Genetics and Molecular Biology, 33, 4, 795-806 (2010)

www.sbg.org.br

Research Article

In silico identification of coffee genome expressed sequences potentially associated with resistance to diseases

Samuel Mazzinghy Alvarenga[1], Eveline Teixeira Caixeta[1,2], Bárbara Hufnagel[1], Flávia Thiebaut[1], Eunize Maciel-Zambolim[1,3], Laércio Zambolim[1,3] and Ney Sussumu Sakiyama[1,4]

[1]*Instituto de Biotecnologia Aplicada à Agropecuária, Universidade Federal de Viçosa, Viçosa, MG, Brazil.*
[2]*Empresa Brasileira de Pesquisa Agropecuária, Embrapa Café, Brasília, DF, Brazil.*
[3]*Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG, Brazil.*
[4]*Departamento de Fitotecnia, Universidade Federal de Viçosa, Viçosa, MG, Brazil.*

Abstract

Sequences potentially associated with coffee resistance to diseases were identified by *in silico* analyses using the database of the Brazilian Coffee Genome Project (BCGP). Keywords corresponding to plant resistance mechanisms to pathogens identified in the literature were used as baits for data mining. Expressed sequence tags (ESTs) related to each of these keywords were identified with tools available in the BCGP bioinformatics platform. A total of 11,300 ESTs were mined. These ESTs were clustered and formed 979 EST-contigs with similarities to chitinases, kinases, cytochrome P450 and nucleotide binding site-leucine rich repeat (NBS-LRR) proteins, as well as with proteins related to disease resistance, pathogenesis, hypersensitivity response (HR) and plant defense responses to diseases. The 140 EST-contigs identified through the keyword NBS-LRR were classified according to function. This classification allowed association of the predicted products of EST-contigs with biological processes, including host defense and apoptosis, and with molecular functions such as nucleotide binding and signal transducer activity. Fisher's exact test was used to examine the significance of differences in contig expression between libraries representing the responses to biotic stress challenges and other libraries from the BCGP. This analysis revealed seven contigs highly similar to catalase, chitinase, protein with a BURP domain and unknown proteins. The involvement of these coffee proteins in plant responses to disease is discussed.

Key words: Coffea, data mining, ESTs, genomics, *in silico*, bioinformatics.

Received: September 2, 2009; Accepted: July 8, 2010.

Introduction

Coffee (*Coffea* spp.), which is produced in more than 60 countries, is one of the most important exportation products in the world. However, its production is frequently adversely affected by diseases and pests. Plant resistance genes (*R* genes) are a very important group of genes that have been used by breeders to develop resistant cultivars as part of strategies for disease control.

Flor (1942) proposed a highly specific recognition system between plant *R* gene products and pathogen elicitor proteins (Avr effectors). In Flor's gene-for-gene theory, the plant with a dominant *R* gene recognizes the pathogen's corresponding dominant *Avr* gene. The presence of the *Avr* gene makes the pathogen non-virulent if the plant has the appropriate *R* gene. If there is no *R* gene in the plant and/or no *Avr* gene in the pathogen, then there is no recognition and disease is initiated (Jia *et al.*, 2000; Deslandes *et al.*, 2003; Belkhadir *et al.*, 2004).

The interaction between the *R* and *Avr* gene product triggers a signal transduction pathway involving several proteins that leads to a plant defense response (Caplan and Dinesh-Kumar, 2006). Some of the proteins induced in response to infection by pathogens include thaumatin, a pathogenesis-related protein, chalcone synthase, a key enzyme in the biosynthesis of diverse flavonoids involved in disease resistance, chitinases and glucanases, which exhibit antifungal activity by attacking components in fungal cell walls, and proteins such as polyphenol oxidases, glucosyl transferases and phytoalexins, known to be involved in plant defense against pests and pathogens.

Plant *R* gene identification and isolation has increased in the last decade (Richter and Ronald, 2000) and the sequencing of several plant genomes has produced a large amount of information on the functions of gene products, and the regulation of biosynthetic and signal transduction pathways (Staskawicz *et al.*, 1995). *R* genes encode at least

Send correspondence to Eveline Teixeira Caixeta. Laboratório de Biotecnologia do Cafeeiro, Instituto de Biotecnologia Aplicada à Agropecuária, Universidade Federal de Viçosa, 36570-001 Viçosa, MG, Brazil. E-mail: eveline.caixeta@embrapa.br.

five types of proteins (R proteins) (Dangl and Jones, 2001) that can be classified according to their structural characteristics. The first R gene to be cloned was the tomato *Pto* gene, which codes for a serine/threonine kinase protein and confers resistance to races of *Pseudomonas syringae* carrying the *AvrPto* avirulence gene (Martin *et al.*, 1993). This gene belongs to the first class (class 1) of R proteins.

Class 2 R proteins include the tomato *Cf9* gene, which confers resistance to *Cladosporium fulvum* and encodes transmembrane proteins, the extracellular N-terminal region of which has leucine-rich repeats (LRR) (Jones *et al.*, 1994). However, these proteins have no significant intracellular region that can serve as a signaling component to activate plant defense mechanisms (Ellis and Jones, 1998).

Class 3, represented by the rice *Xa21* gene (which confers resistance to *Xanthomonas oryzae*), includes proteins with a classic receptor-kinase, one extracellular LRR region, one transmembrane region and a cytosolic serine/threonine kinase domain (Song *et al.*, 1995; Ellis and Jones, 1998).

Most of the resistance genes code for proteins with a variable N-terminal domain followed by a tripartite nucleotide binding site (NBS) and an LRR region (Baker *et al.*, 1997; Ellis *et al.*, 2000; Dangl and Jones, 2001). Based in their N-terminal domain, these proteins can be classified in two groups. The TIR-NBS-LRR group (class 4) has a sequence similar to the Toll protein of *Drosophila* and to the Interleukin-1 Receptor in mammals (TIR). This group TIR-NBS-LRR includes the tobacco *N* gene (which confers resistance to *tobacco mosaic virus* - TMV), the genes *L6* and *M* (resistance to rust) from flax and the genes *RPP5*, *RPP1* (both conferring resistance to *Peronospora parasitica*) and *RPS4* (resistance to *Pseudomonas syringae*) from *Arabidopsis* (Ellis and Jones, 1998; Gassmann *et al.*, 1999).

The other group (class 5) includes non-TIR-NBS-LRR proteins that lack the TIR sequence (Meyers *et al.*, 1999; Pan *et al.*, 2000; Cannon *et al.*, 2002; Richly *et al.*, 2002) and/or contain a leucine zipper domain (LZ) or a coiled coil (CC) domain in their N-terminal region (Pan *et al.*, 2000). CCs are structures formed by 2-5 helixes that display a distinct folding of the amino acid side chains at the helix-helix interface (Lupas, 1996). The CC structure has an organized repetition of seven amino acid residues, two of which have hydrophobic side chains that form an interface for interaction between coils (Young, 2000).

The CC-NBS-LRR class includes the *RPM1* and *RPS2* genes, and the LZ-NBS-LRR class includes the *Arabidopsis RPS5* and tomato *Prf* genes (both confer resistance to *P. syringae*), the *Arabidopsis RPP8* gene (resistance to *P. parasitica*), the tomato *Mi* gene (resistance to the nematode *Meloidogyne incognita*) and the potato *Rx1* gene (resistance to *potato virus* X – PVX) (Ellis and Jones, 1998; Bendahmane *et al.*, 1999). The tomato *I2*, maize *Rp1-D*, lettuce *RGC2* and pepper *Bs2* genes contain neither the TIR nor the LZ region (Tai *et al.*, 1999). In addition to these two regions, some R gene products contain a conserved domain of unknown function, referred to as GLPLAL, between the NBS and LRR regions (Dinesh-Kumar *et al.*, 2000).

ESTs similar to the NBS-LRR class of R proteins have been found in coffee. By using several degenerate primers for two conserved motifs within the NBS regions of R genes from different species, Noir *et al.* (2001) amplified and cloned sequences from nine distinct motifs of NBS proteins in coffee. These authors identified 18 RGAs (Resistance Gene Analogs) in a single coffee accession, a finding that suggested the presence of a large number of R genes in the coffee genome.

Fernandez *et al.* (2004) identified two coffee ESTs highly similar to the proteins DND1 (Defense, No Death) and NDR1 (Non race-specific Disease Resistance), which are resistance signaling components in *Arabidopsis thaliana*. DND1 protein is an ion channel involved in the hypersensitivity reaction signaling pathway whereas NDR1 is a key-component in the signaling pathway of several CC-NBS-LRR proteins (Silva *et al.*, 2006). Guzzo (2004) isolated coffee genes with functions related to plant defense against phytopathogens. These genes are involved in processes such as the hypersensitivity reaction, programmed cell death, anti-microbial protein synthesis and controlled protein degradation. Lin *et al.* (2005) identified coffee genes with putative functions related to disease resistance (such as TIR-NBS-LRR proteins), pathogenesis and other activities.

A better understanding of the mechanisms of disease resistance in coffee plants and the development of molecular tools for identifying these mechanisms may be useful in genetic breeding programs for this crop plant. In this regard, genomic analysis can be a powerful ally in investigating *R* genes and plant-pathogen interactions. The recent application of genome technology to plants has generated considerable information and has provided DNA sequence databases that have allowed the identification of genetic factors related to plant characteristics of agricultural interest.

The use of technologies such as expressed sequence tags (ESTs), which are short sub-sequences of transcribed cDNA sequences, has greatly reduced the time required to locate and describe plant genes. ESTs can be used to identify gene transcripts and are instrumental in gene discovery and gene sequence determination (Adams *et al.*, 1991). Due to their efficiency, ESTs have been widely in creating public databases (Wolfsberg and Landsman, 1997). In plants, ESTs were initially reported for *Arabidopsis* (Höfte *et al.*, 1993) and rice (Yamamoto and Sasaki, 1997), and this was rapidly followed by similar projects for corn (Gai *et al.*, 2000), soybean (Shoemaker *et al.*, 2002), wheat (Lazo *et al.*, 2004), potato (Ronning *et al.*, 2003) and cotton (Udall *et al.*, 2006). The data provided by such studies can be readily accessed via the EST database (dbEST) of the National Center for Biotechnology Information (NCBI), and has facilitated the identification of genes responsible for selected agronomical traits and their subsequent manipulation by molecular genetic techniques.

In the Brazilian Coffee Genome Project (BCGP), approximately 33,000 unigenes were identified in 214,964 ESTs obtained from 37 cDNA libraries of *Coffea arabica*, *Coffea canephora* and *Coffea racemosa* under different physiological conditions. The clustering of these ESTs yielded 17,982 EST-contigs and 32,155 singlets (Vieira *et al.*, 2006). In this work, we used *in silico* analysis to screen the BCGP database for the genes of proteins potentially involved in coffee resistance to diseases.

Material and Methods

Data mining

All of the data from the BCGP, as well as various bioinformatics tools, were provided by the Laboratório de Genômica e Expressão (LGE) and by Embrapa Recursos Genéticos e Biotecnologia (Cenargen).

Initially, we searched the literature to identify proteins related to the mechanisms of plant disease resistance. These proteins were subsequently used as keywords to mine the ESTs in the BCGP. Data mining was done using Gene Projects, a sequence management and manipulation system that is part of the LGE bioinformatics platform (Carazzolle *et al.*, 2007). Based on the keywords found, we ran a search that targeted the annotation of each EST in the LGE database. The ESTs identified in each keyword-based search were placed in virtual folders referred to as “projects”. Each project was identified by a keyword and contained its related ESTs. The projects created were: Chitinase, Cytochrome P450, Glucanase, Glucosyltransferase, HSP (Heat Shock Protein), Hypersensitive, Importin, NBS-LRR (nucleotide binding site-leucine rich repeat), Pathogenesis, Polyphenoloxidase, Phytoalexin, Resistance, Thaumatin and Chalconesyntase. The ESTs of each project were clustered with the CAP3 assembly program (Huang and Madan, 1999) to form EST-contigs that corresponded to consensus sequences with improved length and quality. These EST-contigs were compared to the GenBank nr sequence database (version of August 26, 2002) by BLAST (BLASTX version 2.2.4) (Altschul *et al.*, 1990) with the BLOSUM 62 scoring matrix. The EST clustering and BLAST searches were done through Gene Projects. Each EST-contig annotation was examined in order to obtain relevant information about the putative disease resistance genes. Only EST-contigs with E-values $< e^{-20}$ and scores > 100 were considered for further analysis.

Functional characterization

Functional characterization of the selected EST-contigs from the NBS-LRR Project was done with the Blast2GO software (Conesa *et al.*, 2005). This program finds sequences similar to each EST-contig by Blast searching against the NCBI nr database. The maximum E-value from the best blast hit was set to $1e^{-10}$ and the minimum alignment size (HSP length) was set to 33. Based on the BlastX results, the Blast2GO extracted terms from the Gene Ontology (GO) for each EST-contig. The distribution of GO terms was analyzed at level 3 of the Directed Acyclic Graphs (DAGs). The three categories of electronically designed terms were Molecular Function, Biological Process and Cellular Component. Conserved domains were identified by screening the EST-contigs with InterProScan from the EBI (European Bioinformatics Institute). InterPro is a database that contains information about domains, motifs and regions that are conserved in protein families. The data for InterPro are obtained from other databases that include PROSITE, PRINTS, ProDom, Pfam, SMART, TIGRFAMs, PIRSF, UPERFAMILY and PANTHER. The last step of the program consisted of annotating the ECs (Enzyme Codes) and searching for KEGG (Kyoto Encyclopedia of Genes and Genomes) metabolic maps for the EST-contigs.

Fisher’s exact test

Fisher’s exact test (Fisher, 1922), available from the Cenargen Bioinformatics platform, was used to compare the levels of contig expression between libraries generated following exposure to biotic stress and other libraries from the BCGP. The major libraries generated after biotic stress were RM1 (leaves infected with leaf miner and coffee leaf rust), NS1 (roots infected with nematodes) and RX1 (stems infected with *Xylella* spp.) while the other libraries consisted of the remaining coffee genome libraries (Vieira *et al.*, 2006). A value of $p = 0.05$ indicated significance.

Results and Discussion

Data mining

A total of 11,300 ESTs were mined and deposited in 14 projects. The projects CytochromeP450, Resistance and Chitinase contained the highest number of ESTs, with 2,441 (21.6%), 1,864 (16.5%) and 1,855 (16.4%) ESTs, respectively (Table 1). The projects with the smallest number of ESTs were Phytoalexin, Polyphenoloxidase and Hypersensitive, with 12 (0.10%), 67 (0.59%) and 86 (0.76%) ESTs, respectively.

With the keyword NBS-LRR it was possible to form 140 EST-contigs. The LRR (Leucine-Rich Repeat) domain participates in the interaction of the R protein with its related Avr gene product (elicitor). Sequence variability within the interstitial leucine residues can affect ligand binding and may confer different recognition specificities to each Avr factor (Parniske *et al.*, 1997).

In several plant-pathogen systems, variation in the LRR region sequence is responsible for different recogni-

Table 1 - The number of ESTs and their relative percentages obtained by keyword data mining in 14 keyword projects, the number of clusters (EST-contigs and singlets) formed and the number and percentage of EST-contigs with E-values < e^{-20} and scores > 100.

Project	ESTs	%	EST-contigs Total	Singlets	EST-contigs[1]	%
Chalconesynthase	153	1.35	5	8	5	0.51
Chitinase	1,855	16.41	47	48	45	4.59
CytochromeP450	2,441	21.60	235	202	144	14.70
Glucanase	642	5.68	92	68	88	8.98
Glucosyltransferase	1,286	11.38	130	160	125	12.76
HSP (Heat Shock Protein)	240	2.12	31	27	30	3.06
Hypersensitive	86	0.76	8	4	8	0.81
Importin	532	4.70	11	14	10	1.02
NBS-LRR	826	7.30	160	243	140	14.30
Pathogenesis	979	8.66	63	37	61	6.23
Phytoalexin	12	0.10	3	2	3	0.30
Polyphenoloxidase	67	0.59	4	3	4	0.40
Resistance	1,864	16.49	347	416	300	30.64
Thaumatin	317	2.80	16	7	16	1.63
Total	11,300	100	1,152	1,239	979	100

[1] E-value < e^{-20} and score > 100.

tion or resistance specificities (Parniske *et al.*, 1997; Thomas *et al.*, 1997; Botella *et al.*, 1998; Wang *et al.*, 1998; Ellis *et al.*, 1999). The NBS (nucleotide binding site) domain is believed to participate in the activation of signal transduction compounds, leading to pathogen-specific resistance responses (Aarts *et al.*, 1998; Feys and Parker, 2000; Van Der Biezen *et al.*, 2000). Using the NBS-LRR bait, EST-contigs similar to disease resistance proteins, NBS-LRR proteins, chitinases and other proteins were identified (Table S1). The conserved domains found in the EST-contigs included pfam00931 (NB-ARC domain), COG4886 (LRR protein), cd00116 (LRR_RI, leucine-rich repeats, ribonuclease inhibitor (RI)-like subfamily), smart00220 (S_TKc, serine/threonine protein kinases) and pfam01582 (TIR domain).

The keyword Resistance was used as a general term to obtain ESTs related to plant defense mechanisms and resulted in the identification of 300 EST-contigs that included several classes of disease resistance proteins, such as CC-NBS-LRR, TIR-NBS-LRR and nematode and virus resistance genes (Table S2). The conserved domains observed in the EST-contigs included COG4886 (LRR protein), pfam00931 (NB-ARC domain) and pfam01582 (TIR domain).

Mining with the term Hypersensitive Reaction (HR) yielded eight EST-contigs. The HR involves an extreme cellular response that may cause a high degree of resistance to disease. The HR is an induced defense response that results in the death of a limited number of host cells surrounding the sites of pathogen infection and culminates with the interruption of pathogen multiplication and growth on plant tissues. The HR occurs in response to the recognition of infection by the host, as a consequence of incompatibility between plant and pathogen (Pascholati and Leite, 1994). Of the eight EST-contigs identified, six encoded proteins that were related to the HR (Table S3).

The Pathogenesis project contained 61 EST-contigs that included proteins related to pathogenesis, chitinases, thaumatin family proteins and other proteins. The conserved domains found in the EST-contigs included cd00035 (chitin binding domain), cd00325 (chitinase glyco hydro 19 - glycoside hydrolase family 19 chitinase domain) and pfam00407 (pathogenesis-related protein Bet v I family) (Table S4).

45 EST-contigs were identified with the keyword Chitinase. Chitinases form a large group of enzymes with diverse structures and functions, some of which are related to the resistance of several plant species to pathogens (Sahai and Manocha, 1993; Jackson and Taylor, 1996). These enzymes are among the most expressed genes in coffee (Lin *et al.*, 2005). In a manner functionally similar to glucanases (see below), chitinases weaken the fungus cell wall by hydrolyzing chitin, an N-acetylglucosamine polymer. This hydrolysis results in cell lysis and death (Lin *et al.*, 2005). Analysis of the EST-contigs identified in the Chitinase project revealed the presence of chitinases with several conserved domains, including pfam00704 (glyco hydro 18), cd00035 (ChtBD1, chitin binding domain), cd00325 (chitinase glyco hydro 19, Glycoside hydrolase family 19 chitinase domain) and COG3469 (chitinase) (Table S5).

The CytochromeP450 project contained 144 EST-contigs (Table S6). Cytochrome P450 is involved in cellular oxidation pathways. The proteins encoded by these genes are involved in the biosynthesis of compounds related to defense, such as the *Arabidopsis* gene PAD3, which is required for the synthesis of camalexin during resistance to *Alternaria brassicicola* (Zhou *et al.*, 1999). Takemoto *et al.* (1999) showed that the product of the gene CYP82E1 (cytochrome P450) may be involved in the resistance of tobacco to *Pseudomonas syringae*. Qi *et al.* (2006) demonstrated that the enzyme AsCYP51H10 (a member of the cytochrome P450 family) participates in the production of antimicrobial compounds (avenacins) that confer resistance to diseases in oatmeal.

The keyword Glucanase, derived from the enzyme β1,3-glucanase, identified 88 EST-contigs. This enzyme hydrolyzes β1,3-glucan present in the fungus cell wall, and eventually leads to cell lysis and death (Selitrennikoff, 2001). The extracellular forms of this protein apparently act in the early stages of plant defense by exerting a direct fungicidal action on the hyphae of invading fungi. The release of eliciting oligosaccharides from the fungus wall can activate other local or systemic mechanisms of plant resistance. In contrast, the intracellular forms of β1,3-glucanase apparently act later in plant defense reactions (Boller and Métraux, 1988). The antifungal activity of glucanases has been demonstrated in several cellular enzymatic assays (Stintzi *et al.*, 1993), as well as in transgenic plants (Jach *et al.*, 1995). The EST-contigs identified in the Glucanase project included xyloglucan endotransglycosylase proteins, cellulase, β-1,3-glucanase, proteins from the glycosyl hydrolase family and proteins involved in brassinosteroid regulation (Table S7).

The project HSP (Heat Shock Protein) contained 30 EST-contigs. An increase in heat shock protein expression protects animals and plants against environmental stress. HSP90, a chaperone protein required for several metabolic defense pathways, contributes to the accumulation of resistance proteins in cells (Dreher and Callis, 2007). Analysis of the EST-contigs identified here revealed the presence of several classes of HSP proteins containing the conserved domains pfam00012 (HSP70), cd00298 (alpha-crystallin-Hsps), COG0071 (IbpA, molecular chaperone), cd00189 (TPR, tetratricopeptide repeat domain) and PRK00290 (dnaK, molecular chaperone DnaK) (Table S8).

The project Thaumatin identified 16 EST-contigs associated with thaumatin and pathogenesis-related (PR) proteins. Thaumatin is a PR protein with fungicidal activity against a large number of plant and human pathogens (Selitrennikoff, 2001). Although the precise mechanism of action of this protein is still unclear, various observations indicate that it has an important role in fungus death (Selitrennikoff, 2001). Thaumatin alters the permeability of the fungus cell wall but has little or no effect on the protoplast (Roberts and Selitrennikoff, 1990). This protein binds 1,3 β-glucan and has 1,3 β-glucanase activity *in vitro* (Grenier *et al.*, 1993; Trudel *et al.*, 1998). In tobacco, thaumatin interferes with the regulation of cell wall assembly in *Saccharomyces cerevisiae* (Yun *et al.*, 1997, 1998). Analysis of the EST-contigs identified in the Thaumatin project revealed the presence of the conserved domains Smart 00205 (THN) and pfam00314 (Thaumatin family) (Table S9).

The project Phytoalexin was consisted of three EST-contigs, all of which were annotated for the phytoalexin-deficient 4-1 protein of *Solanum tuberosum* (Table S10). Phytoalexins are synthesized by and accumulate in plant cells after microbial infection (Stoessl, 1986). Phytoalexins inhibit elongation of the germinative tube and radial colony growth, in addition to causing collapse of the membrane system and electrolyte loss. Resistant plants produce high levels of phytoalexins when compared to susceptible ones. The concentration of phytoalexin increases in parallel with that of key enzymes, *e.g.*, chalcone synthase, involved in its biosynthesis (Pascholati and Leite, 1994).

Five EST-contigs were identified with the keyword Chalconesynthase (Table S11). Three EST-contigs had the annotation chalcone synthase, one was annotated for aldo-keto reductase and one for 3-oxoacyl-(acyl-carrier-protein) synthase III. Chalcone synthase, an enzyme that regulates the biosynthesis of phenylpropanoids, catalyzes the first reaction of the biosynthetic pathway of flavonoids and isoflavonoids. Elicitors cause fast transient stimulation of chalcone synthase gene transcription as a basal event in the response to phytoalexins (Ryder *et al.*, 1984).

