence of a sexual cycle, increased virulence and gene transfer is expected in

[1] S. K. and P. H. contributed equally to this paper.

1087-1845/99 $30.00
Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

populations of *P. infestans,* resulting in new challenges for disease control.

Despite their economical importance and singular taxonomic position, *Phytophthora* species, as well as other oomycetes, remain poorly characterized at the molecular level (Judelson, 1996, 1997). Relatively few *Phytophthora* genes have been cloned and functional genetic analyses have only been performed for a single gene, *inf1,* which encodes a member of the elicitin family, a group of structurally related proteins that induce the hypersensitive response in plants and condition avirulence on nonhost *Nicotiana* plants (Kamoun *et al.,* 1994, 1997b, 1998b). Nevertheless, similar to other culturable microorganisms, basic understanding of the genetic makeup of *Phytophthora* can be rapidly increased by high throughput DNA sequencing approaches. In this paper, we describe a pilot-scale cDNA sequencing project that provides the first foray into gene diversity of *Phytophthora.* A total of 1000 expressed sequence tags (ESTs) corresponding to 760 unique sequence sets were identified using random sequencing of clones from a cDNA library constructed from mycelial RNA of *P. infestans.*

MATERIALS AND METHODS

Strains and Culture Conditions

P. infestans strain DDR7602 (US-1 genotype, A1 mating type, INF1 elicitin nonproducer) was used. This strain was isolated in 1976 from potato fields in the former German Democratic Republic (Kamoun *et al.,* 1998a). Routine culturing was conducted on rye sucrose agar media supplemented with 2% sucrose (Caten and Jinks, 1968). RNA isolation was performed on mycelium grown for 4 weeks in the synthetic medium described by Kamoun *et al.* (1994). DDR7602 did not produce sporangia under these conditions.

Construction of cDNA Library

Total RNA from *P. infestans* was isolated from mycelium using the guanidine hydrochloride extraction method (Logemann *et al.,* 1987). Poly(A)$^+$ mRNA was isolated using the oligotex mRNA purification kit (Qiagen, Valencia, CA) and following the manufacturer instructions. The cDNA library was prepared in plasmid pSPORT1 using the Superscript Plasmid System for cDNA Synthesis and Cloning (GIBCO BRL, Gaithersburg, MD) following the manufacturer protocols. Polyadenylated mRNA was used to synthesize oligo(dT) primed cDNAs, which were cloned unidirectionally in *Not*I and *Sal*I digested vector pSPORT1, with the 3′ end of the cDNA adjacent to the *Not*I site. Plasmid ligations were transformed into *Escherichia coli* DH5α. Individual colonies were transferred to 96-well microtiter plates and shipped from Wageningen Agricultural University to the University of Maryland Center for Agricultural Biotechnology, where they were processed for sequencing. EST clones were named by their position in the microtiter plate, preceded by the prefix MY (for mycelial) and the successive number of the microtiter plate (e.g., MY-06-A-04).

DNA Sequencing

Prior to sequencing, all plasmids were isolated from *E. coli* and checked for concentration and presence of an insert by electrophoresis on agarose gels. Sequencing of the cDNA clones was performed from the 5′ end using plasmid DNA as template and the M13 reverse primer. Purified plasmid DNA was sequenced using an Applied Biosystems (ABI) PRISM big dye terminator kit (Perkin-Elmer) and an ABI 377 DNA sequencer at the Center for Agricultural Biotechnology. Raw sequence data were transmitted electronically using the file transfer protocol (ftp) to the National Center for Genomic Resources for subsequent processing and analysis.

Vector Screening Algorithm

We developed a vector-screening algorithm that uses several steps to remove any vector sequence that contaminates the insert sequence. Vector sequences were identified using the BLASTN algorithm, version 2.0.3. (Altschul *et al.,* 1997), sorted and considered one at a time. The working 5′ and 3′ ends of the raw sequence were reset to exclude a similarity region if the following two criteria were met. The expect value (*E*) for the similarity region must be below a specified value, and the distance to the near end of the similarity region must be less than a given interval from the current end of the sequence. We used defaults of 0.001 and 25 nt for the *E* value and overlap interval, respectively. After excluding the vector sequence, restriction sites were considered. If an end of the insert sequence occurred near a restriction site the end was reset to the cleavage point of the restriction site. This was done to facilitate assembly of sequence fragments into contiguous sequences. Only restriction enzymes used in cDNA cloning were considered.

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

The edited insert sequence was then saved to the database, provided that the sequence was not entirely contaminated by vector and that there were no chimeric vector regions within the insert sequence. This edited sequence was used in similarity searches.

Similarity Searches

To search for similarity to other sequences in public databases, we developed an interface to BLAST that adds database connectivity and parses BLAST output into data objects. Each sequence was queried against current versions of NCBI "nr" nonredundant amino acid reference library (ftp://ftp.ncbi.nlm.nih.gov/blast/db) and a nonredundant set of nucleotide sequences compiled from the Genome Sequence Database at NCGR (http://seqsim.ncgr.org). Similarity regions that resulted from BLAST analysis were added to the database as features linked to the query sequence. The BLAST output was also stored intact in the database. We used the default BLAST parameters, but did not store any similarity regions with *E* values above 0.01.

Nonredundant sequences with highly significant matches (*E* values $<10^{-50}$) were queried against the NCBI nonredundant and dBEST databases using the BLASTX and tBLASTX algorithms (Altschul *et al.*, 1997). The BLASTX outputs were examined for biased taxonomic distribution with particular emphasis on plant and fungal sequences. ESTs showing biased BLASTX outputs were reconfirmed using the organism-specific option of the advanced BLAST search page. Sequences showing significantly different similarities to plant and fungal sequences (*E* values differed by at least 10^{10}) were manually analyzed and are listed in Tables 5 and 6. The database releases queried included the entire yeast genome, 12,998 ESTs from *Emericella (Aspergillus) nidulans*, 10,229 ESTs from *N. crassa*, and 1152 ESTs from *Magnaporthe grisea.*

Phylogenetic Analysis

Multiple alignment of 25 elicitin sequences from *P. infestans* (INF1, INF2A, INF2B, INF4, INF5, INF6, and INF7), *P. cryptogea* (CRY-B, CRY-A1, CRY-HAE20, and CRY-HAE26), *P. megasperma* (MGM-B and MGM-A), *P. cinnamomi* (CIN-B), *P. drechsleri* (DRE-B and DRE-A), *P. parasitica* (PARA1), *P. cactorum* (CAC-A), *P. capsici* (CAP-A), *P. sojae* (SOJ1, SOJ2, SOJ3, and SOJ4), and *Pythium vexans* (VEX1 and VEX2) (Kamoun *et al.*, 1997a) was conducted using the program CLUSTAL-X (J. D. Thompson *et al.*, EMBL, Heidelberg, Germany). The phylogeny of the elicitin family was reconstructed using the neighbor joining method based on multiple alignment of the conserved elicitin domains (amino acids 21–118) (Fig. 4). This was done as described in Kamoun *et al.* (1997b) using the programs SEQBOOT (for bootstrap resampling), PROTDIST (for computing distance measures between proteins), NEIGHBOR (for applying the neighbor joining method), and CONSENSE (for computing consensus trees) of the PHYLIP 3.5c software package (J. Felsenstein, University of Washington, Seattle). A total of 1000 bootstrap replications were conducted to determine the statistical significance of the obtained branches.

Data Access

Access to the database is available via the world wide web (http://www.ncgr.org/pgi). Sequence similarity searches against the data described in this work can be performed at http://seqsim.ncgr.org by defining a custom set of targets that include *P. infestans.* A fasta-formatted file containing all sequences available to date is available for distribution upon request to pgi-admin@ncgr.org. Annotated sequence data will be deposited in GenBank.

RESULTS

cDNA Library and Sequencing

As source for the cDNA library we used RNA isolated from *in vitro* grown mycelium of *P. infestans* DDR7602, a strain that does not produce the extracellular elicitor protein INF1 elicitin. DDR7602 was one of five INF1-deficient isolates that were identified from a survey of more than 100 natural isolates of *P. infestans* (Kamoun *et al.*, 1998a). In addition to lacking the INF1 protein, little *inf1* mRNA was detected in DDR7602 (Kamoun *et al.*, 1998a). Because *inf1* mRNA is highly abundant in mycelium of most wild type *P. infestans* isolates (Kamoun *et al.*, 1997b, 1998a), we selected for the sequencing project the INF1-deficient DDR7602 strain in order to avoid repeated sequencing of an abundant cDNA cluster from a known gene.

Polyadenylated mRNA was used to synthesize oligo(dT) primed cDNAs, which were cloned unidirectionally in *Not*I and *Sal*I digested vector pSPORT1. Prior to sequenc-

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

ing, the approximate sizes of the cDNA inserts were determined. Of the tested plasmids, 89.8% contained an insert larger than 100 bp and were selected for sequencing. Based on electrophoresis migration rates relative to supercoiled plasmid markers, the average cDNA insert size was 1002 bp (standard deviation = 611), with the largest insert approaching 6000 bp. Plasmids with inserts exceeding 100 bp were selected for sequence analysis.

A total of 1002 sequences corresponding to the 5′ end of the cDNA inserts were generated using primer M13 reverse. In *Phytophthora* transcripts, the 5′ untranslated regions are typically short, ranging from about 50 to 70 nucleotides (Pieterse *et al.*, 1994). Therefore, the DNA sequences obtained with primer M13 reverse are highly informative and generally overlap an open reading frame even when the cDNA is full length. In total, 669,848 nucleotides were sequenced. Of these, 585,859 nucleotides correspond to *Phytophthora* DNA, and the remaining 83,989 nucleotides to vector sequences. The mean read length was 668 nucleotides per sequence and the average insert read length was 586 nucleotides. In 61 plasmids (6.1%), the entire sequence of the insert was determined in a single pass, indicating inserts smaller than ca. 550–600 bp. The number of aberrant plasmids was low since only two sequences consisted entirely of vector and no chimeric insert/vector clones were identified.

Automated Sequence Analysis

A number of software programs were developed for the automated analysis and storage of the EST sequence data. We developed a relational database and an analysis "pipeline" (Fig. 1). The relational database was adapted from the Genome Sequence Database (GSDB), version 1.0 (Harger *et al.*, 1997). A small subset of the GSDB schema was used to store sequence and feature annotation. In addition, several tables were used to prepare and store analysis results (the "Action" database). Though the two databases have a similar identity, *Phytophthora* sequence data are not currently available from GSDB.

In the analysis pipeline, raw data were automatically processed after transfer via ftp from the sequencing laboratory to NCGR. Inputs to the pipeline were raw sequences and chromatogram traces from ABI 377 DNA sequencers, whereas outputs were annotated sequences and reports that summarize analysis results. A "Gatherer" program examined incoming files for obvious errors before adding the data to the database. Raw sequences were subsequently screened for vector contamination. Sequence regions with similarity to pSPORT1 sequences were removed before further analysis. In general, output from one step of the analysis pipeline served as input to the next step (Fig. 1). Steps in the analysis pipeline included vector

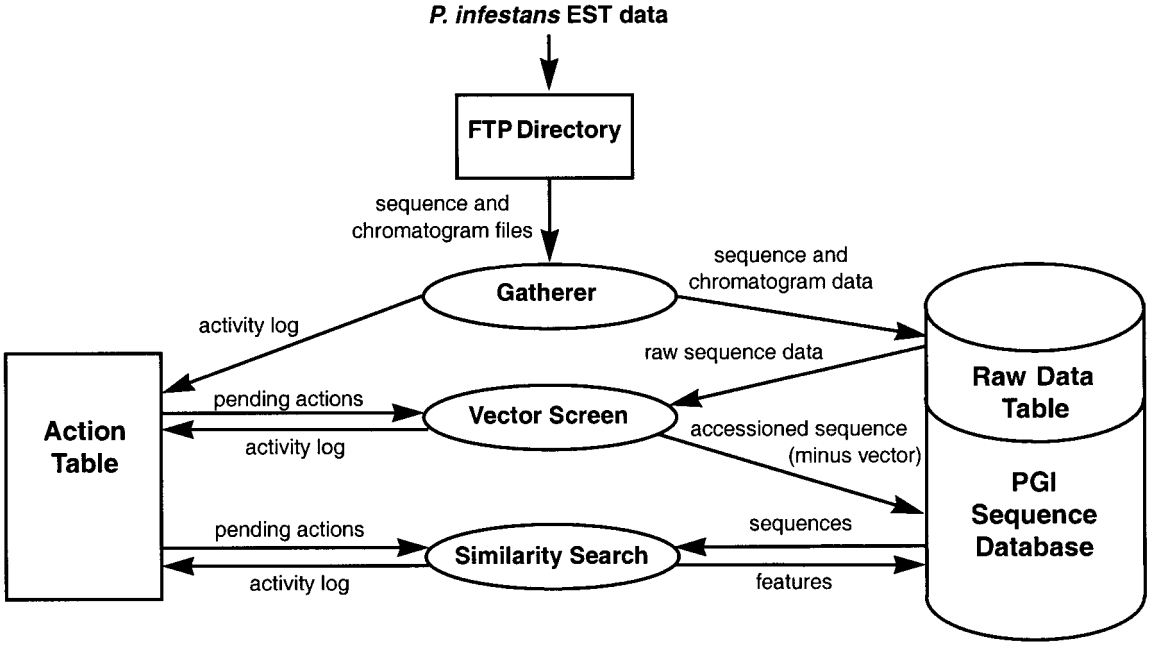

FIG. 1. Overview of the automated data analysis pipeline. See text for discussion.

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

screening and similarity searching. Access to the database is available via the world wide web (http://www.ncgr.org/pgi).

Assembly of EST Sequences

To generate a set of unique sequences from the full set of 1000 clones, we used the BLASTN comparison algorithm (Altschul *et al.*, 1997) to compare the full set of clones with itself. In interpreting the results of the BLAST search, we used the similarity score calculated from the match of a sequence against itself, or that sequence's self-score, to normalize the scores of matches to other sequences. We quantified the degree of similarity for a pair of sequences as the proportion of the score from each match compared to the self-score, which yielded values ranging from 0% (no redundancy) to 100% (absolute redundancy). Based on a preliminary analysis of similarity scores of ESTs corresponding to two known gene families, i.e., the actin and elicitin gene families (Table 1), we grouped sequences by similarity scores such that sequence matches with scores of 10% and greater were considered redundant, and scores below 10% were considered nonredundant. This procedure yielded a set of 760 unique sequences from the full set of 1000 ESTs (Table 2). Of these, 626 sequences remained as singletons and were not identical to any other EST. The remaining 374 sequences corresponded to 134 nonredundant sets of sequences that formed clusters of 2 to 10 sequences.