Four EST-contigs were identified with the keyword Polyphenoloxidase (Table S12). The most relevant annotations were catechol oxidase and polyphenol oxidase. Although the physiological function of polyphenol oxidase in plant cells is still unclear, most reports indicate that this enzyme functions in plant defense against attacks by pathogens and insects. This proposed involvement in defense is based on the enzyme's ability to oxidize phenolic compounds when the tissue is damaged. In this situation, rupture of the cellular compartment containing polyphenoloxidase (plastids) brings the enzyme into contact with phenolic compounds released by rupture of the vacuole (Melo *et al.*, 2006).

The project Importin consisted of 10 EST-contigs, six of which were annotated as importin α (Table S13). Importin α mediates the importation of cytosolic proteins into the nucleus. Plant importin α binds to virulence factors in *Agrobacterium tumefaciens* and several phytopathogenic viruses (Hermsmeier *et al.*, 2001).

The project Glucosyltransferase generated 125 EST-contigs. Glycosyl transferases catalyze the transfer of glucose residues to several substrates and regulate the activity of compounds that have important functions in plant defense against pathogens, such as salicylic acid (Chong *et al.*, 2002). The EST-contigs identified in the project Glucosyltransferase included proteins from the glucosyl transferase family, saccharose synthase and glucosyl transferases induced by cold (Table S14).

Figure 1 and Table S15 show the number of ESTs identified from the different libraries in each project. The libraries analyzed were NS1 (roots with nematodes), RM1 (leaves with coffee leaf miner and coffee leaf rust), RX1 (branches with *Xylella*) and SS1 (normal tissues). The libraries NS1, RM1 and RX1 were chosen because they were under biotic stress and therefore represented a potential source of ESTs involved in plant defense against stress. The SS1 library served as a control to identify genes that were expressed when the plant was not under stress.

This analysis showed that the RM1 library was best represented in the projects Resistance, Cytochrome P450, Glucanase and Glucosyltransferase. This library was also well represented in the projects NBS-LRR and Importin. These observations confirmed the potential importance of RM1 library ESTs in plant defense against coffee leaf miner and coffee leaf rust.

The RX1 library was also well represented in the projects Resistance, Cytochrome P450 and Glucosyltransferase. RX1 was best represented in the projects NBS-LRR, Pathogenesis, and Importin. This result indicates that ESTs mined in this library are important in coffee plant defense against *Xylella* because they are potentially related to resistance process against this pathogen.

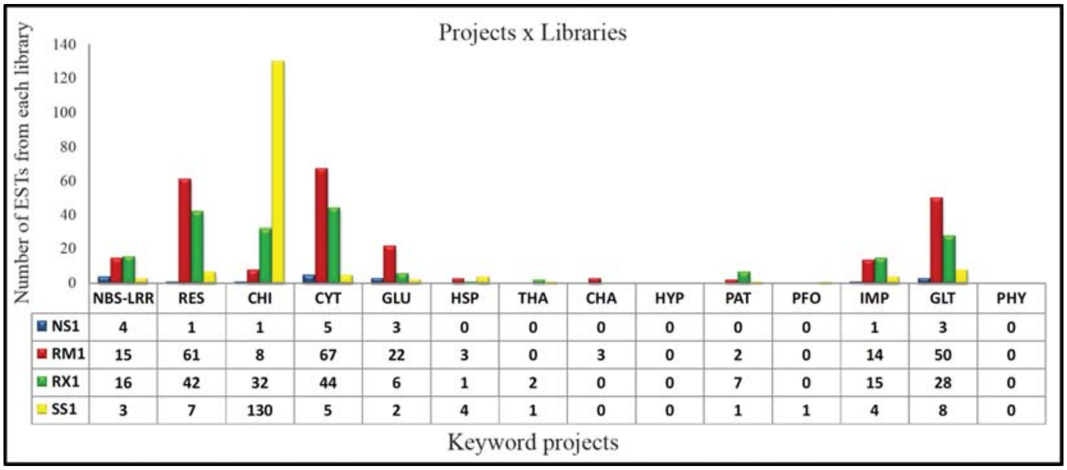
The projects Resistance, Chitinase, Cytochrome P450 and Glucosyltransferase contained the highest number of ESTs from the libraries of interest. The search term "resistance" is a good general term for plant defense mechanisms, and it was expected that the three libraries obtained under stress conditions (NS1, RM1, RX1) would be well-represented by ESTs for stress-related genes. This was indeed the case for the RM1 and RX1 libraries, whereas the NS1 library had only one EST in these projects while the SS1 library had seven. The low number of ESTs in the NS1 library may reflect a weak expression of genes related to the keyword "resistance", or may indicate that the genes were not sampled. Further studies may explain these results.

Chitinases, which are related to the resistance of several plant species to pathogens (Sahai and Manocha, 1993; Jackson and Taylor, 1996), were substantially expressed in normal conditions. This finding agrees with Lin *et al.* (2005), who showed that chitinases were among the most expressed genes in coffee. According to these authors, the fact that the chitinases are highly expressed and represented by an extended gene family in coffee may reflect a greater need for fungal resistance. This need may be related to the perennial nature of coffee and to the fact that this plant is a tropical species and may therefore be exposed to a large number of pathogens.

Plant cytochrome P450 enzymes are involved in a large number of biosynthetic reactions related to the formation of fatty acids, hormones, defense compounds and other molecules. Up to August 2009 (the last available update) 11,294 cytochrome P450 sequences, distributed in 977 families, had been identified; 3,284 of these sequences (excluding variants and pseudogenes) were from plant species (see Cytochrome P450 homepage in Internet Resoures). Hence, a large number of the cytochrome P450 ESTs identified here may be involved in biological processes unrelated to plant defense against diseases.

No ESTs were identified by the Hypersensitive and Phytoalexin projects in any of the libraries analyzed. This suggests that genes related to these terms are either not expressed in these libraries or that their expression was too low to be detected.

The 979 EST-contigs potentially associated with coffee plant resistance to diseases identified here provide a useful database for expanding our understanding of coffee plant resistance to pathogens and pests. The domains detected in the sequences described here served as: (1) recog-



	NBS-LRR	RES	CHI	CYT	GLU	HSP	THA	CHA	HYP	PAT	PFO	IMP	GLT	PHY
🟦 NS1	4	1	1	5	3	0	0	0	0	0	0	1	3	0
🟥 RM1	15	61	8	67	22	3	0	3	0	2	0	14	50	0
🟩 RX1	16	42	32	44	6	1	2	0	0	7	0	15	28	0
🟨 SS1	3	7	130	5	2	4	1	0	0	1	1	4	8	0

Figure 1 - Number of ESTs from NS1 (roots infected with nematodes), RM1 (leaves infected with leaf miner and coffee leaf rust), RX1 (stems infected with *Xylella* spp.) and SS1 (well-watered field plants) libraries that were present in the created projects. CHA – chalconesynthase, CHI – chitinase, CYT – cytochrome P450, GLT – glucosyltransferase, GLU – glucanase, HYP – hypersensitive, IMP – importin, NBS-LRR – nucleotide binding site-leucine rich repeat, PAT – pathogenesis, PFO – polyphenoloxidase, PHY – phytoalexin, RES – resistance and THA – thaumatin.

nition sites for protein-protein interactions, (2) effector binding sites to produce protein conformational changes, (3) sites for ion signal modulation that resulted in target regions alterations and consequent protein activation or inactivation, and (4) recognition sites for elicitor molecules released by pathogens. In addition, the EST-contigs identified here may be useful as probes in the physical mapping of coffee and as candidate genes for the development of molecular markers and the identification of RGAs (Resistance Genes Analogs) in coffee breeding programs.

Functional categorization

The NBS-LRR project generated 826 ESTs that, after clustering, resulted in 160 EST-contigs and 243 singlets. 140 of these EST-contigs had a score > 100 and an E-value < e^{-20}. Most of the 140 EST-contigs consisted of two or three ESTs (Figure S1). The size of the EST-contigs ranged from 469 bp (EST-contig 49: putative NBS-LRR type disease resistance protein of *Prunus persica*) to 3281 bp (EST-contig 50: putative leucine-rich repeat transmembrane, putative protein kinase of *Gossypium hirsutum*) (Table S1).

The best hits for species in the BlastX searches for EST-contigs were for *Vitis vinifera* and *Populus trichocarpa* (Figure S2). A recent study of the evolution and genomic composition of *C. canephora* revealed considerable conservation of the microcollinearity between this species and *V. vinifera* (Guyot *et al.*, 2009). These authors also reported a high level of conservation between the genomes of *C. canephora* and other Dicotyledon species such as *Solanum lycopersicon* and *Populus trichocarpa*. Species from the Asterid I group, such as *Solanum* sp. and *Lycopersicon* sp., known to be related to coffee (Lin *et al.*, 2005), also showed a large number of blast hits.

BlastX of the EST-contigs done through Blast2GO resulted in hits with annotations for NBS-LRR, CC-NBS-LRR, TIR-NBS-LRR, disease resistance protein, leucine-rich repeat transmembrane protein kinase, leucine-rich repeat receptor-like protein kinase and others. The average E-value for the 20 best BlastX hits ranged from e^{-10} to e^{-180} (Figure S3). The average similarity level for the 20 best hits ranged from 43% to 99% (Figure S4). These data show that the results obtained in the annotation were significant and thus highly reliable.

Of the 140 EST-contigs analyzed, 89 were associated with a conserved domain. Analysis of the InterProScan database revealed similarities with domains such as IPR002182 (PF00931 – domain NB-ARC), IPR001611 (PF00560 – domain LRR), IPR007271 (PTHR10231 – nucleotide-sugar transporter) and IPR000719 (PS50011 – domain protein chitinase), all of which have functions related to plant defense against pathogens. No proteins with enzymatic activity (no attached EC) and no metabolic maps (no attached KEGG map) were associated with these EST-contigs.

GO functional classification terms were retrieved for 138 EST-contigs (98.6%): one contig had 11 GO terms, 39 (27.9%) had only one term and two (1.42%) had no term (Figure S5). Since most of the biological functions for the DNA sequences and corresponding proteins were inferred by electronic annotation of the GO terms (Figure S6) it was necessary to screen several databases in order to obtain the greatest amount of information before depositing the sequence in a database.

The Cellular Component category contained terms such as "intrinsic to the membrane" and "plasma membrane" (Figure 2). Known R genes, such as *Cf9* from tomato, which confers resistance to *Cladosporium fulvum*, code for transmembrane proteins in which the extracellular N-terminal region consists of an LRR domain (Jones *et al.*, 1994).

The terms associated with the Molecular Function category included "signal transducer activity", "nucleotide binding" and others (Figure 3). Nucleotide binding activity has been associated with the NBS domain, which has the molecular function of binding ATP or GTP in various organisms (Saraste *et al.*, 1990; Traut, 1994). The N-terminal regions, together with the NBS domain, are believed to participate in the activation of signal transduction pathways involved in pathogen-specific resistance responses (Aarts *et al.*, 1998; Feys and Parker, 2000; Van Der Biezen *et al.*, 2000).

The Biological Process category contained terms such as "defense response", "apoptosis" and "tyrosine kinase protein transmembrane receptor signaling pathway" (Figure 4). The *Pto* gene from tomato encodes a seri-

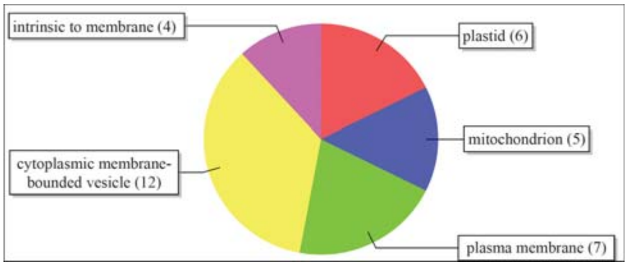


Figure 2 - Distribution of GO terms in the Cellular Component category, level 3.

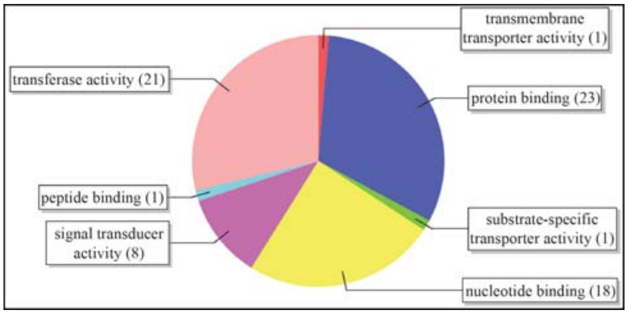


Figure 3 - Distribution of GO terms in the Molecular Function category, level 3.

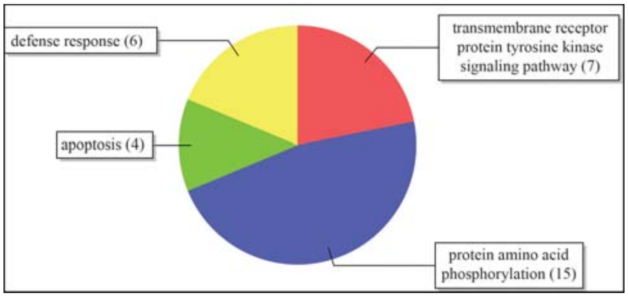


Figure 4 - Distribution of GO terms in the Biological Process category, level 3.

ne/threonine kinase protein that confers resistance to the races of *Pseudomonas syringae* that carry the avirulence gene *AvrPto* (Martin *et al.*, 1993). The gene *Xa21* (resistance to *Xanthomonas oryzae*) from rice, on the other hand, encodes proteins with the classic receptor-kinase format, *i.e.*, an extracellular LRR region, a transmembrane region and a cytosolic serine/threonine domain (Song *et al.*, 1995; Ellis and Jones, 1998).

Fisher's exact test

In Fisher's exact test, the p value represents the probability of a given cluster having equal expression in both of the groups analyzed, *i.e.*, if $p < 0.05$ then there is a significant difference in the expression of a given cluster between the two groups analyzed. Fisher's exact test identified seven contigs that differed in expression between the two groups analyzed, *i.e.*, biotic stress challenged libraries versus the remaining coffee genome libraries. A search for similarity using BlastX resulted in hits that included catalase, chitinase, a protein with a BURP domain and unknown proteins (Table 2).

In plant defense mechanisms, pathogen recognition is followed by an oxidative burst that triggers the rapid generation and accumulation of reactive oxygen species (ROS). During biotic and abiotic stress ROS may have two very different roles: the exacerbation of damage and the activation of defense responses (Dat *et al.*, 2000). Hydrogen peroxide, a powerful and potentially harmful ROS, damages both plant cells and pathogens (Allan and Fluhr, 1997). Catalase, an enzyme that protects cells from the toxic effects of peroxides, catalyzes the conversion of hydrogen peroxide to water and molecular oxygen (Willekens *et al.*, 1997). Catalase also uses hydrogen peroxide to oxidize toxins, including phenols, formic acid, formaldehyde and alcohols (Resende *et al.*, 2003). Dixon *et al.* (1994) showed that in soybean (*Glycine max*) catalase blocked the accumulation of phytoalexin glyceolin elicited by *Verticillium dahlia* Kleb. Treatment with catalase also blocks the induction of cell death in several systems in response to various avirulence signals (Lamb and Dixon, 1997). Catalase activity usually increases in response to high concentrations of hydrogen peroxide during virus infection (Riedle-Bauer, 2000). Lignin deposition tends to increase the resistance of plant cell walls to digestive enzymes of pathogens. In plant defense mechanisms, catalase acts in conjunction with peroxidases to accelerate the oxidation of phenolic compounds that are precursors in lignin synthesis (Margis-Pinheiro *et al.*, 1993). Based on these considerations, we inferred that contigs 8478 and 9073, which putatively encode catalases, are somehow involved in plant defense against pathogens.

Contig 14592 showed high similarity to a class III chitinase from *Coffea arabica* that contains a conserved domain called hevamine. This conserved domain hydrolyzes the linear polysaccharide chains of chitin and peptidoglycan and is therefore important for defense against pathogenic bacteria and fungi. Class III chitinases belong to family 18 of glycosyl hydrolases (Hamel e*t al.*, 1997) and include bifunctional enzymes with lysozyme activity (Jekel *et al.*, 1991). As mentioned earlier, chitinases degrade chitin and are widely distributed in many species of higher plants (Park *et al.*, 2002). Much of the evidence for the suggested roles of chitinases in plant defense has been based on dramatic and rapid enhancement of enzyme levels in the HR during induced host resistance, and in pathogen-infected tissues (Punja and Zhang, 1993).

The BlastX result for contigs 13986 and 431 was, in both cases, a member of the cysteine protease inhibitor

Table 2 - Contigs differentially expressed between libraries containing the responses to biotic stress challenges and other libraries from the Brazilian Coffee Genome Project. Differential expression was confirmed by Fisher's exact test, with the p values indicated in the last column.

Cluster	# reads	Length	BlastX	Score	E-value	GenBank Record	p value
Contig 10650	15	1376	hypothetical protein [*Populus trichocarpa*]	99.4	8.00E-19	ref|XP_002319603.1|	0.00079
Contig 13908	30	1181	BURP domain-containing protein [*Solanum tuberosum*]	289	1.00E-76	gb|ACD49738.1|	0.00114
Contig 13986	9	794	cysteine protease inhibitor family protein / cystatin family protein [*Arabidopsis thaliana*]	46.2	0.003	ref|NP_193383.1|	0.0007
Contig 14592	131	1290	class III chitinase [*Coffea arabica*]	447	1.00E-123	emb|CAJ43737.1|	0.00263
Contig 431	33	821	cysteine protease inhibitor family protein / cystatin family protein [*Arabidopsis thaliana*]	45.8	0.005	ref|NP_193383.1|	0.00158
Contig 8478	65	2154	catalase [*Prunus avium*]	788	0.00	gb|ABM47415.1|	0.00193
Contig 9073	347	1819	catalase [*Nicotiana plumbaginifolia*]	915	0.00	emb|CAA85426.1|	0.00482

family from *Arabidopsis thaliana*. However, the E-values of 0.003 and 0.005 for contigs 13986 and 431, respectively, were above the threshold usually adopted in similarity searches. This finding indicates that these contigs encode a *Coffea* protein with no significant similarity to any protein present in the NCBI nr database at the moment. The BlastX results for contig 10650 indicated that there was a protein in the database that had significant similarity to the product of this contig. However, the identity and function of this protein are unknown. Contig 13908 was highly similar to a BURP domain from *Solanum tuberosum*; BURP domains are believed to have important, fundamental functional roles (Hattori *et al.*, 1998), although no role in defense has yet been attributed to this domain.

In conclusion, the results of this study provide important information on resistance genes in the coffee genome since the terms identified in the GO analysis are probably related to plant defense mechanisms. The information generated by this genomic analysis also expands our understanding of the possible defense responses triggered in coffee plants in response to pathogens.

Acknowledgments

This work was supported by the Brazilian funding agencies CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) and FINEP (Financiadora de Estudos e Projetos), and by the CBP&D/Café (Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café).

References

Aarts N, Metz M, Holub E, Staskawicz BJ, Daniels MJ and Parker JE (1998) Different requirements for EDS1 and NDR1 by disease resistance genes define at least two R gene-mediated signaling pathways in *Arabidopsis*. Proc Natl Acad Sci USA 95:10306-10311.

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, *et al.* (1991) Complementary DNA sequencing: Expressed sequence tags and human genome project. Science 252:1651-1656.

Allan AC and Fluhr R (1997) Two distinct sources of elicited reactive oxygen species in tobacco epidermal cells. Plant Cell 9:1559-1572.

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

Baker B, Zambryski P, Staskawicz B and Dinesh-Kumar SP (1997) Signaling in plant-microbe interactions. Science 276:726-733.

Belkhadir Y, Subramaniam R and Dangl J (2004) Plant disease resistance protein signaling: NBS-LRR proteins and their partners. Curr Opin Plant Biol 7:391-399.

Bendahmane A, Kanyuka K and Baulcombe DC (1999) The Rx gene from potato controls separate virus resistance and cell death responses. Plant Cell 11:781-791.

Boller T, Métraux JP (1988) Extracellular localization of chitinases in cucumber. Physiol Mol Plant Pathol 33:11-16.

Botella MA, Parker JE, Frost LN, Bittner-Eddy PD, Beynon JL, Daniels MJ, Holub EB and Jones JDG (1998) Three genes of the *Arabidopsis* RPP1 complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. Plant Cell 10:1847-1860.

Cannon SB, Zhu H, Baumgarten AM, Spangler R, May G, Cook DR and Young ND (2002) Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. J Mol Evol 54:548-562.

Caplan J and Dinesh-Kumar SP (2006) Recognition and signal transduction associated with R gene-mediated resistance. In: Loebenstein G and Carr JP (eds) Natural Resistance Mechanisms of Plants to Viruses. Springer, Amsterdam, pp 73-98.

Carazzolle MF, Formighieri EF, Digiampietri LA, Araujo MRR, Costa GGL and Pereira GAG (2007) Gene projects: A genome Web tool for ongoing mining and annotation applied to CitEST. Genet Mol Biol 30:1030-1036.

Chong J, Balyz R, Schitt C, Beffa R, Fritig B and Saindrenan P (2002) Downregulation of a pathogen-responsive tobacco UDP-Glc: Phenylpropanoid glucosyltransferase reduces scopoletin glucoside accumulation, enhances oxidative stress, and weakens virus resistance. Plant Cell 14:1093-1107.

Conesa A, Götz S, García-Gomez JM, Terol J, Talón M and Robles M (2005) Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674-3676.

Dangl JL and Jones JD (2001) Plant pathogens and integrated defense responses to infection. Nature 411:826-833.

Dat J, Vandenabeele S, Vranová E, Montagu MV, Inzé D and Breusegem FV (2000) Dual action of the active oxygen species during plant stress responses. Cell Mol Life Sci 57:779-795.

Deslandes L, Olivier J, Peeters N, Feng DX, Khounlotham M, Boucher C, Somssich I, Genin S and Marco Y (2003) Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. Proc Natl Acad Sci USA 13:8024-8029.

Dinesh-Kumar SP, Tham W and Baker BJ (2000) Structure-function analysis of the tobacco mosaic virus resistance gene N. Proc Natl Acad Sci USA 97:14789-14794.

Dixon RA, Harrison MJ and Lamb CJ (1994) Early events in the activation of plant defense responses. Annu Rev Phytopathol 32:479-501.

Dreher K and Callis J (2007) Ubiquitin, hormones and biotic stress in plants. Ann Bot 99:787-822.

Ellis J and Jones D (1998) Structure and function of proteins controlling strain-specific pathogen resistance in plants. Curr Opin Plant Biol 1:288-293.

Ellis J, Dodds P and Pryor T (2000) Structure, function and evolution of plant disease resistance genes. Curr Opin Plant Biol 3:278-284.

Ellis JG, Lawrence GJ, Luck JE and Dodds PN (1999) Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. Plant Cell 11:495-506.

Fernandez D, Santos P, Agostini C, Bon MC, Petitot AS, Silva MC, Guerra-Guimarães L, Ribeiro A, Argout X and Nicole

M (2004) Coffee (*Coffea arabica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*). Mol Plant Pathol 5:527-536.

Feys BJ and Parker JE (2000) Interplay of signaling pathways in plant disease resistance. Trends Genet 16:449-455.

Fisher RA (1922) On the interpretation of χ^2 from contingency tables, and the calculation of P. J Royal Stat Soc 85:87-94.

Flor HH (1942) Inheritance of pathogenicity in *Melampsora lini*. Phytopathology 32:653-669.

Gai XW, Lal S, Xing LQ, Brendel V and Walbot V (2000) Gene discovery using the maize genome database ZmDB. Nucleic Acids Res 28:94-96.

Gassmann W, Hinsch ME and Staskawicz BJ (1999) The *Arabidopsis* RPS4 bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes. Plant J 20:265-277.