TABLE 1
Ratios of BLASTN/Self BLASTN Scores for ESTs Corresponding to the Elicitin and Actin Gene Families

PGI ID numbers	Putative identity	No. of ESTs in cluster	BLASTN/ Self-BLASTN scores (%)[a] Mean	Range
MY-01-C-05 (PGI:S:345)	Elicitin-like *inf5*	9	81.0	65.9–98.5
MY-01-D-04 (PGI:S:390)	Elicitin-like *inf6*	8	84.6	68.8–94.8
MY-01-D-05 (PGI:S:244)	Elicitin-like *inf2*	4	64.7	47.9–88.6
MY-02-A-12 (PGI:S:374)	Actin	6	40.9	24.4–51.3

[a] Degree of similarity was estimated by calculating the proportion of the BLASTN score of each pair of sequences compared to the self-BLASTN score of the reference sequence. Values range from 0% (no redundancy) to 100% (absolute redundancy). All ESTs outside the designated four clusters showed scores below 10%.

TABLE 2
Redundancy of cDNA Clones

	No. of duplicates									
Redundancy of cDNA clones	1	2	3	4	5	6	7	8	9	10
No. of clones	626	81	29	14	3	3	0	1	2	1

Identification of P. infestans cDNA Homologs

To identify homologs of *P. infestans* genes, each edited EST sequence (without vector) was automatically queried against the NCBI nonredundant protein database using the BLASTX algorithm (Altschul *et al.*, 1997). The BLASTX outputs were stored in a database and were accessible to all participants via the world wide web. Figure 2 shows the proportion of EST sequences with highly significant (*E* values $<10^{-20}$), moderately significant (10^{-5} to 10^{-19}), or no significant similarity ($E > 10^{-5}$) to known protein sequences. The BLASTX outputs for a total of 432 nonredundant sequences with significant ($E < 10^{-5}$) matches were then examined manually, and 14 sequences were discarded because their high BLASTX scores reflected low sequence complexity. The remaining 419 sets of nonredundant sequences, which correspond to a total of 632 ESTs (63.2%), were identified as those showing

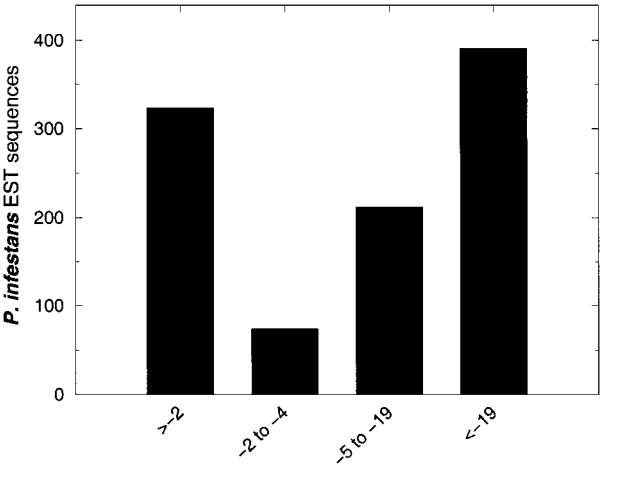

FIG. 2. Sequence similarities between translations of *Phytophthora infestans* ESTs (single pass cDNA sequences) and the best match in the NCBI protein sequence database. Similarity scores were determined using the BLASTX software. The frequency of the resulting *E* values is shown.

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

significant matches to sequences in the NCBI database and saved for functional classification.

Functional Classification of P. infestans Genes

A functional classification scheme of oomycete genes was devised based on both the Expressed Gene Anatomy Database (EGAD) (The Institute for Genomic Research (TIGR; http://www.tigr.org/tdb/egad/egad.html)) and the functional catalog of plant genes (Bevan *et al.*, 1998) (Table 3). New categories that are relevant to *Phytophthora* and other oomycetes have been added. Category 11 comprises elicitors, avirulence factors, and pathogenicity factors, whereas components involved in cell defense have been grouped in category 12. A putative cellular identity and role for 324 nonredundant sequences with significant matches was assigned based on manual examination of their highest matches in the database (for a list, see http://www.ncgr.org/pgi/reports/PinfestansESTmatches.html). A summary of the assignment of these ESTs to functional categories as well as their relative abundance is listed in Table 3 and is illustrated in Fig. 3. All major functional categories were represented by at least several ESTs. The majority of the identified cDNAs were related to protein synthesis, protein destination, metabolism, and energy. The least frequent cDNAs were those involved in cell growth/cell division/DNA synthesis, cell defense, transcription, and intracellular traffic. A total of 95 ESTs could not be classified or could not be assigned a putative identity unambiguously.

Identity of Most Abundant cDNA Clones

Besides cDNAs encoding ribosomal proteins, no single clone or group of related clones was overwhelmingly abundant in the *P. infestans* EST set. All of the most abundant cDNA clones showed significant matches in the database and could be putatively identified (Table 4). The most abundant cDNA cluster reached a frequency of 1.0% and contained sequences encoding translation elongation factor EF-1 alpha. Other relatively abundant clusters comprised two different elicitin-like cDNAs, actin cDNAs, and cDNAs containing mucin-like sequences.

Novel Elicitin-Like cDNAs

Even though the *P. infestans* isolate examined in this study was deficient in mRNA of the major elicitin INF1, a total of 23 ESTs with sequence similarity to elicitins were identified. These correspond to five different classes, and were among the most abundant transcripts identified in this pilot-scale sequencing project (average redundancy = 4.6; 2.3% of total ESTs). To further characterize these elicitin-like cDNAs, we fully sequenced six representative plasmids containing full-length cDNA inserts. One of the identified classes appeared to correspond to two closely related elicitin-like genes previously identified as *inf2A* and *inf2B* in a different strain of *P. infestans* (Kamoun *et al.*, 1997a). In addition, DNA sequence analysis revealed that the other clones contain cDNA sequences of four novel, elicitin-like genes, named *inf4*, *inf5*, *inf6*, and *inf7*. Open reading frames corresponding to predicted proteins of 118, 184, 183, and 199 amino acids were found for *inf4*, *inf5*, *inf6*, and *inf7*, respectively. Multiple alignment of all seven *P. infestans* elicitin and elicitin-like sequences (INF proteins) indicate a high degree of similarity between the different proteins. To investigate the relationships between the novel INF proteins and other elicitins, we reconstructed the phylogeny of 25 members of the elicitin family using the neighbor joining method (Fig. 4). The resulting tree clearly indicates that the novel INF4, INF5, INF6, and INF7 are distinct from the five previously defined classes of elicitins (Kamoun *et al.*, 1997a).