Grenier J, Potvin C and Asselin A (1993) Barley pathogenesis-related proteins with fungal cell wall lytic activity inhibit the growth of yeasts. Plant Physiol 103:1277-1283.

Guyot R, Mare M, Viader V, Hamon P, Coriton O, Bustamante-Porras J, Poncet V, Campa C, Hamon S and Kochko A (2009) Microcollinearity in an ethylene receptor coding gene region of the *Coffea canephora* genome is extensively conserved with *Vitis vinifera* and other distant dicotyledonous sequenced genomes. BMC Plant Biol 9:22-36.

Guzzo SD (2004) Aspectos Bioquímicos e Moleculares da Resistência Sistêmica Adquirida em Cafeeiro contra *Hemileia vastatrix*. Centro de Energia Nuclear na Agricultura, São Paulo, 236 pp.

Hamel F, Boivin R, Tremblay C and Bellemare G (1997) Structural and evolutionary relationships among chitinases of flowering plants. J Mol Evol 44:614-624.

Hattori J, Boutilier KA, van Lookeren Campagne MM and Miki BL (1998) A conserved BURP domain defines a novel group of plant proteins with unusual primary structures. Mol Gen Genet 259:424-428.

Hermsmeier D, Schittko U and Baldwin IT (2001) Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. I. Large-scale changes in the accumulation of growth- and defense-related plant mRNAs. Plant Physiol 125:683-700.

Höfte H, Desprez T, Amselem J, Chiapello H, Caboche M, Moisan A, Jourjon MF, Charpenteau JL, Berthomieu P, Guerrier D, *et al.* (1993) An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. Plant J 4:1051-1061.

Huang X and Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res 9:868-877.

Jach G, Gornhardt B, Mundy J, Logemann J, Pinsdorf E, Leah R, Schell J and Maas C (1995) Enhanced quantitative resistance against fungal disease by combinatorial expression of different barley antifungal proteins in transgenic tobacco. Plant J 8:97-109.

Jackson AO and Taylor CB (1996) Plant-microbe interactions: Life and death at the interface. Plant Cell 8:1651-1668.

Jekel PA, Hartmann BH and Beintema JJ (1991) The primary structure of hevamine, an enzyme with lysozyme/chitinase activity from *Hevea brasiliensis* latex. Eur J Biochem 200:123-130.

Jia Y, McAdams SA, Bryan GT, Hershey HP and Valent B (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. EMBO J 19:4004-4014.

Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ and Jones JDG (1994) Isolation of the tomato Cf-9 gene for resistance to *Cladosporiom fulvum* by transposon tagging. Science 266:789-793.

Lamb C and Dixon RA (1997) The oxidative burst in plant disease resistance. Annu Rev Plant Physiol Plant Mol Biol 48:251-275.

Lazo GR, Chao S, Hummel DD, Edwards H, Crossman CC, Lui N, Matthews DE, Carollo VL, Hane DL, You FM, *et al.* (2004) Development of an expressed sequence tag (EST) resource for wheat (*Triticum aestivum* L.): EST generation, unigene analysis, probe selection and bioinformatics for a 16,000-locus bin-delineated map. Genetics 168:585-593.

Lin C, Mueller LA, Carthy JM, Crouzillat D, Pétiard V and Tanksley SD (2005) Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. Theor Appl Genet 112:114-130.

Lupas A (1996) Coiled coils: New structures and new functions. Trends Biochem Sci 21:375-382.

Margis-Pinheiro M, Martin C, Didierjan L and Burkard G (1993) Differential expression of bean chitinase genes by virus infection, chemical treatment and UV irradiation. Plant Mol Biol 22:659-668.

Martin G, Brommonschenkel S, Chunwongse J, Frary A, Ganal M, Spivey R, Wu T, Earle E and Tanksley S (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262:1432-1436.

Melo GA, Shimizu MM and Mazzafera P (2006) Polyphenoloxidase activity in coffee leaves and its role in resistance against the coffee leaf miner and coffee leaf rust. Phytochemistry 67:277-285.

Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW and Young ND (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant J 20:317-332.

Noir S, Combes MC, Anthony F and Lashermes P (2001) Origin, diversity and evolution of NBS-type disease-resistance gene homologues in coffee trees (*Coffea* L.). Mol Genet Genomics 265:654-662.

Pan Q, Wendel J and Fluhr R (2000) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. J Mol Evol 50:203-213.

Park SM, Kim DH, Truong NH and Itoh Y (2002) Heterologous expression and characterization of class III chitinases from rice (*Oryza sativa* L.). Enzym Microb Tech 30:697-702.

Parniske M, Hammondkosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BBH and Jones JDG (1997) Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the Cf-4/9 locus of tomato. Cell 91:821-832.

Pascholati SF and Leite B (1994) Mecanismos bioquímicos de resistência às doenças. Revis Anu Patol Plantas 2:1-51.

Punja ZK and Zhang YY (1993) Plant chitinases and their roles in resistance to fungal diseases. J Nematol 25:526-540.

Qi X, Bakht S, Qin B, Leggett M, Hemmings A, Mellon F, Eagles J, Werck-Reichhart D, Schaller H, Lesot A, *et al.* (2006) A different function for a member of an ancient and highly conserved cytochrome P450 family: From essential sterols to plant defense. Proc Natl Acad Sci USA 103:18848-18853.

Resende MLV, Salgado SML and Chaves ZM (2003) Reactive oxygen species on plant defense responses to pathogens. Fitopatol Bras 28:123-130.

Richly E, Kurth J and Leister D (2002) Mode of amplification and reorganization of resistance genes during recent *Arabidopsis thaliana* evolution. Mol Biol Evol 19:76-84.

Richter TE and Ronald PC (2000) The evolution of disease resistance genes. Plant Mol Biol 42:195-204.

Riedle-Bauer M (2000) Role of reactive oxygen species and antioxidant enzymes in systemic virus infections of plants. J Phytopathol 148:297-302.

Roberts W and Selitrennikoff CP (1990) Zeamatin, an antifungal protein from maize with membrane-permeabilizing activity. J Genet Microbiol 136:1771-1778.

Ronning CM, Stegalkina SS, Ascenzi RA, Bougri O, Hart AL, Utterbach TR, Vanaken SE, Riedmuller SB, White JA, Cho J, *et al.* (2003) Comparative analyses of potato expressed sequence tag libraries. Plant Physiol 131:419-429.

Ryder TB, Cramer CL, Bell JN, Robbins MP, Dixon RA and Lamb CJ (1984) Elicitor rapidly induces chalcone synthase mRNA in *Phaseolus vulgaris* cells at the onset of the phytoalexin defense response. Proc Natl Acad Sci USA 81:5724-5728.

Sahai AS and Manocha MS (1993) Chitinases of plants and fungi: Their involvement in morphogenesis and host-parasite interaction. FEMS Microbiol Rev 11:317-338.

Selitrennikoff CP (2001) Antifungal proteins. Appl Environ Microbiol 67:2883-2894.

Shoemaker R, Keim P, Vodkin L, Retzel E, Clifton SW, Waterston R, Smoller D, Coryell V, Khanna A, Erpelding J, *et al.* (2002) A compilation of soybean ESTs: Generation and analysis. Genome 45:329-338.

Silva MC, Várzea V, Guerra-Guimarães L, Azinheira HG, Fernandez D, Petitot A, Bertrand B, Lashermes P and Nicole M (2006) Coffee resistance to the main diseases: Leaf rust and coffee berry disease. Braz J Plant Physiol 18:119-147.

Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai W, Zhu LH, *et al.* (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. Science 270:1804-1806.

Staskawicz BJ, Ausubel FM, Baker BJ, Ellis JG and Jones JD (1995) Molecular genetics of plant disease resistance. Science 268:661-667.

Stintzi A, Heitz T, Prasad V, Wiedemann-Merdinoglu S, Kauffmann S, Geoffroy P, Legrand M and Fritig B (1993) Plant 'pathogenesis-related' proteins and their role in defense against pathogens. Biochimie 75:687-706.

Stoessl A (1986) Secondary plant metabolites in plant disease resistance. Part II: Phytoalexins. Fitopatol Bras 11:25-53.

Tai TH, Dahlbeck D, Clark ET, Gajiwala P, Pasion R, Whalen MC, Stall RE and Staskawicz BJ (1999) Expression of the Bs2 pepper gene confers resistance to bacterial spot disease in tomato. Proc Natl Acad Sci USA 96:14153-14158.

Takemoto D, Hayashi M, Doke N, Nishimura M and Kawakita K (1999) Molecular cloning of a defense-response-related cytochrome p450 gene from tobacco. Plant Cell Physiol 40:1232-1242.

Thomas CM, Jones DA, Parniske M, Harrison K, Balint-Kurti PJ, Hatzixanthis K and Jones JDG (1997) Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. Plant Cell 9:2209-2224.

Trudel J, Grenier J, Potvin C and Asselin A (1998) Several thaumatin-like proteins bind to β -1,3-glucans. Plant Physiol 118:1431-1438.

Udall JA, Swanson JM, Haller K, Rapp RA, Sparks ME, Hatfield J, Yu Y, Wu Y, Dowd C, Arpat AB, *et al.* (2006) A global assembly of cotton ESTs. Genome Res 16:441-450.

Van Der Biezen EA, Sun J, Coleman MJ, Bibb MJ and Jones JDG (2000) *Arabidopsis* RelA/SpoT homologs implicate (p)ppGpp in plant signaling. Proc Natl Acad Sci USA 97:3747-3752.

Vieira LGE, Andrade AC, Colombo CA, Moraes AHA, Metha A, Oliveira AC, Labate CA, Marino CL, Monteiro-Vitorello CB, Monte DC, *et al.* (2006) Brazilian coffee genome project: An EST-based genomic resource. Braz J Plant Physiol 18:95-108.

Wang G, Ruan D, Song W, Sideris S, Chen L, Pi L, Zhang S, Zhang Z, Fauquet C, Gaut BS, *et al.* (1998) Xa21D encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. Plant Cell 10:765-779.

Willekens H, Chamnongpol S, Davey M, Schraudner M, Langebartels C, Montagu MV, Inzé D and Camp WV (1997) Catalase is a sink for H_2O_2 and is indispensable for stress defense in C3 plants. EMBO J 16:4806-4816.

Wolfsberg TG and Landsman DA (1997) Comparison of expressed sequence tags (ESTs) to human genomic sequences. Nucleic Acids Res 25:1626-1632.

Yamamoto K and Sasaki TL (1997) Large-scale EST sequencing in rice. Plant Mol Biol 35:135-144.

Young ND (2000) The genetic architecture of resistance. Curr Opin Plant Biol 3:285-290.

Yun D, Zhao Y, Pardo JM, Narasimhan ML, Damsz B, Lee H, Abad LR, D'urzo MP, Hasegawa PM and Bressan RA (1997) Stress proteins on the yeast cell surface determine resistance to osmotin, a plant antifungal protein. Proc Natl Acad Sci USA 94:7082-7087.

Yun DJ, Ibeas JI, Lee H, Coca MA, Narasimhan ML, Uesono Y, Hasegawa PM, Pardo JM and Bressan RA (1998) Osmotin, a plant antifungal protein, subverts signal transduction to enhance fungal cell susceptibility. Mol Cell 1:807-817.

Zhou N, Tootle TL and Glazebrook J (1999) *Arabidopsis* PAD3, a gene required for camalexin biosynthesis, encodes a putative cytochrome P450 monooxygenase. Plant Cell 11:2419-2428.

Internet Resources

Cenargen, https://alanine.cenargen.embrapa.br/cafEST/ (November 19, 2008).

Cytochrome P450 Homepage, http://drnelson.utmem.edu/CytochromeP450.html (August 26, 2009).

dbEST (ESTs database), http://www.ncbi.nlm.nih.gov/dbEST/index.html (March 4, 2009).

EBI (European Bioinformatics Institute), http://www.ebi.ac.uk/ (May 26, 2009).

InterPro, http://www.ebi. ac.uk/interpro/ (May 26, 2009).

InterProScan, http://www.ebi.ac.uk/Tools/InterProScan/ (May 26, 2009).

LGE (Laboratório de Genômica e Expressão), http://www.lge.ibi.unicamp.br/cafe/ (August 6, 2007).

NCBI (National Center for Biotechnology Information), http://www.ncbi.nlm.nih.gov/ (March 4, 2009).

Supplementary Material

The following online material is available for this article:

- Table S1 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project NBS-LRR, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S2 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Resistance, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S3 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Hypersensitive, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S4 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Pathogenesis, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S5 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Chitinase, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins

- Table S6 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Cytochrome P450, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins

- Table S7 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Glucanase, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S8 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project HSP (Heat Shock Protein), and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S9 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Thaumatin, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S10 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Phytoalexin, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S11 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Chalconesynthase, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S12 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Polyphenoloxydase, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S13 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Importin, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins

- Table S14 - EST-contigs with E-values < e^{-20} and scores > 100 obtained in the project Glucosyltransferase, and their blast hits, scores, E-values, sizes, number of reads and conserved domains from putative proteins.

- Table S15 - ESTs from the NS1, RM1, RX1 and SS1 libraries shown in Figure 1.

- Figure S1 - Distribution of the number of reads in the 140 EST-contigs with E-values < e^{-20} and scores > 100 formed after clustering.

- Figure S2 - Species with the most blast hits with the EST-contigs.

- Figure S3 - Average E-value distribution of the 20 best hits of the 140 EST-contigs submitted to BlastX by Blast2GO.

- Figure S4 - Average similarity value distribution of the 20 best hits of the 140 EST-contigs submitted to BlastX by Blast2GO.

- Figure S5 - Distribution of the number of GO terms (Cellular Component, Molecular Function and Biological Process) for the 140 EST-contigs analyzed by Blast2GO.

- Figure S6 - Evidence code (EC) distribution for the EST-contig blast hits.

This material is available as part of the online article from http://www.scielo.br/gmb.

Associate Editor: Sandro José de Souza

Table S1: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project NBS-LRR, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

NBS-LRR						
Contig	BLAST NR	Score	e-value	Length	Reads	Conserved Domains
1	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	168	1.00E-40	696	2	No CD has been identified
2	gi|55297132|dbj|BAD68775.1| putative NBS-LRR disease resistance protein [Oryza sativa (japonica cultivar-group)]	116	8.00E-25	801	2	pfam00931
3	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	143	3.00E-33	623	2	pfam00931
4	gi|5734781|gb|AAD50046.1| Very similar to disease resistance proteins [Arabidopsis thaliana]	201	2.00E-50	985	2	pfam00931
5	gi|37781280|gb|AAP44392.1| nematode resistance-like protein [Solanum tuberosum]	222	1.00E-56	831	2	pfam01582, pfam00931, COG4886
6	gi|5231014|gb|AAD41050.1| NBS/LRR disease resistance protein RPM1 [Arabidopsis thaliana]	164	4.00E-39	1121	2	pfam00931, COG4886
7	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	132	9.00E-30	756	2	pfam00931
9	gi|15219278|ref|NP_175742.1| disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	106	1.00E-21	953	2	pfam00931, pfam01576
10	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	157	2.00E-37	617	2	No CD has been identified
11	gi|47900744|gb|AAT39316.1| putative resistance complex protein I2C-2 [Solanum demissum]	124	3.00E-27	762	2	pfam00931
12	gi|47824998|gb|AAT38770.1| putative late blight resistance protein [Solanum demissum]	125	1.00E-27	858	2	pfam00931
13	gi|50941633|ref|XP_480344.1| putative RGH1A [Oryza sativa (japonica cultivar-group)]	113	5.00E-24	782	2	pfam00931, COG4886
14	gi|21436417|gb|AAM51409.1| unknown protein [Arabidopsis thaliana]	359	6.00E-98	844	3	pfam08263, cd00116
15	gi|50428772|gb|AAT77096.1| Fom-2 protein [Cucumis melo]	166	1.00E-39	1033	3	No CD has been identified
16	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	139	2.00E-49	1241	2	pfam00931
17	gi|61105091|gb|AAX38325.1| receptor-like protein kinase [Lycopersicon esculentum]	270	7.00E-71	1118	3	cd00116
18	gi|27311539|gb|AAO00735.1| expressed protein [Arabidopsis thaliana]	194	3.00E-48	874	2	cd00192
20	gi|62733933|gb|AAX96042.1| Leucine Rich Repeat, putative [Oryza sativa]	171	2.00E-41	711	2	cd00116
22	gi|13249030|gb|AAK16647.1| F-box containing protein TIR1 [Populus tremula x Populus tremuloides]	431	0	1065	4	cd00116
24	gi|50428772|gb|AAT77096.1| Fom-2 protein [Cucumis melo]	252	2.00E-65	1852	3	No CD has been identified
25	gi|18181933|dbj|BAB83871.1| disease resistance protein [Arabidopsis thaliana]	172	2.00E-45	867	4	pfam00931
26	gi|52075880|dbj|BAD45826.1| putative CMP-sialic acid transporter [Oryza sativa]	179	1.00E-88	654	2	pfam04142
27	gi|9758146|dbj|BAB08703.1| disease resistance protein [Arabidopsis thaliana]	110	7.00E-23	912	4	pfam00931, COG4886
28	gi|50939001|ref|XP_479028.1| putative disease resistance protein RPH8 [Oryza sativa (japonica cultivar-group)]	227	2.00E-77	1443	3	pfam00931, COG4886
29	gi|48057658|gb|AAT39957.1| putative late blight resistance protein [Solanum demissum]	206	2.00E-79	1266	4	pfam00931
30	gi|21554189|gb|AAM63268.1| putative leucine-rich repeat disease resistance protein [Arabidopsis thaliana]	432	0	1335	7	cd00116
31	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	168	1.00E-40	803	3	No CD has been identified
32	gi|37718761|gb|AAR01633.1| putative leucine-rich repeat receptor kinase [Oryza sativa]	312	0	1833	4	cd00192, smart00220
33	gi|6456755|gb|AAF09256.1| disease resistance protein BS2 [Capsicum chacoense]	358	4.00E-97	2540	15	pfam00931
35	gi|50939001|ref|XP_479028.1| putative disease resistance protein RPH8 [Oryza sativa (japonica cultivar-group)]	325	4.00E-87	1857	7	pfam00931, COG4886
36	gi|21805750|gb|AAM76770.1| hypothetical protein [Arabidopsis thaliana]	176	9.00E-43	965	5	pfam07320
37	gi|29839503|sp|P59584|RP8HA_ARATH Disease resistance protein RPH8A (RPP8 homolog A) [Arabidopsis thaliana]	293	2.00E-77	2164	5	pfam00931
38	gi|15240263|ref|NP_200956.1| leucine-rich repeat transmembrane protein kinase, putative [Arabidopsis thaliana]	523	0	1416	3	cd00180, cd00116, smart00220, COG4886
39	gi|7270910|emb|CAB80590.1| receptor protein kinase-like protein [Arabidopsis thaliana]	184	5.00E-45	1239	2	cd00192, cd00116, smart00220
40	gi|39636723|gb|AAR29070.1| blight resistance protein RGA1 [Solanum bulbocastanum]	248	4.00E-64	1540	2	pfam00931, COG4886
41	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	145	2.00E-33	785	2	No CD has been identified
42	gi|26450926|dbj|BAC42570.1| putative receptor protein kinase [Arabidopsis thaliana]	286	9.00E-76	1307	3	pfam08263, cd00192, smart00220, COG4886
44	gi|50939001|ref|XP_479028.1| putative disease resistance protein RPH8 [Oryza sativa (japonica cultivar-group)]	291	4.00E-77	1981	6	pfam00931, COG4886
45	gi|50945887|ref|XP_482471.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]	137	2.00E-34	1313	4	pfam00931, COG4886
46	gi|4835246|emb|CAB42924.1| putative disease resistance protein [Arabidopsis thaliana]	216	1.00E-54	1500	2	pfam00931, PRK11281
47	gi|24459865|emb|CAC82608.1| disease resistance-like protein [Coffea canephora]	157	6.00E-37	1025	3	pfam00931
48	gi|30689664|ref|NP_195056.2| disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	652	0	2255	13	pfam00931
49	gi|47059739|gb|AAT09451.1| putative NBS-LRR type disease resistance protein [Prunus persica]	461	0	3281	15	pfam00931, COG4886
50	gi|49333393|gb|AAT64032.1| putative leucine-rich repeat transmembrane protein; putative protein kinase [Gossypium hirsutum]	227	9.00E-59	469	3	pfam08263, cd00192, smart00220, COG4886
51	gi|9651941|gb|AAF91322.1| receptor-like protein kinase 1 [Glycine max]	531	0	831	3	cd00180, cd00116, smart00220, pfam08263
52	gi|20466151|gb|AAM20393.1| transport inhibitor response 1, putative [Arabidopsis thaliana]	222	7.00E-57	641	2	No CD has been identified
53	gi|39636816|gb|AAR29076.1| blight resistance protein T118 [Solanum tarijense]	126	5.00E-30	768	4	pfam00931, COG4886
54	gi|50945887|ref|XP_482471.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]	246	1.00E-63	1619	5	pfam00931, COG4886
55	gi|7573427|emb|CAB87743.1| transport inhibitor response 1 (TIR1) [Arabidopsis thaliana]	280	3.00E-74	812	5	No CD has been identified
56	gi|14573459|gb|AAK68074.1| somatic embryogenesis receptor-like kinase 3 [Arabidopsis thaliana]	210	2.00E-53	1118	3	cd00180, COG4886, smart00220, pfam08263
57	gi|9279736|dbj|BAB01326.1| receptor-like kinase [Arabidopsis thaliana]	382	0	816	2	cd00192, cd00116, pfam08263
58	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	194	6.00E-48	1246	2	pfam00931
59	gi|25403295|pir||B86398 protein T7N9.24 [imported] - [Arabidopsis thaliana]	132	3.00E-60	1059	2	smart00255, pfam00931, COG4886, cd00116
60	gi|34909106|ref|NP_915900.1| putative NBS-LRR type resistance protein [Oryza sativa]	100	3.00E-25	900	2	pfam00931, COG4886
61	gi|17979045|gb|AAL49790.1| unknown protein [Arabidopsis thaliana]	291	1.00E-77	859	2	cd00180, cd00116, smart00220, pfam08263