Taxonomic Identity of Homologs of P. infestans cDNAs

Considering the taxonomic position of oomycetes as stramenopile protists, it was interesting to systematically examine the taxonomic identity of homologs of *P. infestans* cDNAs. For this purpose, nonredundant *P. infestans* sequences were queried against the NCBI nonredundant and dBEST databases using the BLASTX and tBLASTX algorithms (Altschul *et al.*, 1997). The BLASTX outputs were then examined for biased taxonomic distribution with particular emphasis on plant and fungal sequences. ESTs showing biased BLASTX outputs (more than 10^{10} difference in *E* value between plant and fungal hits) were retested using the organism-specific option of the advanced BLAST search page. A total of 11 *P. infestans* sequences showing significantly higher similarity to plant sequences than to fungal sequences were identified (Table 5). In contrast, 7 sequences showed higher similarity to fungal sequences than to plant sequences (Table 6). It should be noted that 7 of these 18 ESTs gave the highest

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

TABLE 3

Assignment to Functional Categories and Relative Abundance of Identified *Phytophthora infestans* ESTs

Functional category	No. of ESTs	
	Nonredundant	Total
01 Metabolism		
01.01 Amino acid metabolism	23	30
01.02 Nitrogen and sulfur metabolism	0	0
01.03 Nucleotide metabolism	8	9
01.04 Phosphate metabolism	1	1
01.05 Carbohydrate metabolism	11	14
01.06 Lipid and sterol metabolism	14	14
01.07 Biosynthesis of vitamins, cofactors, and prosthetic groups	1	1
02 Energy		
02.01 Glycolysis	7	12
02.04 Gluconeogenesis	1	1
02.07 Pentose phosphate pathway	2	3
02.10 TCA pathway	13	19
02.13 Respiration	9	12
02.16 Fermentation	0	0
03 Cell growth/Cell division/DNA synthesis		
03.01 Cell growth	0	0
03.13 Meiosis	0	0
03.16 DNA synthesis/replication	1	2
03.19 Recombination/DNA repair	1	2
03.22 Cell cycle/cell cycle control	3	3
03.25 Cytokinesis	1	2
03.26 Growth regulators	1	1
03.99 Other proteins	0	0
04 Transcription		
04.01 rRNA synthesis	0	0
04.10 tRNA synthesis	0	0
04.19 mRNA synthesis		
04.19.01 General transcription factors	1	1
04.19.04 Transcriptional control	4	4
04.19.07 Chromatin modifier	0	0
04.22 mRNA processing	4	5
04.31 RNA transport	1	2
04.99 Other proteins involved in transcription	7	9
05 Protein synthesis		
05.01 Ribosomal proteins	63	137
05.04 Translational factors	12	23
05.07 Translational control	0	0
05.10 tRNA-synthetases	3	3
05.99 Other proteins involved in protein synthesis	0	0
06 Protein destination		
06.01 Folding and stabilization of proteins	24	40
06.04 Targeting/sorting/translocation of proteins	2	2
06.07 Modification of proteins	5	5
06.10 Assembly of protein complexes	2	2
06.13 Proteolysis	20	27
06.99 Other proteins involved in protein destination	0	0
07 Transport facilitators		
07.01 Ion channels	2	5
07.04 Ion transporters	2	2
07.07 Sugar/carbohydrate transporters	1	2
07.10 Amino acid transporters	2	2
07.13 Lipid transporters	0	0
07.16 Purine/pyrimidine transporters	0	0
07.22 Transport ATPases	3	3
07.25 ABC transporters	7	7
07.99 Other transport facilitators	3	5
08 Intracellular traffic		
08.01 Nuclear transport	7	8
08.04 Mitochondrial transport	2	5
08.07 Vesicular transport	7	9
08.10 Peroxisomal transport	0	0
08.13 Vacuolar transport	0	0
08.16 Extracellular transport	0	0
08.19 Cellular import	0	0
08.99 Other proteins involved in intracellular traffic	0	0
09 Cellular organization and biogenesis		
09.01 Cell wall/plasma membrane	6	13
09.02 Cytoplasm	0	0
09.04 Cytoskeleton	10	15
09.07 ER/golgi	1	2
09.10 Nuclear	0	0
09.13 Chromosome structure	1	1
09.16 Mitochondrial	1	1
09.19 Peroxisomal	1	1
09.25 Vacuolar	0	0
09.99 Other proteins involved in cellular organization and biogenesis	0	0
10 Signal transduction		
10.01 Receptors	1	1
10.04 Mediators	1	3
10.04.04 Kinases	14	16
10.04.07 Phosphatases	3	3
10.04.10 G proteins	2	2
10.99 Other signal transduction proteins	1	1
11 Elictors, avirulence, and pathogenicity factors		
11.01 Elicitins	0	0
11.02 Elicitin-like proteins	5	23
11.03 Other elicitor proteins	2	2
11.04 Avirulence factors	0	0
11.10 Pathogenicity factors	0	0
12 Cell defense		
12.01 Stress response	2	2
12.02 Detoxification	9	17
13 Classification not yet clear-cut	22	29
14 Unclassified proteins	56	66
Total without ribosomal proteins	355	498
Total	419	632

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

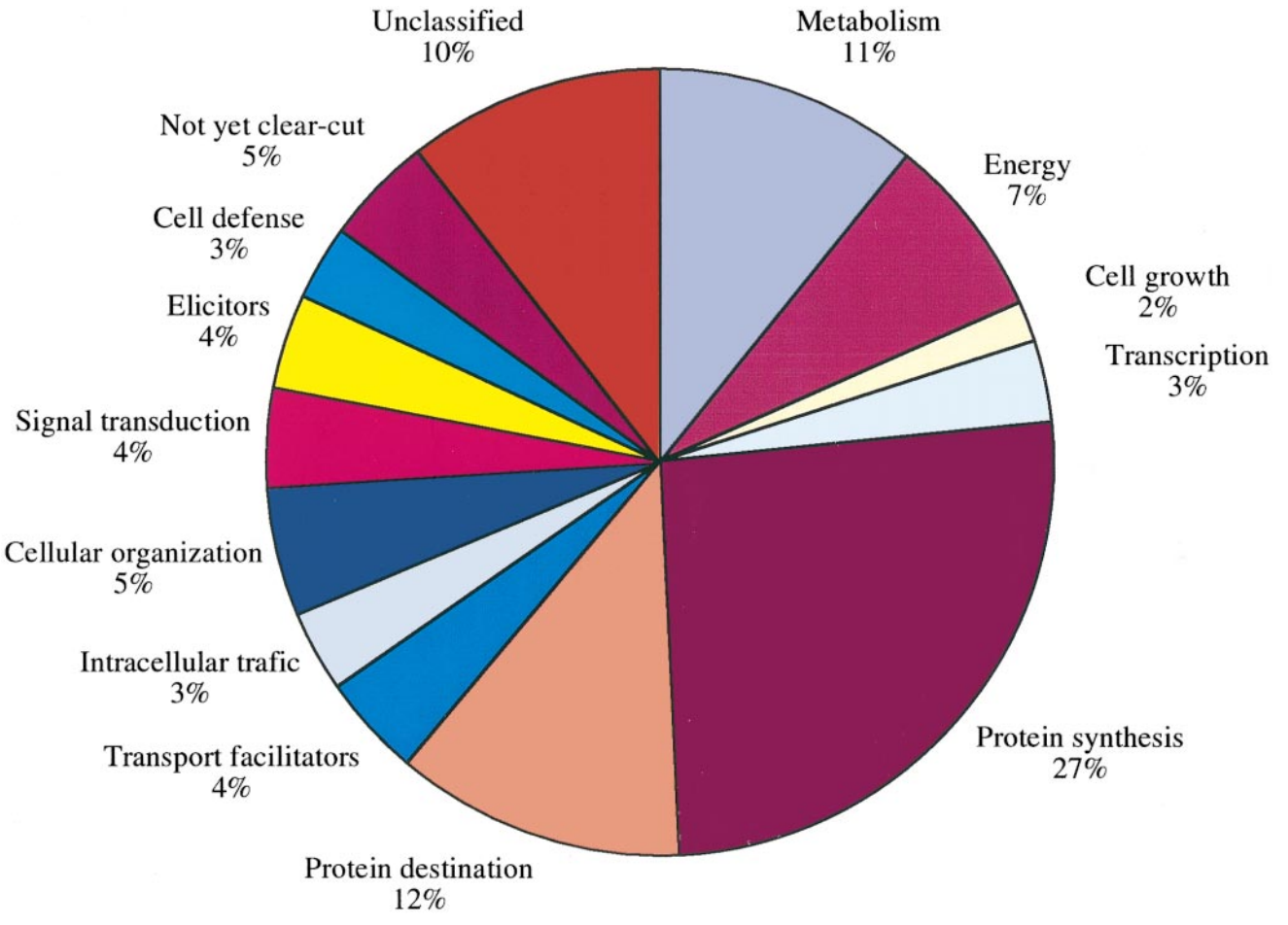

FIG. 3. Classification of 632 *Phytophthora infestans* ESTs from the mycelial library. ESTs with significant matches to known sequences (BLASTX *E* values $<10^{-5}$) were classified into functional categories as described in Table 3. The percentage of ESTs in each of the 14 categories is shown.

TABLE 4
Identity of Most Abundant *P. infestans* cDNA Clusters

PGI ID Nos.	Putative identity	Hits[a]	Frequency[b] (%)
MY-02-H-09 (PGI:S:127)	Elongation factor EF-1 alpha	10	1.0
MY-01-C-05 (PGI:S:345)	Elicitin-like *inf5*	9	0.9
MY-02-H-12 (PGI:S:75)	40S ribosomal protein	9	0.9
MY-01-D-04 (PGI:S:390)	Elicitin-like *inf6*	8	0.8
MY-02-B-11 (PGI:S:237)	Mucin-like	6	0.6
MY-02-A-12 (PGI:S:374)	Actin	6	0.6
MY-03-H-06 (PGI:S:399)	60S ribosomal protein	6	0.6

[a] Total number of ESTs in each cluster.
[b] Frequency of ESTs in examined mycelial library.

BLASTX *E* values with plant sequences, whereas none of the 18 ESTs had a fungal sequence as a best hit.

DISCUSSION

A Pilot Scale Phytophthora cDNA Sequencing Project

In this study, we describe a pilot-scale cDNA sequencing project that provides the first foray into gene diversity of *Phytophthora.* A total of 1000 ESTs corresponding to 760 unique sequences were identified using random sequencing of clones from a cDNA library constructed from mycelial RNA of *P. infestans.* For this project, we elected

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

FIG. 4. Phylogeny reconstruction of the elicitin family from *Phytophthora* spp. and *Pythium vexans.* The phylogenetic tree was constructed using the neighbor-joining method based on multiple alignment of residues 21–118 (elicitin domain) of 25 elicitin and elicitin-like proteins (Kamoun *et al.,* 1997a; this study). Bootstrap values above 50% from 1000 replications are indicated at the nodes. The length of the branches reflects weighted amino-acid substitutions, and the scale bar represents 10% weighted sequence divergence. INF7 was used as an outgroup. The five previously defined classes of elicitins (Kamoun *et al.,* 1997a) representing main clusters of the tree are indicated.

to use a collaborative approach that involved several "wet" laboratories as well as an *in silico* laboratory unit. cDNA library construction and clone sorting, DNA sequencing, and automated data analyses were performed at three separate sites, which already had in place the appropriate equipment and qualified personnel. We found this multi-unit approach to be quite effective since there was no need for the participating groups to acquire novel technologies, such as DNA sequencing or bioinformatics. Perhaps, this type of collaborative approach will prove more appropriate than a centralized unit in establishing genomics projects of nonmodel organisms or organisms that are notoriously difficult to manipulate.

Assembly of EST Sequences

To generate a set of unique sequences from the full set of 1000 clones, we used an original approach based on the BLASTN sequence comparison algorithm (Altschul *et al.,* 1997) and normalization with the reference sequence self-BLASTN score. We tested this approach using redundant ESTs corresponding to two known gene families, the elicitin and actin families. Using a BLASTN/self-BLASTN cutoff score of 10%, we were able to group the 23 ESTs corresponding to the complex elicitin family into five clusters (*inf2, inf4, inf5, inf6, inf7*) and the 6 ESTs corresponding to actin A into a separate cluster. However, this approach did not distinguish between ESTs from the closely related *inf2A* and *inf2B,* which share a 450-bp region of high identity at the 5′ end of the transcripts (437 identical residues out of 450 bp or 97%; Kamoun *et al.,* 1997a). Therefore, one should be aware that each identified cluster may not necessarily correspond to a unique gene. A detailed comparison of this approach with other widely used algorithms will be published elsewhere.

Redundancy

The rate of redundancy of *P. infestans* cDNA clones observed in this study is similar to the rates obtained from other EST sequencing programs of similar scale. Of the 1000 EST sequences, 626 are unique and do not match any other sequence. The remaining 374 sequences correspond to redundant cDNAs and form clusters ranging from 2 to 10 ESTs. This redundancy rate of 37% is similar to the rate obtained with 1409 clones of *N. crassa* (49%), 4809 clones from the cambial tissue of poplar (53%), 833 clones from developing xylem tissue of poplar (23%), and 1216 clones from flower buds of chinese cabbage (26%) (Lim *et al.,* 1996; Nelson *et al.,* 1997; Sterky *et al.,* 1998). Surprisingly, no highly abundant transcripts were identified in the examined *P. infestans* library. The most abundant *P. infestans* cDNA clone encodes elongation factor 1-alpha and was present only at a frequency of 1%. In other EST projects, the most frequent cDNA species ranged from 2.2% in a cDNA library made from a human liver cell line to 19.3% in a library made from poplar developing xylem tissue (Okubo *et al.,* 1992; Sterky *et al.,* 1998). At this stage, the significance of this finding remains unclear. Large-scale sequencing of cDNA clones of other developmental

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

TABLE 5
P. infestans cDNAs Showing Higher Similarity to Plant Genes Than to Fungal Genes