63	gi|44717048|gb|AAN62015.2| leucine-rich repeat protein [Capsicum annum]	262	9.00E-69	821	4	pfam08263
66	gi|38566726|emb|CAE76632.1| leucine rich repeat protein [Cicer arietinum]	500	0	1315	3	cd00116, pfam08263
67	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	194	1.00E-56	1239	3	No CD has been identified
68	gi|3360289|gb|AAC27894.1| leucine-rich repeat transmembrane protein kinase 1 [Zea mays]	407	0	1182	3	cd00180, cd00116, smart00220, pfam08263
69	gi|46095229|gb|AAS80152.1| FOM-2 [Cucumis melo]	177	3.00E-43	1022	3	No CD has been identified
70	gi|40218005|gb|AAR82926.1| coronatine-insensitive 1 [Lycopersicon esculentum]	469	0	1442	3	No CD has been identified
71	gi|38564276|gb|AAR23717.1| At4g22730 [Arabidopsis thaliana]	343	6.00E-93	1078	3	cd00192, pfam08263, smart00220, COG4886
72	gi|19699206|gb|AAL90969.1| At1g21410/F24J8_17 [Arabidopsis thaliana]	502	0	1333	3	smart00367, cd00116
73	gi|8515762|gb|AAF76163.1| RGC1 [Solanum tuberosum]	290	6.00E-77	1239	5	pfam00931
74	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	149	7.00E-35	691	2	pfam00931
75	gi|27311539|gb|AAO00735.1| expressed protein [Arabidopsis thaliana]	468	0	1235	5	cd00192
76	gi|8547232|gb|AAF76308.1| Prf [Lycopersicon pimpinellifolium]	196	9.00E-62	744	2	pfam00931
77	gi|50428772|gb|AAT77096.1| Fom-2 protein [Cucumis melo]	166	1.00E-39	1384	2	No CD has been identified
78	gi|20259553|gb|AAM14119.1| putative receptor protein kinase [Arabidopsis thaliana]	429	0	895	2	cd00180, COG4886, smart00220, cd00116
79	gi|37718761|gb|AAR01633.1| putative leucine-rich repeat receptor kinase [Oryza sativa]	217	3.00E-60	955	2	cd00192, smart00220
80	gi|28393326|gb|AAO42089.1| putative receptor protein kinase [Arabidopsis thaliana]	297	7.00E-90	1021	5	cd00180,cd00116, smart00220, pfam08263
81	gi|42566272|ref|NP_192248.2| leucine-rich repeat transmembrane protein, putative [Arabidopsis thaliana]	194	9.00E-49	518	2	cd00192, smart00220, COG4886
82	gi|5541686|emb|CAB51192.1| putative protein [Arabidopsis thaliana]	254	6.00E-84	2331	8	pfam00931
83	gi|50939001|ref|XP_479028.1| putative disease resistance protein RPH8 [Oryza sativa (japonica cultivar-group)]	390	0	2376	4	pfam00931, COG4886
84	gi|42566272|ref|NP_192248.2| leucine-rich repeat transmembrane protein RPH8A [Oryza sativa]	167	2.00E-40	702	2	cd00192, smart00220, COG4886
86	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	357	7.00E-97	1571	7	pfam00931
87	gi|5231014|gb|AAD41050.1| NBS/LRR disease resistance protein RPM1 [Arabidopsis thaliana]	114	3.00E-24	694	2	pfam00931, COG4886
88	gi|50428772|gb|AAT77096.1| Fom-2 protein [Cucumis melo]	105	1.00E-21	672	2	No CD has been identified
89	gi|21239384|gb|AAM44275.1| receptor-like kinase RHG4 [Glycine max]	375	0	1395	2	cd00180, smart00220, cd00116
90	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	128	1.00E-28	559	2	pfam00931
91	gi|42572433|ref|NP_974312.1| leucine-rich repeat transmembrane protein kinase, putative [Arabidopsis thaliana]	257	5.00E-67	1325	5	cd00192, smart00220, COG4886
92	gi|15218941|ref|NP_176789.1| leucine-rich repeat protein kinase, putative (TMK1) [Arabidopsis thaliana]	556	0	1049	4	cd00180, cd00116, smart00220, COG4886
93	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	124	4.00E-27	802	2	pfam00931
94	gi|50945887|ref|XP_482471.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]	207	8.00E-52	1693	3	pfam00931, COG4886
95	gi|24459841|emb|CAC82597.1| disease resistance-like protein [Coffea arabica]	260	2.00E-68	778	2	pfam00931
96	gi|32364526|gb|AAP80292.1| resistance protein Tsu5 [Arabidopsis thaliana]	105	1.00E-21	931	2	pfam00931
97	gi|24459845|emb|CAC82610.1| disease resistance-like protein [Coffea arabica]	161	2.00E-58	1047	2	pfam00931
99	gi|7672732|gb|AAF66615.1| LRR receptor-like protein kinase [Nicotiana tabacum]	364	2.00E-99	766	5	cd00180, smart00220
100	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	208	2.00E-52	953	3	pfam00931
101	gi|4689223|gb|AAD27815.1| disease resistance protein I2 [Lycopersicon esculentum]	130	4.00E-29	715	2	pfam00931
102	gi|55771369|dbj|BAD72536.1| LIM domain containing protein-like [Oryza sativa]	438	0	885	3	smart00132
103	gi|21689699|gb|AAM67471.1| putative F-box family protein AtFBL3 [Arabidopsis thaliana]	321	2.00E-86	783	5	No CD has been identified
104	gi|47059739|gb|AAT09451.1| putative NBS-LRR type disease resistance protein [Prunus persica]	200	7.00E-50	1107	4	pfam00931, COG4886
105	gi|42569070|ref|NP_179220.2| leucine-rich repeat transmembrane protein kinase, putative [Arabidopsis thaliana]	183	7.00E-45	1102	3	cd00192, smart00220, COG4886
106	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	148	2.00E-34	729	2	pfam00931
107	gi|5734781|gb|AAD50046.1| Very similar to disease resistance proteins [Arabidopsis thaliana]	142	5.00E-33	655	2	pfam00931
108	gi|16323486|gb|AAL15237.1| putative protein translocation complex Sec61 gamma chain [Arabidopsis thaliana]	102	5.00E-21	548	2	COG2443
109	gi|24459853|emb|CAC82602.1| disease resistance-like protein [Coffea arabica]	300	7.00E-80	1131	5	pfam00931
113	gi|30698151|ref|NP_201372.2| leucine-rich repeat transmembrane protein kinase, putative [Arabidopsis thaliana]	232	3.00E-62	860	3	cd00192, smart00220, COG4886
114	gi|46095229|gb|AAS80152.1| FOM-2 [Cucumis melo]	104	8.00E-21	1558	3	No CD has been identified
115	gi|11994124|dbj|BAB01126.1| receptor protein kinase [Arabidopsis thaliana]	274	3.00E-72	848	2	cd00192, cd00116, smart00220, pfam08263
116	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	216	1.00E-54	1229	4	No CD has been identified
117	gi|50899184|ref|XP_450380.1| putative disease related protein 2 [Oryza sativa]	138	1.00E-65	1886	4	pfam00931, COG4886
119	gi|13249030|gb|AAK16647.1| F-box containing protein TIR1 [Populus tremula x Populus tremuloides]	200	2.00E-70	771	3	cd00116
120	gi|62733050|gb|AAX95167.1| receptor-like protein kinase [Oryza sativa]	329	5.00E-89	735	3	cd00192, smart00220, COG4886
121	gi|53828509|gb|AAU94364.1| At1g27190 [Arabidopsis thaliana]	256	5.00E-72	712	2	cd00192, cd05086
122	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	139	1.00E-31	934	3	No CD has been identified
123	gi|50945889|ref|XP_482472.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]	145	9.00E-34	779	4	pfam00931, COG4886
124	gi|9758140|dbj|BAB08632.1| disease resistance protein-like [Arabidopsis thaliana]	169	1.00E-40	936	2	pfam00931, COG4886
125	gi|5478530|gb|AAD43920.1| UVB-resistance protein UVR8 [Arabidopsis thaliana]	453	0	802	4	pfam00415, COG5184
126	gi|48057658|gb|AAT39957.1| putative late blight resistance protein [Solanum demissum]	164	2.00E-39	786	2	pfam00931
127	gi|30385250|gb|AAP23944.1| leucine-rich repeat protein [x Citrofortunella mitis]	360	7.00E-98	1380	20	cd00116, pfam08263
128	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	105	3.00E-21	1325	3	pfam00931
129	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	205	2.00E-51	1058	2	pfam00931
131	gi|34913790|ref|NP_918242.1| putative LRR [Oryza sativa (japonica cultivar-group)]	143	1.00E-34	799	2	pfam00931, COG4886
132	gi|50905715|ref|XP_464346.1| putative protein kinase 2 [Oryza sativa]	184	6.00E-48	824	2	cd00180, smart00220
133	gi|50918423|ref|XP_469608.1| hypothetical protein [Oryza sativa (japonica cultivar-group)]	334	2.00E-90	1019	5	cd02970
134	gi|50925601|ref|XP_472985.1| OSJNBa0084K20.12 [Oryza sativa (japonica cultivar-group)]	236	6.00E-61	730	2	pfam04142

135	gi|55297132|dbj|BAD68775.1| putative NBS-LRR disease resistance protein [Oryza sativa (japonica cultivar-group)]	124	3.00E-27	814	3	pfam00931
136	gi|53689783|gb|AAU89792.1| late blight resistance protein-like [Solanum tuberosum]	152	8.00E-36	803	2	pfam00931, smart00534
137	gi|24459843|emb|CAC82598.1| disease resistance-like protein [Coffea arabica]	138	4.00E-31	1237	3	pfam00931
138	gi|39636816|gb|AAR29076.1| blight resistance protein T118 [Solanum tarijense]	172	2.00E-41	1124	2	pfam00931, COG4886
139	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	332	1.00E-89	1366	2	pfam00931
140	gi|62734129|gb|AAX96238.1| NB-ARC domain, putative [Oryza sativa (japonica cultivar-group)]	111	4.00E-23	1259	3	pfam00931
141	gi|15222893|ref|NP_175437.1| disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	108	1.00E-22	711	2	pfam00931
142	gi|38603954|gb|AAR24722.1| At5g49400 [Arabidopsis thaliana]	197	3.00E-49	709	2	No CD has been identified
144	gi|7270269|emb|CAB80038.1| putative protein [Arabidopsis thaliana]	470	0	990	2	smart00256
145	gi|15222893|ref|NP_175437.1| disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	338	3.00E-91	1906	5	pfam00931
146	gi|50941263|ref|XP_480159.1| putative U3 snoRNP protein IMP4 [Oryza sativa]	388	0	744	4	pfam04427
147	gi|6456755|gb|AAF09256.1| disease resistance protein BS2 [Capsicum chacoense]	216	6.00E-98	2019	6	pfam00931
148	gi|34896972|ref|NP_909832.1| putative leucine-rich repeat protein [Oryza sativa]	303	6.00E-81	1102	36	cd00116, pfam08263
149	gi|15223744|ref|NP_172891.1| leucine-rich repeat transmembrane protein kinase, putative [Arabidopsis thaliana]	339	8.00E-92	1150	2	cd00192, smart00220, COG4886
150	gi|22652532|gb|AAN03742.1| NBS-LRR-like protein [Oryza sativa (japonica cultivar-group)]	222	4.00E-56	2251	6	pfam00931
151	gi|20259113|gb|AAM14272.1| unknown protein [Arabidopsis thaliana]	643	0	1958	18	No CD has been identified
152	gi|5080812|gb|AAD39321.1| Putative disease resistance protein [Arabidopsis thaliana]	181	5.00E-44	1580	3	pfam00931
154	gi|62321425|dbj|BAD94804.1| putative protein [Arabidopsis thaliana]	170	4.00E-41	870	3	No CD has been identified
155	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	227	5.00E-58	1120	2	pfam00931
156	gi|50945887|ref|XP_482471.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]	145	1.00E-51	1558	4	pfam00931, COG4886
157	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	241	7.00E-62	1745	6	pfam00931
159	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	146	5.00E-34	639	2	pfam00931

Table S2: EST-Contigs with e-value $< e^{-20}$ and score > 100 obtained in the Project Resistance, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Resistance						
Contig	**BLAST NR**	**Score**	**e-value**	**Length**	**Reads**	**Conserved Domains**
2	gi|118489526|gb|ABK96565.1|[118489526] unknown [Populus trichocarpa x Populus deltoides]	159	7.00E-38	454	2	No CD has been identified
4	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	168	1.00E-40	696	2	No CD has been identified
5	gi|55297132|dbj|BAD68775.1| putative NBS-LRR disease resistance protein [Oryza sativa (japonica cultivar-group)]	116	1.00E-24	839	3	pfam00931
6	gi|82581291|dbj|BAE48710.1| NRAMP family metal transporter [Chengiopanax sciadophylloides]	404	0	1922	6	PRK00701
7	gi|82581291|dbj|BAE48710.1| NRAMP family metal transporter [Chengiopanax sciadophylloides]	480	0	1170	14	PRK00701
8	gi|15228832|ref|NP_191829.1| ATMRP10 [Arabidopsis thaliana]	343	3.00E-93	848	4	cd03244, cd03250, pfam00664, cd03271, cd03233, COG1132
9	gi|15220789|ref|NP_175748.1| kinase [Arabidopsis thaliana]	635	0	1342	3	cd00192, COG4886
11	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	143	4.00E-33	623	2	pfam00931
12	gi|40846374|gb|AAR92477.1| putative WRKY transcription factor 30 [Vitis aestivalis]	198	1.00E-49	767	2	pfam03106
13	gi|157358243|emb|CAO65880.1|[157358243] unnamed protein product [Vitis vinifera]	502	0	1242	5	No CD has been identified
14	gi|5734781|gb|AAD50046.1| Very similar to disease resistance proteins [Arabidopsis thaliana]	201	3.00E-50	985	2	pfam00931
15	gi|79316282|ref|NP_001030932.1| ATP binding [Arabidopsis thaliana]	293	5.00E-78	809	6	cd00550, pfam02374
16	gi|15220948|ref|NP_172851.1| GN (GNOM) [Arabidopsis thaliana]	309	0	1306	5	cd00171, COG5307
17	gi|41052472|dbj|BAD07483.1| PDR-type ABC transporter 1 [Nicotiana tabacum]	169	4.00E-65	1098	5	cd03232, cd03233, cd03234, pfam01061, pfam08370, cd03226, pfam00005
18	gi|67933183|ref|ZP_00526308.1| regulatory protein, ArsR [Solibacter usitatus Ellin6076]	125	2.00E-27	682	2	cd00090
19	gi|77382823|gb|ABA74336.1| Peptidase C39, bacteriocin processing [Pseudomonas fluorescens PfO-1]	277	3.00E-73	827	18	cd02423
20	gi|34104782|gb|AAQ61139.1| probable manganese transport transmembrane protein [Chromobacterium violaceum ATCC 12472]	346	5.00E-94	798	7	COG1914
21	gi|37781280|gb|AAP44392.1| nematode resistance-like protein [Solanum tuberosum]	222	1.00E-56	831	2	pfam01582, pfam00931, COG4886
22	gi|5231014|gb|AAD41050.1| NBS/LRR disease resistance protein RPM1 [Arabidopsis thaliana]	164	6.00E-39	1121	2	pfam00931, COG4886
23	gi|50928783|ref|XP_473919.1| OSJNBa0058K23.17 [Oryza sativa (japonica cultivar-group)]	395	0	1327	10	COG1132, pfam00664, cd03250, cd03244, cd03231
24	gi|14715462|dbj|BAB62040.1| CjMDR1 [Coptis japonica]	265	1.00E-69	769	2	cd03249, pfam00664, COG1132
25	gi|34898274|ref|NP_910483.1| similar to NBS-LRR type resistance gene [Oryza sativa (japonica cultivar-group)]	84	9.00E-24	946	2	pfam00931, COG4886
26	gi|15237456|ref|NP_199466.1| ATPase, coupled to transmembrane movement of substances [Arabidopsis thaliana]	432	0	994	3	cd03249, pfam00664, COG1132, cd03270
29	gi|72117527|gb|AAZ59790.1| ABC transporter related [Ralstonia eutropha JMP134]	157	1.00E-37	283	2	cd03230, COG1131
30	gi|15217776|ref|NP_174115.1| ATPase, coupled to transmembrane movement of substances [Arabidopsis thaliana]	361	0	1034	2	cd03249, pfam00664, COG1132, cd03270
31	gi|15230349|ref|NP_188563.1| protein binding [Arabidopsis thaliana]	236	8.00E-61	740	2	cd00116
32	gi|42563438|ref|NP_186900.3| Ran GTPase binding / chromatin binding [Arabidopsis thaliana]	350	0	1542	4	pfam00415, COG5184
33	gi|9623337|gb|AAF90123.1| Rar1 [Hordeum vulgare]	162	1.00E-38	761	5	pfam04968
35	gi|68557365|ref|ZP_00596706.1| Drug resistance transporter EmrB/QacA subfamily [Ralstonia metallidurans CH34]	279	7.00E-75	777	6	PRK10504, pfam07690
36	gi|16418843|gb|AAL19297.1| putative diguanylate cyclase/phosphodiesterase domain 1 [Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67]	185	1.00E-45	741	4	cd01948, COG4943
37	gi|30678939|ref|NP_187029.2| ankyrin repeat family protein / regulator of chromosome condensation (RCC1) family protein [Arabidopsis thaliana]	328	3.00E-88	1369	6	cd00204, pfam00415, COG5184
38	gi|15219278|ref|NP_175742.1| disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	106	1.00E-21	953	2	pfam00931, pfam01576
39	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	157	3.00E-37	617	2	No CD has been identified
40	gi|50904597|ref|XP_463787.1| putative senescence-associated protein [Oryza sativa (japonica cultivar-group)]	138	2.00E-31	737	6	cd00158
41	gi|47900744|gb|AAT39316.1| putative resistance complex protein I2C-2, putative [Solanum demissum]	124	3.00E-27	762	2	pfam00931
43	gi|15235780|ref|NP_194004.1| leucine-rich repeat transmembrane protein kinase, putative [Arabidopsis thaliana]	264	3.00E-69	762	3	cd00192, COG4886, pfam08263, smart00220
44	gi|13377502|gb|AAK20740.1| LRK33 [Triticum aestivum]	269	1.00E-70	897	2	cd00192, smart00220
45	gi|77554673|gb|ABA97469.1| expressed protein [Oryza sativa (japonica cultivar-group)]	355	1.00E-96	701	2	pfam00415, COG5184
46	gi|47824998|gb|AAT38770.1| putative late blight resistance protein [Solanum demissum]	125	2.00E-27	858	2	pfam00931
47	gi|68213281|ref|ZP_00565114.1| ABC-2 [Methylobacillus flagellatus KT]	204	2.00E-51	662	3	COG0842
48	gi|77383604|gb|ABA75117.1| Two component Heavy metal response transcriptional regulator, winged helix family [Pseudomonas fluorescens PfO-1]	343	3.00E-93	817	11	cd00156, cd00383, PRK09836
49	gi|50941633|ref|XP_480344.1| putative RGH1A [Oryza sativa (japonica cultivar-group)]	113	6.00E-24	782	2	pfam00931
50	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	121	5.00E-26	1079	4	pfam00931
51	gi|15232373|ref|NP_188718.1| leucine-rich repeat family protein [Arabidopsis thaliana]	359	8.00E-98	844	3	pfam08263, cd00116
52	gi|50428772|gb|AAT77096.1| Fom-2 protein [Cucumis melo]	166	1.00E-39	1033	3	No CD has been identified
53	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	139	2.00E-49	1241	2	pfam00931
54	gi|21326114|gb|AAM47580.1| putative ABC-transporter-like protein [Sorghum bicolor]	333	2.00E-91	806	2	cd03249, pfam00664, COG1132
56	gi|18394385|ref|NP_564003.1| ATP binding / kinase/ protein kinase/ protein [Arabidopsis thaliana]	368	0	1222	5	cd00192
58	gi|55056944|emb|CAH39854.1| PDR-like ABC transporter [Nicotiana tabacum]	331	3.00E-99	774	3	pfam01061