PGI ID	Putative identity	Best plant hit (*E* value)[a]	Best fungal hit (*E* value)[a,b]	Other best hit (*E* value)[a]
MY-03-E-01 (PGI:S:34)	ADP/ATP translocase	sp|P31692 *Chlorella kessleri* ADP/ATP translocase (3e-49)	sp|P02723 *Neurospora* ADP/ATP translocase (5e-28)	gi|2655149 *Drosophila* ADP/ATP translocase (3e-57)
MY-06-E-10 (PGI:S:689)	Aldehydede hydrogenase	gi|927428 *Linum usitatissimum* unnamed orf (6e-76)	sp|P32872 yeast aldehyde dehydrogenase (8e-05)	gi|730598 *Bacillus* 1-pyrroline-5-carboxylate dehydrogenase (2e-12)
MY-05-B-09 (PGI:S:607)	Argininosuccinate lyase	sp|P22675 *Chlamydomonas* argininosuccinate lyase (4e-61)	gi|431758 *Candidaalbicans* argininosuccinate lyase (2e-47)	pir|WZHURS Human argininosuccinate lyase (4e-64)
MY-04-G-12 (PGI:S:580)	Fructose-1 6-biphosphatase	sp|P46267 *Brassicanapus* fructose 1,6-bisphosphatase (8e-61)	sp|P09201 Yeast fructose bisphosphatase (7e-49)	sp|P00636 Pig fructose 1,6-bisphosphatase (6e-55)
MY-12-C-02 (PGI:S:1147)	Glutamate decarboxylase	gi|2494175 *Arabidopsis* glutamate decarboxylase (1e-83)	gi|2494176 Yeast putative glutamate decarboxylase (2e-36)	gi|1652140 *Synechocystis* sp. glutamate decarboxylase (2e-74)
MY-11-F-04 (PGI:S:1100)	Intracellular kinase	gi|1617200 tobacco Shaggy-like kinase (7e-58)	sp|P38615 yeast serine/threonine protein kinase MDS1/RIM11 (4e-35)	gi|2117783 *Xenopus* intracellular kinase (9e-60)
MY-01-C-01 (PGI:S:345)	K^+ channel protein	gi|2494112 *Arabidopsis* ATHKCP (5e-61)	pir|S61978 yeast YPL088w (2e-18)	pir|A53131 *Bos* potassium channel subunit (3e-48)
MY-12-H-10 (PGI:S:1198)	Malate dehydrogenase	gi|2827082 *Medicago sativa* malate dehydrogenase (6e-47)	sp|P17505 yeast malate dehydrogenase (1e-10)	sp|P11708 pig malate dehydrogenase (3e-50)
MY-07-F-01 (PGI:S:811)	2-Phosphoglycerate dehydratase	gi|533474 *Mesembryanthemum crystallinum* 2-phospho-D-glycerate hydrolase (8e-61)	sp|Q12560 *Aspergillus oryzae* enolase (3e-41)	gi|1839192 *Calliobothrium* sp. enolase (4e-52)
MY-04-B-01 (PGI:S:19)	Protein phosphatase	gi|2623199 *Arabidopsis* protein phosphatase X (2e-64)	sp|P48580 *Neurospora* ser/thr protein phosphatase (3e-51)	sp|P11084 rabbit ser/thr protein phosphatase (1e-63)
MY-07-G-04 (PGI:S:841)	Related to Argonaute protein	gi|2149640 *Arabidopsis* Argonaute protein (9e-66)	emb|CAA19275| *Schizosaccharomyces pombe* argonaute-like protein (7e-46)	gi|1200282 *Caenorhabditis elegans* F48F7.1 (8e-60)

[a] The accession number, homolog description, and *E* value obtained with BLASTX algorithm are indicated for the best hit obtained for plants, fungi, and other organisms.
[b] Including the entire yeast genome and filamentous fungi ESTs deposited in dBEST.

stages of *P. infestans* will be needed to get a more complete picture of the relative abundance of various transcripts in the life cycle of *P. infestans.*

Elicitin Gene Family

For this EST project, we selected a *P. infestans* strain that is deficient in the extracellular elicitor protein INF1 and that contained little or no *inf1* mRNA (Kamoun *et al.*, 1998a). As expected, we did not identify ESTs that correspond to the *inf1* gene. However, five classes of elicitin-like cDNAs were found. One of these corresponds to the previously identified class comprising the elicitin-like genes *inf2A* and *inf2B* (Kamoun *et al.*, 1997a). Four additional classes of elicitin-like proteins were novel and were clearly distinct from other members of the elicitin protein family of *Phytophthora* and *Pythium* based on phylogenetic analysis (Fig. 4). Interestingly, three of these elicitin-like cDNAs (*inf5*, *inf6*, and *inf7*) were not identified using other gene discovery approaches, such as polymerase chain reaction (PCR) amplification with degenerate primers or low-stringency hybridizations (S. Kamoun and F. Govers, unpublished data). The isolation of novel elicitin-like ESTs confirms the potential of large-scale

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

TABLE 6
P. infestans cDNAs Showing Higher Similarity to Fungal Genes Than to Plant Genes[a]

PGI ID	Putative identity	Best plant hit (*E* value)[a]	Best fungal hit (*E* value)[a]	Other best hit (*E* value)[a]
MY-10-G-09 (PGI:S:1063)	Acyltransferase	emb|CAB16844.1 *Arabidopsis* serine C-palmitoyltransferase like protein (9e-19)	gi|729923 yeast serine C-palmitoyltransferase (3e-51)	gi|2564249 human serine palmitoyltransferase (6e-67)
MY-10-D-12 (PGI:S:1001)	Cystathionine beta-synthase	gi|231970 spinach cysteine synthase (2e-28)	gi|416161 yeast cystathionine beta-synthase (1e-51)	gi|206600 rat cystathionine beta-synthase (1e-66)
MY-03-F-11 (PGI:S:17)	DEAD box helicase	sp|P41380 *Nicotiana plumbaginifolia* translation initiation factor eIF-4A (5e-27)	sp|Q07478 yeast probable ATP-dependent RNA helicase (1e-59)	gi|476338 *Caenorhabditis elegans* putative RNA helicase (7e-59)
MY-12-E-01 (PGI:S:1150)	Dynein	sp|Q39575 *Chlamydomonas reinhardtii* dynein gamma heavy chain subunit (5e-37)	gi|1169441 *Neurospora* dynein heavy chain (2e-69)	gi|461975 *Dictyostelium discoideum* dynein heavy chain (2e-74)
MY-06-E-01 (PGI:S:714)	Phosphoglycerate kinase	gi|1730512 *Arabidopsis* phosphoglycerate kinase (1e-38)	gi|1078667 *Neurospora* phosphoglycerate kinase (8e-73)	gi|3043929 *Aplysia californica* phosphoglycerate kinase (1e-77)
MY-11-C-10 (PGI:S:1307)	Transaldolase	gi|2078350 potato transaldolase (1e-04)	gi|4602 yeast transaldolase (2e-60)	gi|2073541 human transaldolase (5e-71)
MY-07-B-10 (PGI:S:793)	Ubiquitin-conjugating enzyme	gi|464981 tomato ubiquitin-conjugating enzyme (5e-35)	gi|1717864 yeast ubiquitin-conjugating enzyme (5e-59)	gi|2501432 human ubiquitin-conjugating enzyme (2e-64)

[a] For explanations see Table 5.

sequencing approaches for the discovery of novel sequences even in well-studied gene families.