<table>
<tr><td>59</td><td>gi|62733933|gb|AAX96042.1| Leucine Rich Repeat, putative [Oryza sativa]</td><td>171</td><td>2.00E-41</td><td>711</td><td>2</td><td>cd00116</td></tr>
<tr><td>60</td><td>gi|79318957|ref|NP_001031116.1| ATMRP1 (Arabidopsis thaliana multidrug resistance-associated protein 1); xenobiotic-transporting ATPase. [Arabidopsis thaliana]</td><td>100</td><td>6.00E-20</td><td>783</td><td>2</td><td>cd03244, cd03250, pfam00664, cd03297, COG1132</td></tr>
<tr><td>62</td><td>gi|15227133|ref|NP_182301.1| ATMRP4 [Arabidopsis thaliana]</td><td>429</td><td>0</td><td>1079</td><td>5</td><td>cd03244, cd03250, pfam00664, cd03271, cd03233, COG1132</td></tr>
<tr><td>65</td><td>gi|77383868|gb|ABA75381.1| Secretion protein HlyD [Pseudomonas fluorescens PfO-1]</td><td>425</td><td>0</td><td>744</td><td>6</td><td>PRK10476, pfam00529</td></tr>
<tr><td>66</td><td>gi|72117689|gb|AAZ59952.1| Acriflavin resistance protein [Ralstonia eutropha JMP134]</td><td>128</td><td>9.00E-29</td><td>496</td><td>2</td><td>pfam02355, PRK10614</td></tr>
<tr><td>67</td><td>gi|54287494|gb|AAV31238.1| putative 26S proteasome non-ATPase regulatory subunit 14 [Oryza sativa (japonica cultivar-group)]</td><td>536</td><td>0</td><td>1247</td><td>12</td><td>smart00232</td></tr>
<tr><td>68</td><td>gi|50428772|gb|AAT77096.1| Fom-2 protein [Cucumis melo]</td><td>252</td><td>3.00E-65</td><td>1852</td><td>3</td><td>No CD has been identified</td></tr>
<tr><td>69</td><td>gi|15217954|ref|NP_176135.1| disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]</td><td>172</td><td>2.00E-45</td><td>867</td><td>4</td><td>pfam00931</td></tr>
<tr><td>70</td><td>gi|15238507|ref|NP_198395.1| disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]</td><td>110</td><td>9.00E-23</td><td>912</td><td>4</td><td>pfam00931, COG4886</td></tr>
<tr><td>71</td><td>gi|50939001|ref|XP_479028.1| putative disease resistance protein RPH8A [Oryza sativa (japonica cultivar-group)]</td><td>227</td><td>2.00E-77</td><td>1443</td><td>3</td><td>pfam00931, COG4886</td></tr>
<tr><td>72</td><td>gi|15220948|ref|NP_172851.1| GN (GNOM) [Arabidopsis thaliana]</td><td>416</td><td>0</td><td>753</td><td>5</td><td>cd00171, COG5307</td></tr>
<tr><td>73</td><td>gi|84453202|dbj|BAE71198.1| putative transporter-like protein [Trifolium pratense]</td><td>186</td><td>8.00E-46</td><td>799</td><td>6</td><td>PRK09874, COG2814, pfam07690</td></tr>
<tr><td>74</td><td>gi|21554189|gb|AAM63268.1| putative leucine-rich repeat disease resistance protein [Arabidopsis thaliana]</td><td>432</td><td>0</td><td>1386</td><td>7</td><td>cd00116</td></tr>
<tr><td>75</td><td>gi|48093465|gb|AAT40109.1| putative UDP-glucuronate decarboxylase 3 [Nicotiana tabacum]</td><td>704</td><td>0</td><td>2073</td><td>18</td><td>PRK10084, COG0451</td></tr>
<tr><td>76</td><td>gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]</td><td>168</td><td>2.00E-40</td><td>803</td><td>3</td><td>No CD has been identified</td></tr>
<tr><td>77</td><td>gi|48788902|ref|ZP_00284881.1| COG0654: 2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases [Burkholderia fungorum LB400]</td><td>228</td><td>1.00E-58</td><td>788</td><td>12</td><td>COG0654, PRK06185</td></tr>
<tr><td>78</td><td>gi|37534400|ref|NP_921502.1| hypothetical protein [Oryza sativa (japonica cultivar-group)]</td><td>172</td><td>2.00E-57</td><td>784</td><td>2</td><td>pfam03732, pfam08284, pfam00665, pfam00385</td></tr>
<tr><td>79</td><td>gi|6456755|gb|AAF09256.1| disease resistance protein BS2 [Capsicum chacoense]</td><td>358</td><td>6.00E-97</td><td>2533</td><td>15</td><td>pfam00931</td></tr>
<tr><td>81</td><td>gi|15227133|ref|NP_182301.1| ATMRP4 [Arabidopsis thaliana]</td><td>595</td><td>0</td><td>1353</td><td>3</td><td>cd03244, cd03250, pfam00664, cd03271, cd03233, COG1132</td></tr>
<tr><td>82</td><td>gi|50939001|ref|XP_479028.1| putative disease resistance protein RPH8A [Oryza sativa (japonica cultivar-group)]</td><td>325</td><td>5.00E-87</td><td>1857</td><td>7</td><td>pfam00931, COG4886</td></tr>
<tr><td>83</td><td>gi|15225931|ref|NP_182138.1| unknown protein [Arabidopsis thaliana]</td><td>176</td><td>1.00E-42</td><td>965</td><td>5</td><td>pfam07320</td></tr>
<tr><td>84</td><td>gi|29839503|sp|P59584|RP8HA_ARATH Disease resistance protein RPH8A (RPP8 homolog A) [Arabidopsis thaliana]</td><td>293</td><td>2.00E-77</td><td>2164</td><td>5</td><td>pfam00931</td></tr>
<tr><td>85</td><td>gi|79594244|ref|NP_850245.2| hydrolase [Arabidopsis thaliana]</td><td>298</td><td>1.00E-79</td><td>902</td><td>2</td><td>cd01311</td></tr>
<tr><td>86</td><td>gi|68344526|gb|AAY92132.1| multidrug RND efflux transporter, permease protein MdtB [Pseudomonas fluorescens Pf-5]</td><td>292</td><td>3.00E-78</td><td>510</td><td>11</td><td>pfam02355, PRK10503</td></tr>
<tr><td>87</td><td>gi|39636723|gb|AAR29070.1| blight resistance protein RGA1 [Solanum bulbocastanum]</td><td>248</td><td>5.00E-64</td><td>1540</td><td>2</td><td>pfam00931, COG4886</td></tr>
<tr><td>88</td><td>gi|55773917|dbj|BAD72522.1| putative multidrug resistance-associated protein 7 [Oryza sativa (japonica cultivar-group)]</td><td>142</td><td>6.00E-33</td><td>537</td><td>2</td><td>cd03244, cd03250, pfam00664, cd03215, COG1132</td></tr>
<tr><td>89</td><td>gi|5824321|emb|CAB54139.1| ATPase [Solanum tuberosum]</td><td>372</td><td>0</td><td>1121</td><td>4</td><td>cd0055</td></tr>
<tr><td>90</td><td>gi|5478530|gb|AAD43920.1| UVB-resistance protein UVR8 [Arabidopsis thaliana]</td><td>523</td><td>0</td><td>1308</td><td>8</td><td>pfam00415, COG5184</td></tr>
<tr><td>92</td><td>gi|68342685|gb|AAY90291.1| fusaric acid resistance protein, putative [Pseudomonas fluorescens Pf-5]</td><td>478</td><td>0</td><td>798</td><td>9</td><td>pfam04632, COG4129, COG1289</td></tr>
<tr><td>93</td><td>gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]</td><td>145</td><td>2.00E-33</td><td>785</td><td>2</td><td>No CD has been identified</td></tr>
<tr><td>94</td><td>gi|15220442|ref|NP_172010.1|AXR1 (AUXIN RESISTANT 1); small protein activating enzyme [Arabidopsis thaliana]</td><td>311</td><td>2.00E-83</td><td>831</td><td>2</td><td>cd01493</td></tr>
<tr><td>97</td><td>gi|15240747|ref|NP_196345.1| EMS1 (EXCESS MICROSPOROCYTES1); kinase [Arabidopsis thaliana]</td><td>164</td><td>4.00E-39</td><td>805</td><td>3</td><td>cd00192, COG4886, cd00116, smart00220</td></tr>
<tr><td>98</td><td>gi|50939001|ref|XP_479028.1| putative disease resistance protein RPH8A [Oryza sativa (japonica cultivar-group)]</td><td>291</td><td>5.00E-77</td><td>1981</td><td>6</td><td>pfam00931, COG4886</td></tr>
<tr><td>99</td><td>gi|18416049|ref|NP_567674.1| unknown protein [Arabidopsis thaliana]</td><td>652</td><td>0</td><td>1469</td><td>6</td><td>pfam03105, pfam07690</td></tr>
<tr><td>100</td><td>gi|15227133|ref|NP_182301.1| ATMRP4 [Arabidopsis thaliana]</td><td>978</td><td>0</td><td>2599</td><td>10</td><td>cd03244, cd03250, pfam00664, cd03271, cd03233, COG1132</td></tr>
<tr><td>101</td><td>gi|15228004|ref|NP_181808.1| protein binding [Arabidopsis thaliana]</td><td>119</td><td>1.00E-25</td><td>794</td><td>2</td><td>COG0842</td></tr>
<tr><td>102</td><td>gi|56381949|gb|AAV85693.1| At5g59250 [Arabidopsis thaliana]</td><td>244</td><td>4.00E-63</td><td>881</td><td>2</td><td>pfam00083</td></tr>
<tr><td>103</td><td>gi|51971421|dbj|BAD44375.1| unnamed protein product [Arabidopsis thaliana]</td><td>390</td><td>0</td><td>875</td><td>3</td><td>cd03249, pfam00664</td></tr>
<tr><td>104</td><td>gi|50945887|ref|XP_482471.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]</td><td>137</td><td>3.00E-34</td><td>1313</td><td>4</td><td>pfam00931, COG4886</td></tr>
<tr><td>105</td><td>gi|14626935|gb|AAK70805.1| leucine-rich repeat resistance protein-like protein [Gossypium hirsutum]</td><td>332</td><td>2.00E-93</td><td>873</td><td>4</td><td>COG4886</td></tr>
<tr><td>107</td><td>gi|38327504|gb|AAR17783.1| ribosomal protein L3 [Lycopersicon esculentum]</td><td>457</td><td>0</td><td>819</td><td>2</td><td>PRK04231</td></tr>
<tr><td>108</td><td>gi|15230357|ref|NP_190664.1| ATP binding [Arabidopsis thaliana]</td><td>216</td><td>1.00E-54</td><td>1500</td><td>2</td><td>pfam00931, PRK11281</td></tr>
<tr><td>109</td><td>gi|15231620|ref|NP_191462.1| TT12 (TRANSPARENT TESTA 12); antiporter/transporter [Arabidopsis thaliana]</td><td>338</td><td>1.00E-91</td><td>872</td><td>9</td><td>pfam01554, COG0534</td></tr>
<tr><td>110</td><td>gi|74040324|gb|AAZ95005.1| late blight resistance protein Rpi-blb2 [Solanum bulbocastanum]</td><td>157</td><td>6.00E-37</td><td>1023</td><td>4</td><td>pfam00931</td></tr>
<tr><td>111</td><td>gi|77556936|gb|ABA99732.1| disease resistance, putative [Oryza sativa]</td><td>671</td><td>0</td><td>2255</td><td>13</td><td>pfam00931</td></tr>
<tr><td>112</td><td>gi|18407955|ref|NP_566879.1| MLO3; calmodulin binding [Arabidopsis thaliana]</td><td>275</td><td>1.00E-72</td><td>820</td><td>2</td><td>pfam03094</td></tr>
<tr><td>113</td><td>gi|33573112|emb|CAE36756.1| acriflavine resistance protein B [Bordetella parapertussis]</td><td>350</td><td>2.00E-95</td><td>613</td><td>2</td><td>pfam02355, PRK10555</td></tr>
<tr><td>114</td><td>gi|68344778|gb|AAY92384.1| undecaprenol kinase, putative [Pseudomonas fluorescens Pf-5]</td><td>263</td><td>6.00E-69</td><td>848</td><td>3</td><td>PRK00281</td></tr>
<tr><td>115</td><td>gi|47059739|gb|AAT09451.1| putative NBS-LRR type disease resistance protein [Prunus persica]</td><td>461</td><td>0</td><td>3281</td><td>15</td><td>pfam00931, COG4886</td></tr>
<tr><td>116</td><td>gi|56784769|dbj|BAD81990.1| putative terbinafine resistance locus protein [Oryza sativa Japonica Group]</td><td>174</td><td>1.00E-42</td><td>598</td><td>3</td><td>pfam04893</td></tr>
<tr><td>117</td><td>gi|50906405|ref|XP_464691.1| putative MAP3K delta-1 protein kinase [Oryza sativa (japonica cultivar-group)]</td><td>276</td><td>5.00E-73</td><td>850</td><td>4</td><td>cd00180, smart00220</td></tr>
</table>

118	gi|32400274|emb|CAE00640.1| putative mitogen-activated protein kinase 1 [Medicago sativa]	196	5.00E-49	702	2	cd00180, smart00220
119	gi|68345576|gb|AAY93182.1| outer membrane efflux protein [Pseudomonas fluorescens Pf-5]	310	2.00E-83	576	3	pfam02321, COG1538
120	gi|75704070|gb|ABA23746.1| ABC transporter-like [Anabaena variabilis ATCC 29413]	264	2.00E-69	598	2	cd03253, COG1132
121	gi|15220982|ref|NP_171690.1| PFC1 (PALEFACE 1) [Arabidopsis thaliana]	141	2.00E-37	687	2	PRK00274
122	gi|49533774|gb|AAT66773.1| putative late blight resistance protein [Solanum demissum]	313	1.00E-83	1266	5	pfam00931
123	gi|39636816|gb|AAR29076.1| blight resistance protein T118 [Solanum tarijense]	126	7.00E-30	768	4	pfam00931, COG4886
124	gi|34913374|ref|NP_918034.1| putative DNA damage repair protein [Oryza sativa]	201	2.00E-50	632	2	smart00361
125	gi|79328183|ref|NP_001031909.1| carboxylic ester hydrolase [Arabidopsis thaliana]	371	0	1328	14	pfam02230
126	gi|79328183|ref|NP_001031909.1| carboxylic ester hydrolase [Arabidopsis thaliana]	345	1.00E-93	868	6	pfam02230
128	gi|18394385|ref|NP_564003.1| ATP binding / kinase/ protein kinase/ protein [Arabidopsis thaliana]	404	0	760	2	cd00192
129	gi|14330718|emb|CAC40827.1| HcrVf3 protein [Malus floribunda]	145	8.00E-42	742	2	pfam08263, cd00116, COG4886
130	gi|15225286|ref|NP_180201.1| ER (ERECTA) [Arabidopsis thaliana]	179	2.00E-59	643	2	cd00180, pfam08263, cd00116, cd05086
131	gi|55139525|gb|AAV41396.1| peru 2 [Lycopersicon peruvianum]	194	2.00E-48	695	2	pfam08263
132	gi|14269079|gb|AAK58012.1| verticillium wilt disease resistance protein Ve2 [Lycopersicon esculentum]	165	2.00E-39	740	2	pfam08263, cd00116
133	gi|15235429|ref|NP_192169.1| calmodulin binding [Arabidopsis thaliana]	499	0	1828	4	pfam03094
134	gi|50945887|ref|XP_482471.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]	246	1.00E-63	1619	5	pfam00931, COG4886
135	gi|18478787|gb|AAL73330.1| putative receptor-like protein kinase RLPK1 [Glycine max]	101	2.00E-20	663	2	cd00192
136	gi|15225814|ref|NP_180259.1| ATPase, coupled to transmembrane movement of substances [Arabidopsis thaliana]	756	0	1725	6	cd03232, cd03233, pfam01061, pfam08370, cd03262, pfam00005
137	gi|79330795|ref|NP_001032070.1| MLO11; calmodulin binding [Arabidopsis thaliana]	183	3.00E-68	778	3	pfam03094
138	gi|28436071|gb|AAO41731.1| cytoplasmic ribosomal protein S14 [Brassica napus]	234	2.00E-60	826	11	PRK09607
139	gi|33637487|gb|AAQ23899.1| RSH2 [Nicotiana tabacum]	1100	0	2333	10	pfam04607, pfam01966
140	gi|15231046|ref|NP_191408.1| phosphoric ester hydrolase [Arabidopsis thaliana]	381	0	1145	4	cd03388
141	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	217	5.00E-55	962	2	pfam00931
142	gi|77689672|ref|ZP_00804853.1| Peptidase M15D, vanX D-ala-D-ala dipeptidase [Rhodopseudomonas palustris BisB5]	119	1.00E-25	673	2	COG2173
143	gi|42733544|dbj|BAD11207.1| multidrug resistance-associated protein [Thlaspi caerulescens]	226	5.00E-72	665	2	cd03244, cd03250, pfam00664, cd03224, COG1132
144	gi|33576997|emb|CAE34077.1| putative ABC transporter [Bordetella bronchiseptica RB50]	212	3.00E-54	337	4	cd03249, pfam00664, COG1132
145	gi|82735577|ref|ZP_00898439.1| major facilitator family transporter [Pseudomonas putida F1]	414	0	690	6	COG2814, pfam0769
146	gi|50918091|ref|XP_469442.1| putative protein kinase [Oryza sativa (japonica cultivar-group)]	197	2.00E-49	740	2	cd00180, smart00220
147	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	194	8.00E-48	1246	2	pfam00931
148	gi|10121909|gb|AAG13419.1| T7N9.24 [Arabidopsis thaliana]	132	4.00E-60	1059	2	smart00255, cd00116, pfam00931, COG4886
150	gi|27368843|emb|CAD59579.1| MDR-like ABC transporter [Oryza sativa (japonica cultivar-group)]	144	4.00E-53	1311	2	cd03249, pfam00664, COG1132
151	gi|49086916|gb|AAT51392.1| PA4599 [synthetic construct]	183	2.00E-45	501	3	PRK09578, pfam00529
152	gi|34909106|ref|NP_915900.1| putative NBS-LRR type resistance protein [Oryza sativa (japonica cultivar-group)]	100	5.00E-25	900	2	pfam00931, COG4886
153	gi|71040679|gb|AAZ20288.1| disease resistance-responsive family protein [Arachis hypogaea]	146	4.00E-34	612	2	pfam03018
154	gi|41052474|dbj|BAD07484.1| PDR-type ABC transporter 2 [Nicotiana tabacum]	380	0	835	3	cd03232, cd03233, pfam01061, pfam08370, pfam00005
155	gi|33572422|emb|CAE41988.1| muramoyltetrapeptide carboxypeptidase [Bordetella pertussis Tohama I]	211	6.00E-54	508	2	PRK11253
156	gi|32400274|emb|CAE00640.1| putative mitogen-activated protein kinase 1 [Medicago sativa]	139	7.00E-32	702	2	cd00180, smart00220
159	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	269	2.00E-70	1318	3	pfam00931
160	gi|22331862|ref|NP_191473.2| ATMRP14 [Arabidopsis thaliana]	318	3.00E-85	994	2	cd03244, cd03250, pfam00664, cd03296, COG1132
162	gi|77384019|gb|ABA75532.1| Drug resistance transporter EmrB/QacA subfamily [Pseudomonas fluorescens PfO-1]	323	3.00E-87	719	5	PRK10504, pfam07690
163	gi|18423435|ref|NP_568777.1| WRKY27; transcription factor [Arabidopsis thaliana]	140	5.00E-32	787	2	pfam03106
164	gi|4097585|gb|AAD09518.1| NTGP4 [Nicotiana tabacum]	355	2.00E-96	1384	9	cd01852, COG4717
165	gi|33571779|emb|CAE41286.1| acriflavine resistance protein B [Bordetella pertussis Tohama I]	145	5.00E-34	255	2	PRK10555
168	gi|71040679|gb|AAZ20288.1| disease resistance-responsive family protein [Arachis hypogaea]	161	3.00E-38	789	2	pfam03018
169	gi|15235429|ref|NP_192169.1| calmodulin binding [Arabidopsis thaliana]	661	0	1458	4	pfam03094
171	gi|38566726|emb|CAE76632.1| leucine rich repeat protein [Cicer arietinum]	500	0	1315	3	pfam08263, cd00116
172	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	194	2.00E-56	1239	3	No CD has been identified
173	gi|15233846|ref|NP_194179.1| ATP binding / kinase/ protein kinase/ protein [Arabidopsis thaliana]	169	9.00E-41	733	2	cd00192
174	gi|34904948|ref|NP_913821.1| putative zinc transporter [Oryza sativa]	216	2.00E-54	1436	4	COG1230
175	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	105	1.00E-21	675	2	pfam00931
176	gi|50900396|ref|XP_450986.1| putative ABC transporter [Oryza sativa (japonica cultivar-group)]	490	0	1415	6	cd03213, pfam00005, pfam01061
178	gi|12006354|gb|AAG44839.1| putative Hs1pro-1-like receptor [Glycine max]	344	3.00E-93	1029	7	pfam07014, pfam07231
180	gi|42602161|gb|AAS21681.1| receptor-like kinase [Arabidopsis thaliana]	265	3.00E-69	1304	2	cd00192, COG4886, smart00220
182	gi|15223013|ref|NP_177760.1| unknown protein [Arabidopsis thaliana]	269	6.00E-71	686	2	COG5070
183	gi|77689672|ref|ZP_00804853.1| Peptidase M15D, vanX D-ala-D-ala dipeptidase [Rhodopseudomonas palustris BisB5]	120	4.00E-26	684	7	COG2173

184	gi|66734028|gb|AAY53482.1| resistance protein R1 [Solanum demissum]	290	8.00E-77	1239	5	pfam00931
185	gi|47717725|gb|AAT37905.1| multidrug-resistance associated protein 3 [Zea mays]	350	3.00E-95	720	2	cd03244, cd03250, pfam00664, COG1132
187	gi|15230899|ref|NP_188597.1| catalytic/ iron ion binding [Arabidopsis thaliana]	379	0	1367	3	pfam04055, COG0820
188	gi|30687745|ref|NP_173637.3| unknown protein [Arabidopsis thaliana]	197	2.00E-60	770	2	cd00170, pfam03765
189	gi|8547232|gb|AAF76308.1| Prf [Lycopersicon pimpinellifolium]	196	1.00E-61	744	2	pfam00931
190	gi|30267436|gb|AAP21819.1| metal transporter [Lycopersicon esculentum]	783	0	2216	12	pfam01566
191	gi|34905796|ref|NP_914245.1| P0401G10.24 [Oryza sativa (japonica cultivar-group)]	132	1.00E-29	758	2	pfam03656
192	gi|68560019|ref|ZP_00599347.1| Heavy metal efflux pump CzcA [Ralstonia metallidurans CH34]	343	4.00E-93	656	2	COG3696
193	gi|24051889|gb|AAN43154.1| orf, conserved hypothetical protein [Shigella flexneri 2a str.301]	240	5.00E-62	817	7	PRK10995
194	gi|41052474|dbj|BAD07484.1| PDR-type ABC transporter 2 [Nicotiana tabacum]	917	0	2264	29	cd03232, cd03233, pfam01061, pfam08370, pfam00005
195	gi|82794018|gb|ABB91438.1| R-FOM-2 [Cucumis melo]	167	9.00E-40	1384	2	pfam00931
196	gi|15239535|ref|NP_197963.1| kinase/ protein binding [Arabidopsis thaliana]	104	4.00E-21	770	2	pfam08263, cd00116
197	gi|34910256|ref|NP_916475.1| putative MRP-like ABC transporter [Oryza sativa]	357	3.00E-97	788	2	cd03244, cd03250, pfam00664, COG1132
198	gi|22137208|gb|AAM91449.1| At1g53210/F12M16_12 [Arabidopsis thaliana]	306	6.00E-82	724	2	pfam01699, cd00051
199	gi|15231938|ref|NP_188102.1| leucine-rich repeat family protein / protein kinase family protein [Arabidopsis thaliana]	293	1.00E-77	1074	4	cd00192, cd00116
200	gi|12324080|gb|AAG52008.1| unknown protein; 23065-20358 [Arabidopsis thaliana]	236	8.00E-61	1037	3	pfam02535
202	gi|58826317|gb|AAW82883.1| phytoalexin-deficient 4-1 protein [Solanum tuberosum]	528	0	1506	5	cd00519
203	gi|14626935|gb|AAK70805.1| leucine-rich repeat resistance protein-like protein [Gossypium hirsutum]	531	0	1539	8	COG4886
204	gi|21552979|gb|AAM62409.1| Rar1 [Nicotiana tabacum]	349	7.00E-95	905	3	pfam04968
207	gi|15232624|ref|NP_190257.1| ATP binding [Arabidopsis thaliana]	254	8.00E-84	2331	8	pfam00931
208	gi|50939001|ref|XP_479028.1| putative disease resistance protein RPH8 [Oryza sativa (japonica cultivar-group)]	390	0	2376	4	pfam00931, COG4886
209	gi|42572237|ref|NP_974213.1| binding [Arabidopsis thaliana]	381	0	1176	3	cd00204, COG5184, pfam00415
210	gi|33568473|emb|CAE32385.1| putative multidrug resistance protein [Bordetella bronchiseptica RB50]	210	4.00E-53	691	2	PRK10476, COG1566
211	gi|48093465|gb|AAT40109.1| putative UDP-glucuronate decarboxylase 3 [Nicotiana tabacum]	147	6.00E-34	862	7	PRK10084, COG0451
212	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	412	0	1961	9	pfam00931
213	gi|55733942|gb|AAV59449.1| putative MRP-like ABC transporter [Oryza sativa]	298	0	1171	2	cd03244, cd03250, pfam00664, cd03297, COG1132, pfam00005
214	gi|26450926|dbj|BAC42570.1| putative receptor protein kinase [Arabidopsis thaliana]	286	7.00E-76	922	2	cd00192, pfam08263, COG4886, smart00220
215	gi|5231014|gb|AAD41050.1| NBS/LRR disease resistance protein RPM1 [Arabidopsis thaliana]	114	4.00E-24	694	2	pfam00931, COG4886
216	gi|82794018|gb|ABB91438.1| R-FOM-2 [Cucumis melo]	105	1.00E-21	672	2	pfam00931
218	gi|3334219|sp|O23920|HPPD_DAUCA 4-hydroxyphenylpyruvate dioxygenase (4HPPD) [Daucus carota]	233	8.00E-65	773	2	pfam00903
219	gi|62632825|gb|AAX89383.1| NBS-LRR type disease resistance protein Rps1-k-2 [Glycine max]	156	8.00E-37	848	2	pfam00931, COG4886
220	gi|41052472|dbj|BAD07483.1| PDR-type ABC transporter 1 [Nicotiana tabacum]	408	0	1677	4	cd03232, cd03233, cd03234, pfam01061, pfam08370, cd03226, pfam00005
221	gi|15229564|ref|NP_189044.1| unknown protein [Arabidopsis thaliana]	108	4.00E-23	382	2	pfam03018
222	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	124	5.00E-27	802	2	pfam00931
223	gi|38489222|gb|AAR21296.1| NDR1-like protein [Nicotiana benthamiana]	206	6.00E-52	703	4	pfam07320
224	gi|50945887|ref|XP_482471.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]	207	1.00E-51	1693	3	pfam00931, COG4886
225	gi|30693847|ref|NP_190920.2| UXS1 (UDP-GLUCURONIC ACID DECARBOXYLASE 1); catalytic [Arabidopsis thaliana]	159	2.00E-37	855	2	PRK10084, COG0451
226	gi|52854207|gb|AAU88159.1| disease resistance-like protein [Coffea arabica]	261	1.00E-68	778	2	pfam00931
228	gi|68348514|gb|AAY96120.1| multidrug resistance transporter, Bcr/CflA family [Pseudomonas fluorescens Pf-5]	347	3.00E-94	860	9	PRK11102, pfam07690
229	gi|18414234|ref|NP_568119.1| CDC25 [Arabidopsis thaliana]	211	2.00E-53	765	5	cd01531
230	gi|44888877|gb|AAS48176.1| Osmolarity-sensing histidine-kinase [Pseudomonas fluorescens]	400	0	709	5	cd00075, pfam00672, PRK09467
231	gi|32364526|gb|AAP80292.1| resistance protein Tsu5 [Arabidopsis thaliana]	105	2.00E-21	902	2	pfam00931
232	gi|24459845|emb|CAC82610.1| disease resistance-like protein [Coffea arabica]	161	2.00E-58	1047	2	pfam00931
234	gi|15223416|ref|NP_171656.1| metal ion binding [Arabidopsis thaliana]	100	3.00E-20	640	2	pfam00403
235	gi|83283975|gb|ABC01895.1| transporter-like protein [Solanum tuberosum]	338	1.00E-91	804	4	cd00170, pfam03765
236	gi|29409364|gb|AAM29178.1| biostress-resistance-related protein [Triticum aestivum]	308	2.00E-82	840	2	pfam02230
237	gi|42562729|ref|NP_175747.2| serine/threonine protein kinase-related [Arabidopsis thaliana]	207	6.00E-52	1360	3	cd00192, cd00116
238	gi|18402572|ref|NP_566659.1| unknown protein [Arabidopsis thaliana]	374	0	1240	3	pfam02470, COG1463
239	gi|15230686|ref|NP_187915.1| ATMRP3 [Arabidopsis thaliana]	315	0	1282	2	cd03244, cd03250, pfam00664, COG1132, cd03224
241	gi|34907356|ref|NP_915025.1| putative receptor protein kinase [Oryza sativa]	348	1.00E-94	880	2	pfam08263, COG4886, cd00116, smart00220, cd00180
242	gi|62896437|emb|CAD91352.2| mercuric ion reductase MerA [Pseudomonas fluorescens]	142	3.00E-33	256	21	cd00371, pfam02852, pfam00070, COG1249, PRK10671
243	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	233	1.00E-59	1114	4	pfam00931
244	gi|33238810|gb|AAQ00875.1| PLP-dependent enzyme [Prochlorococcus marinus str. CCMP1375]	207	2.00E-84	1558	3	pfam06838
245	gi|48093461|gb|AAT40107.1| UDP-glucuronate decarboxylase 1 [Nicotiana tabacum]	658	0	1437	10	PRK11908, COG0451
246	gi|4689223|gb|AAD27815.1| disease resistance protein I2 [Lycopersicon esculentum]	130	6.00E-29	715	2	pfam00931
247	gi|15235205|ref|NP_193719.1| ATPP2-A1 [Arabidopsis thaliana]	147	5.00E-34	1009	5	No CD has been identified
248	gi|77380404|gb|ABA71917.1| Secretion protein HlyD [Pseudomonas fluorescens PfO-1]	146	7.00E-34	805	11	PRK10559, COG1566
249	gi|7677046|gb|AAF67003.1| putative Hs1pro-1 homolog [Pisum sativum]	188	1.00E-46	611	2	pfam07014, pfam07231

250	gi|55771369|dbj|BAD72536.1| LIM domain containing protein-like [Oryza sativa]	438	0	885	3	smart00132
251	gi|47059739|gb|AAT09451.1| putative NBS-LRR type disease resistance protein [Prunus persica]	201	5.00E-50	1140	4	pfam00931, COG4886
252	gi|53727578|ref|ZP_00347912.1| COG1566: Multidrug resistance efflux pump [Pseudomonas aeruginosa UCBPP-PA14]	238	2.00E-61	764	12	COG1566, PRK10476
253	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	148	2.00E-34	729	2	pfam00931
254	gi|5734781|gb|AAD50046.1| Very similar to disease resistance proteins [Arabidopsis thaliana]	142	7.00E-33	655	2	pfam00931
255	gi|55773694|dbj|BAD72277.1| 40S ribosomal protein S30-like [Oryza sativa]	103	2.00E-21	476	2	pfam04758
257	gi|68342709|gb|AAY90315.1| multidrug efflux RND transporter [Pseudomonas fluorescens Pf-5]	348	1.00E-94	885	15	pfam02355, PRK09579
258	gi|24459853|emb|CAC82602.1| disease resistance-like protein [Coffea arabica]	300	9.00E-80	1131	5	pfam00931
259	gi|30698733|ref|NP_177218.3| ATTAP1; ATPase, coupled to transmembrane movement of substances [Arabidopsis thaliana]	454	0	899	3	cd03249, pfam00664, COG1132
262	gi|3894387|gb|AAC78593.1| Hcr2-0B [Lycopersicon esculentum]	127	4.00E-28	720	2	cd00116, pfam08263
264	gi|82794018|gb|ABB91438.1| R-FOM-2 [Cucumis melo]	104	8.00E-21	1558	3	pfam00931
265	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	216	2.00E-54	1229	4	No CD has been identified
266	gi|30698151|ref|NP_201372.2| ATP binding / kinase/ protein serine/threonine kinase [Arabidopsis thaliana]	200	1.00E-83	1204	3	cd00116, pfam08263, COG4886, cd00180
267	gi|77967250|gb|ABB08630.1| Acriflavin resistance protein [Burkholderia sp. 383]	105	1.00E-21	593	3	COG0841
268	gi|50899184|ref|XP_450380.1| putative disease related protein 2 [Oryza sativa]	138	2.00E-65	1886	4	pfam00931, COG4886
270	gi|33564744|emb|CAE44069.1| AcrB/AcrD/AcrF family protein [Bordetella pertussis Tohama I]	147	4.00E-34	643	4	PRK10503
271	gi|15235205|ref|NP_193719.1| ATPP2-A1 [Arabidopsis thaliana]	150	3.00E-35	689	2	No CD has been identified
272	gi|15217785|ref|NP_171753.1| ATPase, coupled to transmembrane movement of substances [Arabidopsis thaliana]	321	2.00E-86	804	5	cd03249, pfam00664, COG1132
273	gi|46389753|dbj|BAD15107.1| hypothetical protein [Nicotiana tabacum]	139	2.00E-31	934	3	No CD has been identified
274	gi|50945889|ref|XP_482472.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]	145	1.00E-33	779	4	pfam00931, COG4886
275	gi|66737320|gb|AAY54606.1| NRG1 [Nicotiana benthamiana]	224	4.00E-57	936	2	pfam05659, pfam00931, COG4886
276	gi|49533774|gb|AAT66773.1| putative late blight resistance protein, identical [Solanum demissum]	164	3.00E-39	786	2	pfam00931
277	gi|14715462|dbj|BAB62040.1| CjMDR1 [Coptis japonica]	645	0	1569	12	cd03249, pfam00664, COG1132
278	gi|30690355|ref|NP_850885.1| D111/G-patch domain-containing protein [Arabidopsis thaliana]	168	2.00E-40	774	2	smart00443
279	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	105	4.00E-21	1325	3	pfam00931
280	gi|22331862|ref|NP_191473.2| ATMRP14 [Arabidopsis thaliana]	481	0	1093	4	cd03244, cd03250, pfam00664, cd03296, COG1132
281	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	205	2.00E-51	1058	2	pfam00931
282	gi|79319849|ref|NP_001031180.1| unknown protein [Arabidopsis thaliana]	115	2.00E-49	1167	4	pfam04859
283	gi|21239382|gb|AAM44274.1| receptor-like kinase RHG1 [Glycine max]	375	0	793	2	smart00220, pfam08263, COG4886, cd00180
285	gi|18408274|ref|NP_564850.1| ATNAP11 [Arabidopsis thaliana]	466	0	1579	5	cd03261, COG1127
286	gi|34913790|ref|NP_918242.1| putative LRR [Oryza sativa (japonica cultivar-group)]	143	2.00E-34	799	2	pfam00931, COG4886
287	gi|10177549|dbj|BAB10828.1| ABC transporter-like protein [Arabidopsis thaliana]	341	2.00E-92	824	2	cd03249, pfam00664, COG1132
289	gi|24984168|gb|AAN68204.1| ABC transporter, permease/ATP-binding protein, putative [Pseudomonas putida KT2440]	231	2.00E-59	589	11	cd03253,pfam00664, cd03223, COG5265
290	gi|82735577|ref|ZP_00898439.1| major facilitator family transporter [Pseudomonas putida F1]	417	0	699	6	COG2814, pfam07690
291	gi|16974114|emb|CAC95155.1| putative resistance protein [Solanum lycopersicum]	469	0	1466	9	pfam03798
292	gi|55297132|dbj|BAD68775.1| putative NBS-LRR disease resistance protein [Oryza sativa Japonica Group]	124	4.00E-27	814	3	pfam00931
293	gi|53689783|gb|AAU89792.1| late blight resistance protein-like [Solanum tuberosum]	152	1.00E-35	803	2	pfam00931, smart00534
294	gi|77381310|gb|ABA72823.1| Acriflavin resistance protein [Pseudomonas fluorescens PfO-1]	404	0	781	5	COG0841
295	gi|2292907|emb|CAA71179.1| P-glycoprotein homologue [Hordeum vulgare]	438	0	1166	3	cd03249, pfam00664, COG1132, cd03271
296	gi|52854237|gb|AAU88174.1| disease resistance-like protein [Psilanthus wightianus]	143	1.00E-32	1237	3	pfam00931
297	gi|39636816|gb|AAR29076.1| blight resistance protein T118 [Solanum tarijense]	172	2.00E-41	1124	2	pfam00931, COG4886
298	gi|42570865|ref|NP_973506.1| SAE2 (SUMO-ACTIVATING ENZYME 2) [Arabidopsis thaliana]	286	8.00E-76	818	4	cd01489
299	gi|34104782|gb|AAQ61139.1| probable manganese transport transmembrane protein [Chromobacterium violaceum ATCC 12472]	178	1.00E-43	620	4	COG1914
301	gi|82568697|dbj|BAE48662.1| alcohol dehydrogenase [Prunus mume]	172	3.00E-42	509	2	PRK12825, PRK05653
302	gi|8547237|gb|AAF76312.1| Prf [Lycopersicon esculentum]	411	0	1667	5	pfam00931
303	gi|15220948|ref|NP_172851.1| GN (GNOM) [Arabidopsis thaliana]	365	0	942	2	cd00171, COG5307
304	gi|6671365|gb|AAF23176.1| P-glycoprotein [Gossypium hirsutum]	456	0	908	3	cd03249, pfam00664, COG1132, cd03270
305	gi|42567508|ref|NP_195551.2| antiporter/ drug transporter [Arabidopsis thaliana]	251	2.00E-65	851	2	pfam01554
306	gi|62734129|gb|AAX96238.1| NB-ARC domain, putative [Oryza sativa (japonica cultivar-group)]	111	5.00E-23	1259	3	pfam00931
307	gi|47572215|ref|ZP_00242260.1| COG0798: Arsenite efflux pump ACR3 and related permeases [Rubrivivax gelatinosus PM1]	107	1.00E-22	196	4	COG0798
308	gi|34910752|ref|NP_916723.1| P0042A10.5 [Oryza sativa (japonica cultivar-group)]	499	0	1353	2	COG5184, pfam00415
309	gi|15237733|ref|NP_201260.1| unknown protein [Arabidopsis thaliana]	463	0	835	6	No CD has been identified
310	gi|18398444|ref|NP_565416.1| MLO8; calmodulin binding [Arabidopsis thaliana]	602	0	1672	8	pfam03094
311	gi|18400769|ref|NP_566512.1| Ran GTPase binding [Arabidopsis thaliana]	433	0	1028	2	COG5184
312	gi|4204793|gb|AAD10836.1| P-glycoprotein [Solanum tuberosum]	457	0	944	5	cd03249, pfam00664, COG1132
313	gi|15230349|ref|NP_188563.1| protein binding [Arabidopsis thaliana]	206	6.00E-60	1107	4	cd00116
315	gi|15222893|ref|NP_175437.1| ATP binding [Arabidopsis thaliana]	108	2.00E-22	711	2	pfam00931
316	gi|28436071|gb|AAO41731.1| cytoplasmic ribosomal protein S14 [Brassica napus]	232	1.00E-59	889	13	PRK09607
317	gi|14269079|gb|AAK58012.1| verticillium wilt disease resistance protein Ve2	369	0	1255	3	cd00116, pfam08263

	[Lycopersicon esculentum]					
318	gi|15220604|ref|NP_176961.1| ATNAP3 [Arabidopsis thaliana]	361	3.00E-98	1361	11	cd03260, COG1126
319	gi|30267434|gb|AAP21818.1| root-specific metal transporter [Lycopersicon esculentum]	333	5.00E-90	785	2	pfam01566
321	gi|15222893|ref|NP_175437.1| ATP binding [Arabidopsis thaliana]	338	4.00E-91	1906	5	pfam00931
322	gi|42550761|gb|EAA73604.1| hypothetical protein FG04278.1 [Gibberella zeae PH-1]	294	0	733	2	pfam03452
323	gi|50929779|ref|XP_474417.1| OSJNBa0088H09.13 [Oryza sativa (japonica cultivar-group)]	206	2.00E-55	820	4	pfam04939
324	gi|6456755|gb|AAF09256.1| disease resistance protein BS2 [Capsicum chacoense]	216	8.00E-98	2019	6	pfam00931
325	gi|3288065|emb|CAB09800.1| RisS protein [Bordetella bronchiseptica]	391	0	725	2	cd00075, smart00304, smart00388, COG3920
326	gi|15220149|ref|NP_175157.1| NRAMP2; metal ion transporter [Arabidopsis thaliana]	620	0	1205	3	pfam01566
327	gi|7110565|gb|AAF36987.1| viral resistance protein [Arabidopsis thaliana]	125	3.00E-27	1054	2	pfam00931
328	gi|15219396|ref|NP_173131.1| unknown protein [Arabidopsis thaliana]	245	1.00E-70	972	4	No CD has been identified
329	gi|18408943|ref|NP_564921.1| IAR1 (IAA-ALANINE RESISTANT 1); metal ion transporter [Arabidopsis thaliana]	352	1.00E-95	1102	2	pfam02535
330	gi|22652532|gb|AAN03742.1| NBS-LRR-like protein [Oryza sativa (japonica cultivar-group)]	222	6.00E-56	2251	6	pfam00931
331	gi|34911420|ref|NP_917057.1| putative leucine rich repeat containing protein kinase [Oryza sativa (japonica cultivar-group)]	265	1.00E-69	706	2	cd00192, COG4886
332	gi|15218909|ref|NP_176187.1| ATP binding [Arabidopsis thaliana]	181	7.00E-44	1580	3	pfam00931
333	gi|77964582|gb|ABB05963.1| Drug resistance transporter, EmrB/QacA subfamily [Burkholderia sp. 383]	280	2.00E-74	499	2	PRK10504, pfam07690
334	gi|20466556|gb|AAM20595.1| integral membrane protein, putative [Arabidopsis thaliana]	332	8.00E-90	751	3	pfam01554, COG0534
335	gi|79317904|ref|NP_001031037.1| unknown protein [Arabidopsis thaliana]	166	1.00E-39	1104	2	pfam03798
336	gi|66737320|gb|AAY54606.1| NRG1 [Nicotiana benthamiana]	210	6.00E-53	870	3	pfam05659, pfam00931, COG4886
338	gi|77551506|gb|ABA94303.1| AWJL218 protein [Oryza sativa (japonica cultivar-group)]	311	5.00E-83	1800	3	pfam08263, COG4886
339	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	247	8.00E-64	1290	3	pfam00931
340	gi|50945887|ref|XP_482471.1| putative disease resistance gene homolog [Oryza sativa (japonica cultivar-group)]	145	1.00E-51	1558	4	pfam00931, COG4886
341	gi|48057628|gb|AAT39951.1| putative disease resistance protein Prf [Solanum demissum]	241	1.00E-61	2213	8	pfam00931
343	gi|71081904|gb|AAZ23261.1| senescence-associated protein [Nicotiana tabacum]	211	4.00E-53	966	3	cd00158
344	gi|48209881|gb|AAT40487.1| putative disease resistance protein [Solanum demissum]	146	6.00E-34	639	2	pfam00931
345	gi|42571555|ref|NP_973868.1| TGD1 (TRIGALACTOSYLDIACYLGLYCEROL 1) [Arabidopsis thaliana]	474	0	1390	5	pfam02405
346	gi|14331118|emb|CAC40990.1| ABC1 protein [Nicotiana plumbaginifolia]	133	6.00E-30	683	3	cd03232, cd03233, cd03234, pfam01061, pfam08370, cd03259, cd03247
347	gi|71081904|gb|AAZ23261.1| senescence-associated protein [Nicotiana tabacum]	212	1.00E-53	817	4	cd00158

Table S3: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Hypersensitive, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Hypersensitive						
Contig	BLAST NR	Score	e-value	Length	Reads	Conserved Domains
1	gi|21386975|gb|AAM47891.1| hypersensitive-induced response protein [Arabidopsis thaliana]	199	2.00E-49	1075	4	cd03407
2	gi|34484310|gb|AAQ72788.1| hypersensitive-induced response protein [Cucumis sativus]	487	0	1576	16	cd03407
3	gi|11994123|dbj|BAB01125.1| unnamed protein product [Arabidopsis thaliana]	242	3.00E-66	1008	5	cd00128
4	gi|21386975|gb|AAM47891.1| hypersensitive-induced response protein [Arabidopsis thaliana]	475	0	1241	6	cd03407
5	gi|48209898|gb|AAT40492.1| putative hypersensitive-induced reaction protein [Solanum demissum]	514	0	1296	24	cd03407
6	gi|15450553|gb|AAK96454.1| AT5g49480/K6M13_2 [Arabidopsis thaliana]	150	5.00E-35	805	6	cd00051
7	gi|46452120|gb|AAS98165.1| hypersensitive-induced reaction protein [Capsicum annuum]	520	0	1118	19	cd03407
8	gi|34484310|gb|AAQ72788.1| hypersensitive-induced response protein [Cucumis sativus]	139	8.00E-32	790	2	cd03407

Table S4: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Pathogenesis, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Pathogenesis						
Contig	**BLAST NR**	**Score**	**e-value**	**Length**	**Reads**	**Conserved Domains**
1	gi|467689|emb|CAA55128.1| chitinase/lysozyme [Nicotiana tabacum]	284	9.00E-79	808	2	smart00636, pfam00704
2	gi|60280807|gb|AAX18296.1| major allergen Mal d 1.0501 [Malus x domestica]	166	8.00E-40	807	13	pfam00407
3	gi|603882|emb|CAA87072.1| pathogenesis-related protein PR-3 type [Sambucus nigra]	362	6.00E-99	804	3	cd00325, cd00035
4	gi|45773916|gb|AAS76762.1| At3g14067 [Arabidopsis thaliana]	127	3.00E-28	774	3	cd02120, pfam00082, pfam05922
5	gi|3413481|emb|CAA07250.1| serine protease [Lycopersicon esculentum]	218	2.00E-67	787	2	cd02120, pfam00082, pfam05922
6	gi|32401255|gb|AAP80801.1| class VII chitinase precursor [Gossypium hirsutum]	198	2.00E-49	713	3	cd00325
7	gi|21592749|gb|AAM64698.1| putative thaumatin-like protein [Arabidopsis thaliana]	325	2.00E-87	1592	3	smart00205
9	gi|6723683|emb|CAB67120.1| subtilisin-like protease [Lycopersicon esculentum]	230	3.00E-59	754	3	cd02120, pfam00082, pfam05922
10	gi|21592749|gb|AAM64698.1| putative thaumatin-like protein [Arabidopsis thaliana]	304	3.00E-81	911	7	smart00205
11	gi|4586372|dbj|BAA74546.2| thaumatin-like protein SE39b [Nicotiana tabacum]	313	5.00E-84	1011	4	smart00205
12	gi|7531180|sp|O04681|PTI5_LYCES Pathogenesis-related genes transcriptional activator PTI5 (PTO-interacting protein 5) [Lycopersicon esculentum]	158	2.00E-37	819	6	cd00018
13	gi|60280803|gb|AAX18294.1| major allergen Mal d 1.04 [Malus x domestica]	188	1.00E-46	735	2	pfam00407
14	gi|60280807|gb|AAX18296.1| major allergen Mal d 1.0501 [Malus x domestica]	184	6.00E-45	1068	81	pfam00407
15	gi|60280807|gb|AAX18296.1| major allergen Mal d 1.0501 [Malus x domestica]	166	1.00E-39	993	183	pfam00407
16	gi|60280803|gb|AAX18294.1| major allergen Mal d 1.04 [Malus x domestica]	168	2.00E-40	777	4	pfam00407
17	gi|40824063|gb|AAR92335.1| At1g12440 [Arabidopsis thaliana]	160	9.00E-38	1030	12	pfam01428
18	gi|60280803|gb|AAX18294.1| major allergen Mal d 1.04 [Malus x domestica]	165	2.00E-39	824	13	pfam00407
19	gi|60280803|gb|AAX18294.1| major allergen Mal d 1.04 [Malus x domestica]	163	7.00E-39	810	2	pfam00407
20	gi|7269429|emb|CAB79433.1| putative pathogenesis-related protein [Arabidopsis thaliana]	236	9.00E-61	975	3	cd00168
21	gi|60280807|gb|AAX18296.1| major allergen Mal d 1.0501 [Malus x domestica]	183	7.00E-45	830	15	pfam00407
22	gi|60280807|gb|AAX18296.1| major allergen Mal d 1.0501 [Malus x domestica]	166	8.00E-40	803	3	pfam00407
23	gi|71057064|emb|CAI38795.2| thaumatin-like protein [Actinidia deliciosa]	400	0	1017	41	pfam00314
24	gi|34897932|ref|NP_910312.1| ozone-responsive stress-related protein-like [Oryza sativa (japonica cultivar-group)]	138	1.00E-31	572	8	pfam06592
25	gi|53830843|gb|AAU95244.1| putative thaumatin-like protein [Solanum tuberosum]	371	0	975	56	pfam00314
26	gi|71057064|emb|CAI38795.2| thaumatin-like protein [Actinidia deliciosa]	404	0	892	15	pfam00314
27	gi|19976|emb|CAA41439.1| pathogenesis-related protein P2 [Lycopersicon esculentum]	216	4.00E-55	652	2	pfam00967
28	gi|33413141|emb|CAD60273.1| putative pathogenesis related protein 1 precursor [Vitis vinifera]	241	2.00E-62	761	9	cd00168
29	gi|33413141|emb|CAD60273.1| putative pathogenesis related protein 1 precursor [Vitis vinifera]	212	9.00E-54	818	28	cd00168
30	gi|33329390|gb|AAQ10092.1| thaumatin-like protein [Vitis vinifera]	381	0	1813	117	pfam00314
31	gi|14290153|gb|AAK59278.1| thaumatin-like protein [Sambucus nigra]	368	0	880	21	pfam00314
32	gi|60280803|gb|AAX18294.1| major allergen Mal d 1.04 [Malus x domestica]	166	1.00E-39	771	14	pfam00407
33	gi|33413141|emb|CAD60273.1| putative pathogenesis related protein 1 precursor [Vitis vinifera]	213	5.00E-54	657	2	cd00168
34	gi|2677826|gb|AAB97141.1| major allergen protein homolog [Prunus armeniaca]	147	4.00E-34	902	4	pfam00407
35	gi|3901018|emb|CAA10235.1| stress and pathogenesis-related protein [Fagus sylvatica]	214	4.00E-54	822	4	pfam00407
36	gi|6273385|gb|AAF06347.1| SCUTL2 [Vitis vinifera]	228	5.00E-84	925	10	smart00205
37	gi|7531180|sp|O04681|PTI5_LYCES Pathogenesis-related genes transcriptional activator PTI5 (PTO-interacting protein 5) [Lycopersicon esculentum]	172	1.00E-41	1047	3	cd00018
38	gi|32815927|gb|AAP88348.1| At3g12630 [Arabidopsis thaliana]	154	6.00E-36	1280	13	pfam01428
39	gi|4586372|dbj|BAA74546.2| thaumatin-like protein SE39b [Nicotiana tabacum]	281	2.00E-79	1037	3	smart00205
40	gi|30424403|emb|CAD56465.1| putative class 5 chitinase [Medicago truncatula]	219	1.00E-55	867	3	smart00636, pfam00704
41	gi|15221033|ref|NP_173261.1| thaumatin, putative [Arabidopsis thaliana]	399	0	1147	6	smart00205
42	gi|12324220|gb|AAG52086.1| thaumatin-like protein; 9376-10898 [Arabidopsis thaliana]	314	0	858	7	smart00205
43	gi|1236785|emb|CAA57976.1| sts14 [Solanum tuberosum]	175	6.00E-43	507	2	cd00168
44	gi|2501182|sp|Q41350|OLP1_LYCES Osmotin-like protein precursor [Lycopersicon esculentum]	413	0	1133	8	smart00205
45	gi|4510345|gb|AAD21434.1| expressed protein [Arabidopsis thaliana]	166	1.00E-39	811	4	pfam01428
46	gi|5814093|gb|AAD52097.1| receptor-like kinase CHRK1 [Nicotiana tabacum]	250	4.00E-65	812	3	cd00192, smart00636, pfam00704
48	gi|505267|emb|CAA54374.1| chitinase, class V [Nicotiana tabacum]	409	0	1346	15	smart00636, pfam00704
49	gi|50948973|ref|XP_493844.1| putative glutathione S-transferase [Oryza sativa]	115	2.00E-24	781	2	COG0625, cd03055, cd03203
50	gi|60280803|gb|AAX18294.1| major allergen Mal d 1.04 [Malus x domestica]	164	3.00E-39	938	91	pfam00407
51	gi|169363|gb|AAA33773.1| PVPR3 [Phaseolus vulgaris]	139	2.00E-31	910	3	pfam01428
52	gi|60280803|gb|AAX18294.1| major allergen Mal d 1.04 [Malus x domestica]	170	5.00E-41	801	8	pfam00407
53	gi|2677826|gb|AAB97141.1| major allergen protein homolog [Prunus armeniaca]	149	1.00E-34	818	4	pfam00407
54	gi|60280803|gb|AAX18294.1| major allergen Mal d 1.04 [Malus x domestica]	193	6.00E-48	772	2	pfam00407
55	gi|3183991|emb|CAA06414.1| P69F protein [Lycopersicon esculentum]	159	7.00E-38	757	2	cd02120, pfam00082, pfam05922
56	gi|2677826|gb|AAB97141.1| major allergen protein homolog [Prunus armeniaca]	167	5.00E-40	880	18	pfam00407
57	gi|38603816|gb|AAR24653.1| At5g40020 [Arabidopsis thaliana]	148	3.00E-34	890	2	smart00205
58	gi|11385441|gb|AAG34803.1| glutathione S-transferase GST 13 [Glycine max]	248	3.00E-64	1006	5	cd03185, cd03058
59	gi|7106538|dbj|BAA92224.1| similar to PR-10 [Vigna unguiculata]	197	5.00E-49	913	10	pfam00407
60	gi|21595167|gb|AAM66077.1| pathogenesis-related protein-like protein [Arabidopsis thaliana]	114	1.00E-24	515	3	No CD has been identified
61	gi|7270553|emb|CAB81510.1| thaumatin-like protein [Arabidopsis thaliana]	343	6.00E-93	870	4	smart00205
62	gi|60280803|gb|AAX18294.1| major allergen Mal d 1.04 [Malus x domestica]	164	4.00E-39	925	13	pfam00407
63	gi|12323299|gb|AAG51631.1| thaumatin-like protein; 12104-13574 [Arabidopsis thaliana]	261	2.00E-68	765	2	smart00205