The elicitin-like cDNAs *inf5* and *inf6* were among the most abundant ESTs (0.9 and 0.8%, respectively). Interestingly, in INF1 producer strains, *inf1* mRNA also occurred at high levels in *in vitro* grown mycelium, suggesting that high expression levels may be a general feature of elicitin and elicitin-like genes (Kamoun *et al.*, 1997b; S. Kamoun and F. Govers, unpublished data). The observation that elicitin and elicitin-like mRNAs are abundant may reflect the potential role of these proteins as sterol-carrier proteins (Mikes *et al.*, 1997, 1998), a function of essential importance to *Phytophthora* spp. since they cannot synthesize sterols and must assimilate them from external sources (Hendrix, 1970).

Taxonomic Identity of Homologs of *P. infestans* cDNAs

The recent taxonomic positioning of oomycetes as stramenopile protists raises issues of evolutionary relationships to algae and possibly to plants (Kumar and Rzhetsky, 1996; Paquin *et al.*, 1995; Van de Peer and De Wachter, 1997). A survey of the BLAST analysis results revealed plant genes as having the highest similarity to 147 of the 632 matched ESTs (22.9%). In addition, we compared BLASTX outputs for differences in similarities between plant and fungal sequences. Based on a difference in *E* value higher than 10^{10}, a total of 11 *P. infestans* ESTs showed significantly higher similarity to plant sequences than to fungal sequences (Table 5), whereas 7 ESTs were more similar to fungal sequences than to plant sequences (Table 6). Even though numerous *Arabidopsis* sequences are now available in the databases and our analyses included about 25,000 ESTs from filamentous fungi, a shortcoming in the interpretation of these analyses is the currently incomplete sequence data available for both plants and filamentous fungi. Therefore, the biased distribution observed for some high-similarity matches of *P. infestans* sequences may reflect a lack of data rather than true phylogenetic relationships. However, based on this preliminary study, it is clear that *P. infestans* cDNAs are not particularly related to fungal sequences and that a bias toward plant sequences may occur in concordance with the modern taxonomic view of oomycetes.

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

Similarity to Plant Genes

Several ESTs with very high sequence similarity to unique plant genes merit further comment. EST MY-07-G-04 shows very high sequence similarity with the *Arabidopsis thaliana* developmental gene *Argonaute* (*E* value of 10^{-66}). Mutations at this locus result in abnormal development of leaves and floral organs and failure to form axillary meristems (Bohmert *et al.*, 1998). Related genes of unknown function have been identified in other multicellular organisms ranging from *Caenorhabditis elegans* to humans (Bohmert *et al.*, 1998; Moussian *et al.*, 1998). The absence of family members in the sequenced genomes of several bacteria and the budding yeast *S. cerevisiae* has led to the suggestion that *Argonaute* and related genes encode functions specific for multicellular organisms. A second *P. infestans* EST, MY-06-A-05, showed less, but significant, conservation (*E* value of 10^{-21}) to an independent portion of the *Argonaute* gene, suggesting the possibility of a multigene family. The gene with highest similarity score for EST MY-10-E-04 (*E* value of 10^{-46}) encodes a barley protein that is rapidly induced after infection by the powdery mildew fungus (*Erysiphe graminis*) (Walther-Larsen *et al.*, 1993). Interestingly, this protein contains the heat shock HSP90 family signature and is also closely related (*E* value of 10^{-43}) to the protein encoded by the *mod-E1* gene required for vegetative incompatibility in the fungus *Podospora anserina* (Loubradou *et al.*, 1997). Several ESTs showed highest similarity to plant genes involved in salt stress responses. These included the *S*-adenosylmethionine synthetase gene of tomato (Espartero *et al.*, 1994) for MY-11-B-09 (*E* value of 10^{-75}) and the delta 1-pyrroline-5-carboxylate synthetase A gene of *A. thaliana* (Savoure *et al.*, 1995) for MY-13-C-01 (E value of 10^{-41}). The latter gene plays a key role in proline biosynthesis, leading to osmoregulation in plants.

Future Prospects

The information generated in this *P. infestans* EST project demonstrates the power of this gene discovery methodology for an organism that has not been studied extensively at the biochemical, cellular or molecular level. Prior to this work only 48 *P. infestans* DNA sequences were deposited in public databases, of which more than half were noncoding sequences. The finding that almost 37% of the ESTs show no similarity to known sequences, not even to the large number of ESTs that have been generated in the past years from other eukaryotic organisms, suggests that there are many genes yet to be discovered in phylogenetically diverse organisms. In view of this it will be interesting to see how many more unknown sequences can be found among genes expressed in other developmental stages of the life cycle of *P. infestans*, in particular in stages when the pathogen is in close interaction with the host. Therefore, future efforts in *Phytophthora* genomics should focus on large-scale random sequencing of stage-specific cDNA libraries. The knowledge gained will not be limited to sequence information. Valuable gene expression profiling data will also become available.

As opposed to the relatively high percentage of unknown sequences, the high similarity score with plant genes is remarkable and may provide clues to possible functional roles for novel *P. infestans* genes that would not be forthcoming by traditional means. Recent reports of homology dependent gene silencing in *P. infestans* (van West *et al.*, 1999; Kamoun *et al.*, 1998b) suggest that new opportunities for efficient functional analyses of these and other genes with highly intriguing sequence similarities may be close at hand.

Phytophthora infestans is an important plant pathogen. Effective control of late blight is hampered by the limited availability of resistant potato cultivars. In addition, the lack of specific targets for control requires the use of nonspecific, mostly harmful agrochemicals. An EST sequencing approach, such as the project described here, provides the fastest way of identifying components that regulate vital processes such as sporulation, spore germination, and virulence. This knowledge will provide additional insight into the development of novel control strategies, not only for *P. infestans* but also for other oomycete plant pathogens.

ACKNOWLEDGMENTS

Our thanks to Koen de Groot, Laura Rauser, and Ellen Rosenbloom for technical assistance.

REFERENCES

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **17:** 3389–3402.

Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C., Bergkamp, R., Dirkse, W., Van Staveren, M., Stiekema, W., Drost, L.,

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

Ridley, P., Hudson, S. A., Patel, K., Murphy, G., Piffanelli, P., Wedler, H., Wedler, E., Wambutt, R., Weitzenegger, T., Pohl, T. M., Terryn, N., Gielen, J., Villarroel, R., and Chalwatzis, N. 1998. Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* **391:** 485–488.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Coboche, M., and Benning, C. 1998. AGRO1 defines a novel locus of *Arabidopsis* controlling leaf development. *EMBO J.* **17:** 170–180.

Caten, C. E., and Jinks, J. L. 1968. Spontaneous variability of single isolates of *Phytophthora infestans.* I. Cultural variation. *Can. J. Bot.* **46:** 329–347.

Erwin, D. C., and Ribeiro, O. K. 1996. *Phytophthora Diseases Worldwide.* APS Press, St. Paul, MN.

Espartero, J., Pintor-Toro, J. A., and Pardo, J. M. 1994. Differential accumulation of S-adenosylmethionine synthetase transcripts in response to salt stress. *Plant Mol. Biol.* **25:** 217–227.