Table S5: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Chitinase, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Chitinase						
Contig	BLAST NR	Score	e-value	Length	Reads	Conserved Domains
1	gi|467689|emb|CAA55128.1| chitinase/lysozyme [Nicotiana tabacum]	284	9.00E-79	808	2	smart00636, pfam00704
2	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	135	9.00E-35	843	2	pfam00704
3	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	320	5.00E-86	1364	56	pfam00704
4	gi|603882|emb|CAA87072.1| pathogenesis-related protein PR-3 type [Sambucus nigra]	362	6.00E-99	804	3	cd00325, cd00035
5	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	231	4.00E-59	1335	288	pfam00704
6	gi|3451147|emb|CAA09110.1| chitinase [Hevea brasiliensis]	357	1.00E-96	1880	58	COG3469
7	gi|11262118|pir||JC7335 chitinase (EC 3.2.1.14) 1 - cone shell (Conus tulipa) [Tulipa bakeri]	344	2.00E-93	851	3	smart00636
8	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	163	6.00E-39	880	19	pfam00704
9	gi|3790355|dbj|BAA33971.1| chitinase 134 [Nicotiana tabacum]	450	0	1007	11	cd00035
10	gi|37051096|dbj|BAC81645.1| class1 chitinase [Pisum sativum]	473	0	1377	48	cd00325
11	gi|3790355|dbj|BAA33971.1| chitinase 134 [Nicotiana tabacum]	457	0	1050	19	cd00035
12	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	194	4.00E-48	922	35	pfam00704
13	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	224	9.00E-57	1624	62	pfam00704
14	gi|10954033|gb|AAG25709.1| class III acidic chitinase [Malus x domestica]	273	5.00E-72	729	3	COG3469
15	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	206	7.00E-52	829	10	pfam00704
16	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	335	2.00E-90	1499	282	pfam00704
17	gi|1359600|emb|CAA64868.1| chitinase Ib [Castanea sativa]	457	0	1178	11	cd00325
18	gi|167539|gb|AAC37395.1| chitinase [Cucumis sativus]	162	2.00E-38	821	6	COG3469, pfam00704
19	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	339	1.00E-91	1295	155	pfam00704
20	gi|33414050|gb|AAP03087.1| class Ib chitinase [Galega orientalis]	285	2.00E-85	866	5	cd00325, cd00035
22	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	219	1.00E-55	932	4	pfam00704
23	gi|167540|gb|AAC37396.1| ORF 3 [Cucumis sativus]	222	2.00E-56	1142	7	COG3469
24	gi|4835586|dbj|BAA77677.1| acidic chitinase [Glycine max]	268	2.00E-70	1072	8	COG3469
25	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	152	1.00E-35	892	5	pfam00704
26	gi|30424403|emb|CAD56465.1| putative class 5 chitinase [Medicago truncatula]	219	1.00E-55	867	3	smart00636, pfam00704
27	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	228	4.00E-58	1391	136	pfam00704
28	gi|3451147|emb|CAA09110.1| chitinase [Hevea brasiliensis]	345	1.00E-93	994	2	COG3469
29	gi|34908492|ref|NP_915593.1| putative beta-1,3-glucanase [Oryza sativa]	400	0	1068	3	pfam00332
30	gi|762879|dbj|BAA08708.1| chitinase [Psophocarpus tetragonolobus]	155	7.00E-37	619	3	COG3469
31	gi|37051096|dbj|BAC81645.1| class1 chitinase [Pisum sativum]	439	0	1130	12	cd00325
32	gi|167538|gb|AAC37394.1| ORF 1 [Cucumis sativus]	160	2.00E-38	681	3	COG3469
33	gi|5814093|gb|AAD52097.1| receptor-like kinase CHRK1 [Nicotiana tabacum]	250	4.00E-65	812	3	cd00192, smart00636, pfam00704
34	gi|7434972|pir||JE0184 chitinase (EC 3.2.1.14) 2 - cone shell (Conus tulipa) [Tulipa bakeri]	382	0	1071	4	smart00636
35	gi|62999433|gb|AAY25165.1| beta-1,3-glucanase 1 [Ziziphus jujuba]	442	0	932	2	pfam00332
36	gi|505267|emb|CAA54374.1| chitinase, class V [Nicotiana tabacum]	409	0	1346	15	smart00636, pfam00704
37	gi|3790355|dbj|BAA33971.1| chitinase 134 [Nicotiana tabacum]	392	0	859	19	cd00035
38	gi|19775|emb|CAA77656.1| acidic chitinase III [Nicotiana tabacum]	377	0	1244	12	COG3469
40	gi|10954033|gb|AAG25709.1| class III acidic chitinase [Malus x domestica]	239	1.00E-61	1064	12	COG3469
41	gi|15553476|gb|AAL01886.1| chitinase 3-like protein precursor [Trichosanthes kirilowii]	223	1.00E-56	1113	6	pfam00704
42	gi|4835586|dbj|BAA77677.1| acidic chitinase [Glycine max]	160	4.00E-38	760	2	COG3469
43	gi|3790355|dbj|BAA33971.1| chitinase 134 [Nicotiana tabacum]	414	0	936	15	cd00035
44	gi|21592759|gb|AAM64708.1| unknown [Arabidopsis thaliana]	177	3.00E-52	1030	3	cd00030
45	gi|432580|gb|AAB28479.1| acidic class III chitinase SE2 [Beta vulgaris]	389	0	1067	21	COG3469
46	gi|2934696|dbj|BAA25015.1| class III acidic endochitinase [Glycine max]	111	1.00E-23	581	2	COG3469
47	gi|4835586|dbj|BAA77677.1| acidic chitinase [Glycine max]	270	2.00E-70	2662	421	COG3469

Table S6: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Cytochrome P450, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Cytochrome P450

Contig	BLAST NR	Score	e-value	Length	Reads	Conserved Domains
1	gi|147857238|emb|CAN83490.1| hypothetical protein [Vitis vinifera]	394	0	1087	2	pfam00067
2	gi|147860148|emb|CAN78724.1| hypothetical protein [Vitis vinifera]	204	3.00E-51	564	2	pfam00067
3	gi|15229044|ref|NP_190458.1| ATMP2 [Arabidopsis thaliana]	169	1.00E-40	663	3	pfam00173
4	gi|147786937|emb|CAN60080.1| hypothetical protein [Vitis vinifera]	292	2.00E-77	974	3	pfam00067
5	gi|147812439|emb|CAN65790.1| hypothetical protein [Vitis vinifera]	249	1.00E-64	806	5	pfam00067
7	gi|147864286|emb|CAN83013.1| hypothetical protein [Vitis vinifera]	167	4.00E-40	613	6	pfam00067
8	gi|125574606|gb|EAZ15890.1| hypothetical protein OsJ_030099 [Oryza sativa]	208	3.00E-52	813	2	COG2124
9	gi|147777974|emb|CAN74205.1| hypothetical protein [Vitis vinifera]	376	0	1120	4	pfam00067
10	gi|85068668|gb|ABC69414.1| CYP72A57 [Nicotiana tabacum]	400	0	1081	4	pfam00067
11	gi|92893043|gb|ABE91557.1| E-class P450, group I [Medicago truncatula]	446	0	1487	15	pfam00067
12	gi|21536603|gb|AAM60935.1| unknown [Arabidopsis thaliana]	354	4.00E-96	1033	3	No CD has been identified
13	gi|147801850|emb|CAN75347.1| hypothetical protein [Vitis vinifera]	488	0	878	4	pfam00067
14	gi|84514175|gb|ABC59096.1| cytochrome P450 monooxygenase CYP97C10 [Medicago truncatula].	272	7.00E-98	908	3	pfam00067
16	gi|2695711|emb|CAA04703.1| cytochome b5 [Olea europaea]	126	1.00E-27	740	3	pfam00173
17	gi|85068594|gb|ABC69377.1| CYP81B2v2 [Nicotiana tabacum]	295	2.00E-78	688	4	pfam00067
18	gi|147864456|emb|CAN80496.1| hypothetical protein [Vitis vinifera]	282	2.00E-74	904	2	pfam00067
19	gi|63029720|gb|AAY27751.1| allene oxide synthase [Hevea brasiliensis]	275	3.00E-72	935	5	pfam00067
20	gi|147858518|emb|CAN81014.1| hypothetical protein [Vitis vinifera]	171	7.00E-44	739	3	pfam00067
21	gi|18420515|ref|NP_568068.1| electron transport SCO1/SenC family protein [Arabidopsis thaliana]	261	7.00E-68	1156	3	cd02968
22	gi|147784145|emb|CAN72302.1| hypothetical protein [Vitis vinifera]	744	0	2395	14	pfam00067
23	gi|92870560|gb|ABE79844.1| E-class P450, group I [Medicago truncatula]	191	3.00E-47	664	2	pfam00067
24	gi|12321297|gb|AAG50718.1|AC079041_11 cytochrome P450, putative [Arabidopsis thaliana]	844	0	1757	5	pfam00067
25	gi|77461113|ref|YP_350620.1| Cytochrome bd ubiquinol oxidase, subunit II [Pseudomonas fluorescens PfO-1]	226	9.00E-58	848	13	pfam02322, COG1294
26	gi|126445639|ref|YP_001079231.1| bacterioferritin [Burkholderia mallei NCTC 10247]	280	6.00E-74	794	2	cd00907
27	gi|27764531|gb|AAO23063.1| ent-kaurenoic acid oxidase [Pisum sativum]	266	6.00E-70	751	3	COG1294
28	gi|85068674|gb|ABC69417.1| CYP72A54 [Nicotiana tabacum]	723	0	1653	7	pfam00067
29	gi|33603796|ref|NP_891356.1| putative cytochrome oxidase assembly protein [Bordetella bronchiseptica RB50]	268	2.00E-70	745	2	COG1612
30	gi|115453349|ref|NP_001050275.1| Os03g0390400 [Oryza sativa]	156	7.00E-37	623	3	cd00926
31	gi|110433184|gb|ABG74350.1| cytochrome P450 [Capsicum chinense]	714	0	1639	14	pfam00067
32	gi|145337333|ref|NP_177109.3| oxygen binding [Arabidopsis thaliana]	529	0	1057	9	pfam00067
33	gi|18203445|sp|Q9SPB3|RL10_VITRI 60S ribosomal protein L10 (QM protein homolog) [Vitis riparia]	443	0	969	40	PRK04199
34	gi|147852119|emb|CAN80156.1| hypothetical protein [Vitis vinifera]	360	7.00E-98	1030	2	pfam00067
35	gi|50346814|ref|YP_053185.1| cytochrome b6/f complex subunit IV [Nymphaea alba]	327	6.00E-88	957	5	cd00290
36	gi|15235720|ref|NP_195496.1| cytochrome c oxidase-related [Arabidopsis thaliana]	125	2.00E-27	657	8	cd00925
37	gi|121998106|ref|YP_001002893.1| cytochrome B561 [Halorhodospira halophila SL1]	87	6.00E-22	722	2	COG3658
38	gi|147838607|emb|CAN69513.1| hypothetical protein [Vitis vinifera]	239	1.00E-61	738	2	pfam00067
39	gi|5915836|sp|P93531|C71D7_SOLCH Cytochrome P450 71D7 [Solanum chacoense]	228	3.00E-58	681	3	pfam00067
40	gi|85068684|gb|ABC69422.1| CYP72A58 [Nicotiana tabacum]	674	0	1740	4	pfam00067
41	gi|115437178|ref|NP_001043230.1| Os01g0527700 [Oryza sativa]	490	0	1559	7	COG0109
42	gi|85068674|gb|ABC69417.1| CYP72A54 [Nicotiana tabacum]	490	0	2252	7	pfam00067
44	gi|85001719|gb|ABC68413.1| cytochrome P450 monooxygenase CYP76E3 [Glycine max]	238	4.00E-61	948	3	pfam00067
45	gi|115480934|ref|NP_001064060.1| Os10g0118800 [Oryza sativa]	159	5.00E-37	1405	4	pfam03188
46	gi|92867559|gb|ABE77793.1| E-class P450, group I [Medicago truncatula]	339	5.00E-94	945	6	pfam00067
47	gi|1345882|sp|P49098|CYB5_TOBAC Cytochrome b5 [Nicotiana tabacum]	242	2.00E-62	1000	11	pfam00173
48	gi|147777491|emb|CAN62730.1| hypothetical protein [Vitis vinifera]	304	4.00E-81	818	9	pfam00067
49	gi|85068656|gb|ABC69408.1| CYP71AT2v1 [Nicotiana tabacum]	303	3.00E-81	1182	8	pfam00067
51	gi|33601781|ref|NP_889341.1| Putative heme export protein [Bordetella bronchiseptica RB50]	281	1.00E-74	508	3	COG0755
52	gi|18252325|gb|AAL66194.1|AF386512_1 cytochrome P450 [Pyrus communis]	221	3.00E-56	897	2	pfam00067
53	gi|147812439|emb|CAN65790.1| hypothetical protein [Vitis vinifera]	718	0	1820	17	pfam00067
54	gi|18394631|ref|NP_564058.1| rhomboid family protein [Arabidopsis thaliana]	180	1.00E-43	879	3	pfam01694
55	gi|115466706|ref|NP_001056952.1| Os06g0175900 [Oryza sativa]	122	9.00E-27	595	6	No CD has been identified
56	gi|147778583|emb|CAN60309.1| hypothetical protein [Vitis vinifera]	283	1.00E-74	797	5	pfam00067
57	gi|147782531|emb|CAN68429.1| hypothetical protein [Vitis vinifera]	171	3.00E-41	734	2	pfam00067
58	gi|147794657|emb|CAN73509.1| hypothetical protein [Vitis vinifera]	146	1.00E-33	733	2	pfam00067
59	gi|33593694|ref|NP_881338.1| cytochrome ubiquinol oxidase subunit I [Bordetella pertussis Tohama I]	305	6.00E-82	538	4	cd01662, COG0843
60	gi|139538863|gb|ABO77958.1| p-coumaroyl quinate/shikimate 3'-hydroxylase [Coffea canephora]	1034	0	1877	17	pfam00067
61	gi|18252325|gb|AAL66194.1|AF386512_1 cytochrome P450 [Pyrus communis]	162	1.00E-38	588	2	pfam00067
63	gi|1351206|sp|P48522|TCMO_CATRO Trans-cinnamate 4-monooxygenase [Catharanthus roseus]	490	0	1011	12	pfam00067
64	gi|77457812|ref|YP_347317.1| Cytochrome c-type biogenesis protein [Pseudomonas fluorescens PfO-1]	346	9.00E-94	810	10	cd00189, COG4235
65	gi|130270|sp|P00290|PLAS_LACSA Plastocyanin [Lactuca sativa]	177	2.00E-43	582	3	pfam00127
66	gi|147821410|emb|CAN63505.1| hypothetical protein [Vitis vinifera]	135	2.00E-30	710	2	pfam00067
67	gi|71897452|gb|AAZ52550.1| cytochrome P450 [Capsicum annuum]	288	2.00E-76	980	4	pfam00067
68	gi|147833535|emb|CAN68303.1| hypothetical protein [Vitis vinifera]	333	7.00E-90	763	9	pfam00067
69	gi|77461691|ref|YP_351198.1| Cytochrome c, class I [Pseudomonas fluorenscens]	161	1.00E-38	491	21	COG0672, COG2010
70	gi|70730201|ref|YP_259940.1| cytochrome c oxidase, cbb3-type, subunit I [Pseudomonas	441	0	649	2	cd01661

	fluorescens Pf-5]					
71	gi|129593742|gb|ABO31111.1| sphingolipid delta-8 desaturase [Nicotiana tabacum]	798	0	1737	9	cd03506, pfam00173, cd03507
73	gi|147844260|emb|CAN80040.1| hypothetical protein [Vitis vinifera]	697	0	1486	5	pfam00067
75	gi|2695711|emb|CAA04703.1| cytochome b5 [Olea europaea]	244	5.00E-63	830	14	pfam00173
76	gi|147855053|emb|CAN82364.1| hypothetical protein [Vitis vinifera]	159	2.00E-37	722	7	pfam00067
77	gi|147866210|emb|CAN79423.1| hypothetical protein [Vitis vinifera]	313	4.00E-84	686	2	pfam00067
78	gi|18203445|sp|Q9SPB3|RL10_VITRI 60S ribosomal protein L10 (QM protein homolog) [Vitis riparia]	442	0	964	25	PRK04199
79	gi|147777975|emb|CAN74206.1| hypothetical protein [Vitis vinifera]	274	3.00E-72	695	3	pfam00067
80	gi|1351357|sp|P48502|UCR6_SOLTU Ubiquinol-cytochrome c reductase complex 14 kDa protein (CR14) [Solanum tuberosum]	155	3.00E-36	923	17	pfam02271
81	gi|85068668|gb|ABC69414.1| CYP72A57 [Nicotiana tabacum]	328	3.00E-88	873	5	pfam00067
82	gi|30678524|ref|NP_186877.2| flavodoxin family protein [Arabidopsis thaliana]	498	0	1134	4	pfam00258, pfam00175, COG0369
83	gi|147841765|emb|CAN62209.1| hypothetical protein [Vitis vinifera]	205	4.00E-51	1388	31	pfam00067
87	gi|1040729|emb|CAA60621.1| delta-8 sphingolipid desaturase [Helianthus annuus]	694	0	1910	7	cd03506, pfam00173, cd03511
85	gi|147776011|emb|CAN71371.1| hypothetical protein [Vitis vinifera]	290	5.00E-77	739	5	pfam00067
86	gi|147767047|emb|CAN67678.1| hypothetical protein [Vitis vinifera]	555	0	2037	16	pfam00067
141	gi|115484331|ref|NP_001065827.1| Os11g0162200 [Oryza sativa]	115	3.00E-24	796	6	pfam02320
142	gi|147795107|emb|CAN60851.1| hypothetical protein [Vitis vinifera]	286	1.00E-75	914	3	pfam00067
143	gi|1351357|sp|P48502|UCR6_SOLTU Ubiquinol-cytochrome c reductase complex 14 kDa protein (CR14) [Solanum tuberosum]	152	2.00E-35	840	6	pfam02271
144	gi|92872915|gb|ABE81444.1| E-class P450, group I [Medicago truncatula]	278	2.00E-73	819	4	pfam00067
145	gi|66044163|ref|YP_234004.1| FMN-dependent alpha-hydroxy acid dehydrogenase [Pseudomonas syringae pv. syringae B728a]	223	2.00E-57	373	9	cd02809, PRK11197
146	gi|77460775|ref|YP_350282.1| Periplasmic Sensor Signal Transduction Histidine Kinase [Pseudomonas fluorescens PfO-1]	305	2.00E-81	829	8	cd00075, COG4191
147	gi|15241121|ref|NP_200420.1| COX15 (CYTOCHROME C OXIDASE 15) [Arabidopsis thaliana]	137	6.00E-31	868	2	COG1612
149	gi|18420859|ref|NP_568463.1| CYP714A1 (cytochrome P450, family 714, subfamily A, polypeptide 1); oxygen binding [Arabidopsis thaliana]	236	8.00E-61	680	3	pfam00067
150	gi|147821972|emb|CAN77159.1| hypothetical protein [Vitis vinifera]	283	0	1738	8	pfam00067
151	gi|147811196|emb|CAN70159.1| hypothetical protein [Vitis vinifera]	269	2.00E-73	851	3	pfam00067
152	gi|22330642|ref|NP_177656.2| flavodoxin family protein / radical SAM domain-containing protein [Arabidopsis thaliana]	204	3.00E-51	676	2	pfam08608, pfam00258, pfam04055, COG0731
153	gi|124484371|dbj|BAF46296.1| cytochrome P450 [Ipomoea nil]	285	2.00E-75	749	3	pfam00067
154	gi|147840876|emb|CAN73186.1| hypothetical protein [Vitis vinifera]	293	2.00E-77	1260	20	pfam00067
155	gi|145323848|ref|NP_001077513.1| unknown protein [Arabidopsis thaliana]	121	2.00E-26	565	3	No CD has been identified
156	gi|147765656|emb|CAN71501.1| hypothetical protein [Vitis vinifera]	727	0	2096	15	pfam00067
157	gi|1619602|emb|CAA69976.1| MtN3 [Medicago truncatula]	237	5.00E-61	731	2	pfam03083
158	gi|92885617|gb|ABE87844.1| E-class P450, group I [Medicago truncatula]	265	1.00E-85	1055	2	pfam00067
160	gi|147786939|emb|CAN60082.1| hypothetical protein [Vitis vinifera]	264	4.00E-69	745	3	pfam00067
161	gi|85068648|gb|ABC69404.1| CYP71D51v2 [Nicotiana tabacum]	437	0	1575	4	pfam00067
162	gi|46403211|gb|AAS92625.1| coniferylalcohol 5-hydroxylase [Centaurium erythraea]	353	5.00E-96	781	2	pfam00067
163	gi|147843271|emb|CAN80536.1| hypothetical protein [Vitis vinifera]	509	0	1420	5	pfam00067
164	gi|147835182|emb|CAN76753.1| hypothetical protein [Vitis vinifera]	503	0	1213	5	pfam00067
165	gi|48428165|sp|Q9LZQ0|CX5C2_ARATH Cytochrome c oxidase polypeptide Vc-2 (Cytochrome c oxidase subunit 5c-2) [Arabidopsis thaliana]	107	5.00E-22	591	2	pfam05799
166	gi|5915840|sp|O22307|C71DB_LOTJA Cytochrome P450 71D11 [Lotus japonicus]	230	6.00E-59	688	2	pfam00067
167	gi|77459567|ref|YP_349074.1| Transcriptional Regulator, Crp/Fnr family [Pseudomonas fluorescens PfO-1]	255	9.00E-67	603	2	cd00038, cd00092, PRK11161
168	gi|82570227|gb|ABB83676.1| putative p-coumaroyl 3'-hydroxylase CYP98A-C1 [Coffea canephora]	503	0	915	6	pfam00067
169	gi|12331298|emb|CAC24711.1| cytochrome P450 [Solanum tuberosum]	546	0	1655	20	pfam00067
170	gi|37954114|gb|AAP69988.1| ent-kaurene oxidase [Pisum sativum]	592	0	2462	30	pfam00067
171	gi|115446391|ref|NP_001046975.1| Os02g0520800 [Oryza sativa]	272	2.00E-71	773	4	cd03470
173	gi|587562|emb|CAA56520.1| mitochondrial processing peptidase [Solanum tuberosum]	273	1.00E-71	887	5	pfam00675, pfam05193, COG0612
174	gi|147765747|emb|CAN60189.1| hypothetical protein [Vitis vinifera]	356	2.00E-96	1104	4	pfam00067
175	gi|15241121|ref|NP_200420.1| COX15 (CYTOCHROME C OXIDASE 15) [Arabidopsis thaliana]	221	1.00E-61	671	3	COG1612
176	gi|139538863|gb|ABO77958.1| p-coumaroyl quinate/shikimate 3'-hydroxylase [Coffea canephora]	390	0	894	3	pfam00067
177	gi|147772136|emb|CAN75686.1| hypothetical protein [Vitis vinifera]	189	9.00E-47	697	2	pfam00067
178	gi|15982240|emb|CAC91565.1| hydroperoxide lyase [Nicotiana attenuata]	298	1.00E-79	727	2	pfam00067
179	gi|27764531|gb|AAO23063.1| ent-kaurenoic acid oxidase [Pisum sativum]	321	3.00E-86	805	3	COG2124
180	gi|85068684|gb|ABC69422.1| CYP72A58 [Nicotiana tabacum]	660	0	1546	11	pfam00067
181	gi|147774515|emb|CAN76784.1| hypothetical protein [Vitis vinifera]	225	2.00E-57	720	2	pfam00067
182	gi|147795107|emb|CAN60851.1| hypothetical protein [Vitis vinifera]	139	2.00E-31	829	2	pfam00067
183	gi|126355969|ref|ZP_01712975.1| Amine dehydrogenase [Pseudomonas putida]	267	4.00E-70	787	8	pfam06433
184	gi|77460158|ref|YP_349665.1| thiol:disulfide interchange protein DsbE, putative [Pseudomonas fluorescens PfO-1]	246	2.00E-64	458	13	cd02966, pfam01790
185	gi|147842082|emb|CAN62647.1| hypothetical protein [Vitis vinifera]	200	8.00E-50	781	3	pfam00067
186	gi|71081902|gb|AAZ23260.1| cytochrome P450 monooxygenase [Nicotiana tabacum]	115	1.00E-24	595	3	COG2124
187	gi|118022|sp|P00054|CYC_SESIN Cytochrome c [Sesamum indicum]	220	1.00E-55	984	18	COG3474
188	gi|85068648|gb|ABC69404.1| CYP71D51v2 [Nicotiana tabacum]	291	6.00E-77	1163	2	pfam00067
189	gi|27754697|gb|AAO22792.1| putative cytochrome c oxidoreductase [Arabidopsis thaliana]	268	2.00E-70	792	6	pfam03083
191	gi|119897307|ref|YP_932520.1| L-lactate dehydrogenase [Azoarcus sp. BH72]	316	7.00E-85	742	2	cd03332, pfam01070
192	gi|26106071|dbj|BAC41516.1| NADPH-cytochrome P-450 reductase [Ophiorrhiza pumila]	1078	0	2692	22	pfam00258, pfam00175, COG0369
193	gi|85068672|gb|ABC69416.1| CYP71AU1 [Nicotiana tabacum]	422	0	1186	3	pfam00067
194	gi|63029720|gb|AAY27751.1| allene oxide synthase [Hevea brasiliensis]	657	0	2020	31	pfam00067
195	gi|147858518|emb|CAN81014.1| hypothetical protein [Vitis vinifera]	506	0	1851	14	pfam00067
196	gi|15219780|ref|NP_176265.1| B5 #5 (cytochrome b5 family protein #5) [Arabidopsis thaliana]	143	1.00E-32	862	2	pfam00173
197	gi|147794492|emb|CAN62759.1| hypothetical protein [Vitis vinifera]	115	1.00E-24	620	7	pfam00067