Fry, W. E., and Goodwin, S. B. 1997a. Re-emergence of potato and tomato late blight in the United States. *Plant Dis.* **81:** 1349–1357.

Fry, W. E., and Goodwin, S. B. 1997b. Resurgence of the Irish potato famine fungus. *Bioscience* **47:** 363–371.

Harger, C., Skupski, M., Allen, E., Clark, C., Crowley, D., Dickinson, E., Easley, D., Espinosa-Lujan, A., Farmer, A., Fields, C., Flores, L., Harris, L., Keen, G., Manning, M., McLeod, M., O'Neill, J., Pumilia, M., Reinert, R., Rider, D., Rohrlich, J., Romero, Y., Schwertfeger, J., Seluja, G., Siepel, A., and Schad, P. A. 1997. The Genome Sequence DataBase version 1.0 (GSDB): From low pass sequences to complete genomes. *Nucleic Acids Res.* **25:** 18–23.

Hendrix, J. W. 1970. Sterols in growth and reproduction of fungi. *Annu. Rev. Phytopathol.* **8:** 111–130.

Judelson, H. J.1996. Recent advances in the genetics of oomycete plant-pathogens. *Mol. Plant–Microbe Interact.* **9:** 443–449.

Judelson, H. J. 1997. The genetics and biology of *Phytophthora infestans:* Modern approaches to a historical challenge. *Fung. Genet. Biol.* **22:** 65–76.

Kamoun, S., Lindqvist, H., and Govers, F. 1997a. A novel class of elicitin-like genes from *Phytophthora infestans. Mol. Plant–Microbe Interact.* **10:** 1028–1030.

Kamoun, S., van der Lee, T., van den Berg-Velthuis, G., de Groot, K. E., and Govers, F. 1998a. Loss of production of the elicitor protein INF1 in the clonal lineage US-1 of *Phytophthora infestans. Phytopathology* **88:** 1315–1323.

Kamoun, S., van West, P., de Jong, A. J., de Groot, K., Vleeshouwers, V. G. A. A., and Govers, F. 1997b. A gene encoding a protein elicitor of *Phytophthora infestans* is down-regulated during infection of potato. *Mol. Plant–Microbe Interact.* **10:** 13–20.

Kamoun, S., van West, P., Vleeshouwers, V. G. A. A., de Groot, K. E., and Govers, F. 1998b. Resistance of *Nicotiana benthamiana* to *Phytophthora infestans* is mediated by the recognition of the elicitor protein INF1. *Plant Cell* **10:** 1413–1426.

Kamoun, S., Young, M., Forster, H., Coffey, M. D., and Tyler, B. M. 1994. Potential role of elicitins in the interaction between *Phytophthora:* Species and tobacco. *Appl. Env. Microbiol.* **60:** 1593–1598.

Kumar, S., and Rzhetsky, A. 1996. Evolutionary relationships of eukaryotic kingdoms. *J. Mol. Evol.* **42:** 183–193.

Lim, C. O., Kim, H. Y., Kim, M. G., Lee, S. I., Chung, W. S., Park, S. H., Hwang, I., and Cho, M. J. 1996. Expressed sequence tags of Chinese cabbage flower bud cDNA. *Plant Physiol.* **111:** 577–588.

Logemann, J., Schell, J., and Willmitzer, L. 1987. Improved method for isolation of RNA from plant tissue. *Anal. Biochem.* **163:** 16–20.

Loubradou, G., Begueret, J., and Tureq, B. 1997. A mutation in an HSP90 gene affects the sexual cycle and suppresses vegetative incompatibility in the fungus *Podospora anserina. Genetics* **147:** 581–588.

Mikes, V., Milat, M. L., Ponchet, M., Panabieres, F., Ricci, P., and Blein, J. P. 1998. Elicitins, proteinaceous elicitors of plant defense, are a new class of sterol carrier proteins. *Biochem. Biophys. Res. Commun.* **245:** 133–139.

Mikes, V., Milat, M. L., Ponchet, M., Ricci, P., and Blein, J. P. 1997. The fungal elicitor cryptogein is a sterol carrier protein. *FEBS Lett.* **416:** 190–192.

Moussian, B., Schoof, H., Haecker, A., Jurgens, G., and Laux, T. 1998. Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during *Arabidopsis* embryogenesis. *EMBO J.* **17:** 1799–1809.

Nelson, M. A., Kang, S., Braun, E. L., Crawford, M. E., Dolan, P. L., Leonard, P. M., Mitchell, J., Armijo, A. M., Bean, L., Blueyes, E., Cushing, T., Errett, A., Fleharty, M., Gorman, M., Judson, K., Miller, R., Ortega, J., Pavlova, I., Perea, J., Todisco, S., Trujillo, R., Valentine, J., Wells, A., Werner-Washburne, M., and Natvig, D. O. 1997. Expressed sequences from conidial, mycelial, and sexual stages of *Neurospora crassa. Fung. Genet. Biol.* **21:** 348–363.

Okubo, K., Hori, N., Matoba, R., Niiyama, T., Fukushima, A., Kojima, Y., and Matsubara, K. 1992. Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nat. Genet.* **2:** 173–179.

Paquin, B., Roewer, I., Wang, Z., and Lang, B. F. 1995. A robust fungal phylogeny using the mitochondrially encoded NAD5 protein sequence. *Can. J. Bot.* **73** (Suppl. 1): S180–S185.

Pieterse, C. M. J., van West, P., Verbakel, H. M., Brasse, P. W. H. M., van den Berg Velthuis, G. C. M., and Govers, F. 1994. Structure and genomic organization of the *ipiB* and *ipiO* gene clusters of *Phytophthora infestans. Gene* **138:** 67–77.

Savoure, A., Jaoua, S., Hua, X. J., Ardiles, W., Van Montagu, M., and Verbruggen, N. 1995. Isolation, characterization, and chromosomal location of a gene encoding the delta 1-pyrroline-5-carboxylate synthetase in *Arabidopsis thaliana. FEBS Lett.* **372:** 13–19.

Sterky, F., Regan, S., Karlsson, J., Hertzberg, M., Rohde, A., Holmberg, A., Amini, B., Bhalerao, R., Larsson, M., Villarroel, R., Van Montagu, M., Sandberg, G., Olsson, O., Teeri, T. T., Boerjan, W., Gustafsson, P., Uhlen, M., Sundberg, B., and Lundeberg, J. 1998. Gene discovery in the wood-forming tissues of poplar: Analysis of 5,692 expressed sequence tags. *Proc. Natl. Acad. Sci. USA* **95:** 13330–13335.

Van de Peer, Y., and De Wachter, R. 1997. Evolutionary relationships among the eukaryotic crown taxa taking into account site-to-site rate variation in 18S rRNA. *J. Mol. Evol.* **45:** 619–630.

van West, P., Kamoun, S., van't Klooster, J. W., and Govers, F. 1999. Inter nuclear gene silencing in *Phytophthora infestans. Mol. Cell* **3:** 339–348.

Walther-Larsen, H., Brandt, J., Collinge, D. B., and Thordal-Christensen, H. 1993. A pathogen-induced gene of barley encodes a HSP90 homologue showing striking similarity to vertebrate forms resident in the endoplasmic reticulum. *Plant Mol. Biol.* **21:** 1097–1180.

Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.