198	gi|27529728|dbj|BAC53893.1| cytochrome P450 [Petunia x hybrida]	298	0	1645	6	pfam00067
199	gi|157357203|emb|CAO63697.1|[157357203] unnamed protein product [Vitis vinifera]	537	1.00E-150	1704	50	pfam00067
204	gi|115265663|dbj|BAF32924.1| cytochrome bd ubiquinol oxidase, subunit II [Pseudomonas syringae pv. phaseolicola]	228	8.00E-59	520	5	pfam02322, COG1294
211	gi|147801862|emb|CAN74977.1| hypothetical protein [Vitis vinifera]	110	5.00E-23	565	2	pfam00067
220	gi|84578869|dbj|BAE72877.1|[84578869] cytochrome P450 [Verbena x hybrida]	231	3.00E-61	781	2	pfam00067
221	gi|14423327|gb|AAK62346.1|[14423327] elicitor-inducible cytochrome P450 [Nicotiana tabacum]	353	1.00E-100	857	3	pfam00067
225	gi|147769656|emb|CAN74644.1|[147769656] hypothetical protein [Vitis vinifera]	295	2.00E-78	740	2	pfam00067
227	gi|147742913|sp|P84887|AAUA_ALCFA Aralkylamine dehydrogenase light chain precursor (Aromatic amine dehydrogenase) [Alcaligenes faecalis]	215	2.00E-54	681	2	pfam02975
231	gi|147828267|emb|CAN75404.1|[147828267] hypothetical protein [Vitis vinifera]	238	4.00E-68	883	5	pfam00067
232	gi|85068612|gb|ABC69386.1|[85068612] CYP92B2v1 [Nicotiana tabacum]	236	6.00E-61	690	2	pfam00067
234	gi|17065916|emb|CAC80883.1| geraniol 10-hydroxylase [Catharanthus roseus]	172	6.00E-64	595	2	pfam00067
235	gi|125580889|gb|EAZ21820.1|[125580889] hypothetical protein OsJ_005303 [Oryza sativa]	299	1.00E-79	780	2	cd02968

Table S7: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Glucanase, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Glucanase

Contig	BLAST NR	Score	e-value	Length	Reads	Conserved Domains
1	gi|91107165|gb|ABE11608.1| xyloglucan endo-transglycosylase precursor [Solanum chacoense]	212	1.00E-53	665	11	cd02176, pfam06955
2	gi|92868701|gb|ABE78694.1| Glycoside hydrolase, family 17; Virulence factor, pectin lyase fold [Medicago truncatula]	203	5.00E-51	774	3	pfam07983, pfam00332
3	gi|94442926|emb|CAJ91137.1| beta-1,3-glucanase [Platanus x acerifolia]	226	6.00E-58	802	2	pfam00332
4	gi|34909360|ref|NP_916027.1| P0638D12.12 [Oryza sativa (japonica cultivar-group)]	213	1.00E-62	848	4	pfam00332
5	gi|92895371|gb|ABE92687.1| Glycoside transferase, six-hairpin, subgroup [Medicago truncatula]	578	0	1493	3	pfam00759
6	gi|62362434|gb|AAX81588.1| nectarin IV [Nicotiana langsdorffii x Nicotiana sanderae]	392	0	1707	11	No CD has been identified
7	gi|29826242|gb|AAO91861.1| TGB12K interacting protein 2 [Nicotiana tabacum]	311	3.00E-83	1054	13	cd00204
8	gi|3582436|dbj|BAA33065.1| beta-D-glucan exohydrolase [Nicotiana tabacum]	324	5.00E-87	1330	5	COG1472, pfam01915
9	gi|14279169|gb|AAK58515.1| beta-1,3-glucanase-like protein [Olea europaea]	296	2.00E-78	999	2	pfam07983, pfam00332
10	gi|4662638|gb|AAD26909.1| putative beta-1,3-glucanase [Arabidopsis thaliana]	296	1.00E-79	839	6	pfam07983, pfam00332
11	gi|91107165|gb|ABE11608.1| xyloglucan endo-transglycosylase precursor [Solanum chacoense]	483	0	1081	14	cd02176, pfam06955
12	gi|34329342|gb|AAQ63883.1| cellulase [Medicago truncatula]	158	5.00E-37	1372	3	pfam00759
13	gi|33324614|gb|AAQ08018.1| endo-1,4-beta-glucanase [Gossypium hirsutum]	530	0	1111	3	pfam00759
14	gi|56785404|dbj|BAD82640.1| putative elicitor inducible beta-1,3-glucanase NtEIG-E76 [Oryza sativa (japonica cultivar-group)]	321	3.00E-86	839	3	pfam00332
15	gi|9294513|dbj|BAB02775.1| unnamed protein product [Arabidopsis thaliana]	156	1.00E-36	1079	3	COG0412
16	gi|15228047|ref|NP_181224.1| hydrolase, acting on glycosyl bonds / hydrolase, hydrolyzing O-glycosyl compounds [Arabidopsis thaliana]	115	6.00E-48	798	6	cd02176, pfam06955
17	gi|2244740|dbj|BAA21111.1| endo-1,4-beta-glucanase [Gossypium hirsutum]	385	0	1377	11	pfam00759
18	gi|55740509|gb|AAV63847.1| hypothetical protein At1g29380 [Arabidopsis thaliana]	141	3.00E-32	811	2	pfam07983
19	gi|8886865|gb|AAF80590.1| xyloglucan endotransglycosylase XET1 [Asparagus officinalis]	134	2.00E-62	656	2	cd02176, pfam06955
20	gi|55978791|gb|AAV68857.1| hypothetical protein AT1G79480 [Arabidopsis thaliana]	135	3.00E-30	876	2	pfam07983
21	gi|15232590|ref|NP_190241.1| hydrolase, hydrolyzing O-glycosyl compounds [Arabidopsis thaliana]	203	5.00E-51	717	2	cd02176, pfam06955
22	gi|50924081|ref|XP_472401.1| OSJNBa0073L04.8 [Oryza sativa (japonica cultivar-group)]	275	0	1392	2	pfam00332
23	gi|62362438|gb|AAX81590.1| beta-1,3-glucanase [Fragaria x ananassa]	377	0	1469	5	pfam00332
24	gi|11071974|dbj|BAB17320.1| elicitor inducible beta-1,3-glucanase NtEIG-E76 [Nicotiana tabacum]	682	0	1783	11	pfam07983, pfam00332
25	gi|15225764|ref|NP_180858.1| hydrolase, hydrolyzing O-glycosyl compounds [Arabidopsis thaliana]	298	2.00E-79	955	2	cd02176, pfam06955
26	gi|57899486|dbj|BAD86947.1| putative elicitor inducible beta-1,3-glucanase NtEIG-E76 [Oryza sativa (japonica cultivar-group)]	342	9.00E-93	759	2	pfam07983, pfam00332
27	gi|94442926|emb|CAJ91137.1| beta-1,3-glucanase [Platanus x acerifolia]	278	1.00E-73	797	4	pfam00332
28	gi|13560781|gb|AAK30204.1| endoxyloglucan transferase [Daucus carota]	507	0	1282	4	cd02176, pfam06955
29	gi|1885310|emb|CAA62847.1| Endoxyloglucan transferase (EXT) [Hordeum vulgare subsp. vulgare]	140	2.00E-63	1365	4	cd02176, pfam06955
30	gi|33391721|gb|AAQ17461.1| beta-D-glucosidase [Gossypium hirsutum]	591	0	1274	6	COG1472, pfam01915
31	gi|14029149|gb|AAK51119.1| xyloglucan endo-transglycosylase [Carica papaya]	397	0	869	4	cd02176, pfam06955
32	gi|18391291|ref|NP_563892.1| hydrolase, acting on glycosyl bonds / hydrolyzing O-glycosyl compounds [Arabidopsis thaliana]	308	0	885	6	cd02176, pfam06955
33	gi|70927645|gb|AAZ15705.1| endo-alpha-1,4-glucanase [Gossypium hirsutum]	914	0	2119	12	COG1472, pfam01915
34	gi|19911573|dbj|BAB86890.1| syringolide-induced protein 19-1-5 [Glycine max]	290	9.00E-92	747	2	cd02176, pfam06955
35	gi|19911573|dbj|BAB86890.1| syringolide-induced protein 19-1-5 [Glycine max]	466	0	1173	13	cd02176, pfam06955
36	gi|92891730|gb|ABE90942.1| putative glycosyl hydrolase family 17 protein [Medicago truncatula]	163	1.00E-38	951	10	pfam07983
37	gi|90186655|gb|ABD91577.1| beta-1,3-glucanase [Medicago sativa]	218	1.00E-55	707	2	pfam00332
38	gi|18403812|ref|NP_566731.1| hydrolase [Arabidopsis thaliana]	229	2.00E-58	940	5	COG0412
39	gi|37223500|gb|AAQ90287.1| beta-1,3-glucanase, acidic [Coffea arabica]	253	9.00E-66	819	2	pfam07983, pfam00332
40	gi|34908492|ref|NP_915593.1| putative beta-1,3-glucanase [Oryza sativa]	481	0	1549	13	pfam00332
41	gi|37223498|gb|AAQ90286.1| beta-1,3-glucanase, basic [Coffea arabica]	585	0	1412	19	pfam00332
43	gi|92885855|gb|ABE87944.1| Glycosyl hydrolases family 17 [Medicago truncatula]	281	1.00E-74	749	3	pfam07983, pfam00332
44	gi|15235714|ref|NP_195494.1| hydrolase, acting on glycosyl bonds / hydrolyzing O-glycosyl compounds [Arabidopsis thaliana]	101	1.00E-39	1255	5	cd02176, pfam06955
45	gi|90186653|gb|ABD91576.1| beta-1,3-glucanase [Medicago sativa]	193	4.00E-48	656	3	pfam00332
47	gi|29243202|dbj|BAC66186.1| beta-1,3-glucanase [Fragaria x ananassa]	316	1.00E-84	976	4	pfam00332
48	gi|2065531|gb|AAC49704.1| endo-1,4-beta-glucanase [Lycopersicon esculentum]	368	0	1062	5	pfam00759
49	gi|42795460|gb|AAS46240.1| xyloglucan endotransglucosylase-hydrolase XTH5 [Lycopersicon esculentum]	487	0	1300	16	cd02176, pfam06955
50	gi|108710308|gb|ABF98103.1| Glucan endo-1,3-beta-glucosidase 7 precursor, putative, expressed [Oryza sativa (japonica cultivar-group)]	348	4.00E-94	1665	6	pfam07983, pfam00332
51	gi|42795462|gb|AAS46241.1| xyloglucan endotransglucosylase-hydrolase XTH3 [Lycopersicon esculentum]	389	0	1136	12	cd02176, pfam06955
52	gi|285741|dbj|BAA03413.1| EDGP precursor [Daucus carota]	584	0	1567	12	No CD has been identified
53	gi|2290683|gb|AAB65156.1| basic cellulase [Citrus sinensis]	401	0	888	3	pfam00759
54	gi|37223498|gb|AAQ90286.1| beta-1,3-glucanase, basic [Coffea arabica]	641	0	1249	13	pfam00332
55	gi|3582436|dbj|BAA33065.1| beta-D-glucan exohydrolase [Nicotiana tabacum]	285	1.00E-83	636	2	COG1472, pfam01915
56	gi|30678225|ref|NP_178637.2| hydrolase, hydrolyzing O-glycosyl compounds [Arabidopsis thaliana]	308	1.00E-82	697	2	cd02176, pfam06955
57	gi|37223500|gb|AAQ90287.1| beta-1,3-glucanase, acidic [Coffea arabica]	428	0	1425	3	pfam07983, pfam00332
58	gi|15232711|ref|NP_190288.1| hydrolase, hydrolyzing O-glycosyl compounds [Arabidopsis thaliana]	291	3.00E-77	1062	4	cd02176, pfam06955
59	gi|92872472|gb|ABE81087.1| Glycoside transferase, six-hairpin, subgroup [Medicago truncatula]	254	2.00E-66	742	2	pfam00759
60	gi|51854423|gb|AAU10802.1| 'putative endo-1,3;1,4-beta-D-glucanase' [Oryza sativa (japonica cultivar-group)]	232	2.00E-59	1078	5	COG0412
61	gi|70780051|gb|AAZ08349.1| xyloglucan endotransglycosylase/hydrolase XTH5 [Lycopersicon esculentum]	448	0	1276	16	cd02176, pfam06955
62	gi|30690053|ref|NP_195174.3| hydrolase, hydrolyzing O-glycosyl compounds [Arabidopsis thaliana]	304	3.00E-81	981	4	cd02176, pfam06955

63	gi|37223500|gb|AAQ90287.1| beta-1,3-glucanase, acidic [Coffea arabica]	171	7.00E-42	449	5	pfam07983, pfam00332
64	gi|18394637|ref|NP_564059.1| unknown protein [Arabidopsis thaliana]	171	3.00E-41	720	2	pfam07983
65	gi|4165132|gb|AAD08699.1| endo-beta-1,4-D-glucanase [Lycopersicon esculentum]	139	4.00E-41	1020	28	pfam00759
66	gi|11071974|dbj|BAB17320.1| elicitor inducible beta-1,3-glucanase NtEIG-E76 [Nicotiana tabacum]	320	4.00E-86	802	4	pfam07983, pfam00332
67	gi|37625031|gb|AAQ96339.1| putative ankyrin-repeat protein [Vitis aestivalis]	248	3.00E-64	866	2	cd00204
68	gi|15232707|ref|NP_190284.1| hydrolase, hydrolyzing O-glycosyl compounds [Arabidopsis thaliana]	491	0	1033	3	cd02176, pfam06955
69	gi|79328626|ref|NP_001031936.1| hydrolase, hydrolyzing O-glycosyl com.pounds [Arabidopsis thaliana]	215	1.00E-68	971	5	cd02176, pfam06955
70	gi|92891730|gb|ABE90942.1| putative glycosyl hydrolase family 17 protein [Medicago truncatula]	161	3.00E-38	916	5	pfam07983
71	gi|577068|emb|CAA58002.1| xyloglycan endo-transglycosylase [Lycopersicon esculentum]	237	1.00E-77	843	3	cd02176, pfam06955
72	gi|22947852|gb|AAN07898.1| xyloglucan endotransglycosylase [Malus x domestica]	360	4.00E-98	758	2	cd02176, pfam06955
73	gi|285741|dbj|BAA03413.1| EDGP precursor [Daucus carota]	299	9.00E-80	904	3	No CD has been identified
75	gi|42795466|gb|AAS46243.1| xyloglucan endotransglucosylase-hydrolase XTH5 [Lycopersicon esculentum]	466	0	1347	17	cd02176, pfam06955
76	gi|7414433|emb|CAB85903.1| beta-1,3 glucanase [Pisum sativum]	519	0	1725	15	pfam00332
77	gi|4165132|gb|AAD08699.1| endo-beta-1,4-D-glucanase [Lycopersicon esculentum]	248	6.00E-69	782	3	pfam00759
78	gi|2230955|emb|CAA72133.1| endo-1,4-beta-D-glucanase [Lycopersicon esculentum]	291	2.00E-77	734	2	pfam00759
79	gi|473102|emb|CAA82271.1| beta-1,3-glucanase [Nicotiana tabacum]	316	5.00E-85	743	2	pfam00332
81	gi|29500899|emb|CAD87533.1| putative xyloglucan endotransglycosylase - hydrolase XTH5 [Lycopersicon esculentum]	330	0	864	2	cd02176, pfam06955
82	gi|87240526|gb|ABD32384.1| Peptidase A1, pepsin [Medicago truncatula]	351	5.00E-95	1393	8	pfam00026
83	gi|70927645|gb|AAZ15705.1| endo-alpha-1,4-glucanase [Gossypium hirsutum]	430	0	1071	7	COG1472, pfam01915
84	gi|18403820|ref|NP_566732.1| hydrolase [Arabidopsis thaliana]	252	2.00E-65	1010	13	COG0412
85	gi|18403820|ref|NP_566732.1| hydrolase [Arabidopsis thaliana]	273	1.00E-71	1171	16	COG0412
86	gi|577068|emb|CAA58002.1| xyloglycan endo-transglycosylase [Lycopersicon esculentum]	379	0	891	2	cd02176, pfam06955
87	gi|4165132|gb|AAD08699.1| endo-beta-1,4-D-glucanase [Lycopersicon esculentum]	635	0	1333	8	pfam00759
88	gi|32482806|gb|AAP84703.1| putative xyloglucanase inhibitor [Solanum tuberosum]	540	0	1599	9	No CD has been identified
89	gi|16903351|gb|AAL30453.1| endo-beta-1,4-glucanase precursor [Nicotiana tabacum]	309	8.00E-83	732	2	pfam00759
90	gi|22331064|ref|NP_566459.2| FUS9 (FUSCA 9); ubiquitin conjugating enzyme [Arabidopsis thaliana]	261	3.00E-68	996	4	cd00195
91	gi|29826242|gb|AAO91861.1| TGB12K interacting protein 2 [Nicotiana tabacum]	482	0	1516	29	cd00204
92	gi|50944207|ref|XP_481631.1| putative glycosyl hydrolase [Oryza sativa]	160	3.00E-38	613	2	pfam00332

Table S8: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project HSP (Heat Shock Protein), and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

HSP (Heat Shock Protein)						
Contig	BLAST NR	Score	e-value	Length	Reads	Conserved Domains
1	gi|19618|emb|CAA41547.1| heat shock protein [Medicago sativa]	128	2.00E-28	721	3	cd00298
2	gi|558606|emb|CAA50022.1| Nthsp18p [Nicotiana tabacum]	219	6.00E-56	781	4	cd00298
3	gi|37704443|gb|AAR01526.1| cytosolic class I small heat shock protein 3B [Nicotiana tabacum]	132	2.00E-30	416	3	cd00298
4	gi|15225187|ref|NP_180771.1| HSP70T-2; ATP binding [Arabidopsis thaliana]	354	2.00E-96	749	2	pfam00012
5	gi|4456758|emb|CAB36910.1| heat shock protein 17.4 [Quercus suber]	233	7.00E-60	782	4	cd00298, COG0071
6	gi|30687816|ref|NP_189160.3| ROF1 (ROTAMASE FKBP 1); FK506 binding/calmodulin binding/peptidyl-prolyl cis-trans isomerase [Arabidopsis thaliana]	818	0	1995	11	cd00189, pfam00254
7	gi|15148884|gb|AAK84869.1| small heat stress protein class CIII [Lycopersicon esculentum]	164	4.00E-39	891	5	cd00298
8	gi|18415982|ref|NP_567663.1| unknown protein [Arabidopsis thaliana]	285	3.00E-75	1627	40	cd00189, smart00727
9	gi|1143427|emb|CAA52149.1| heat shock protein 70 [Cucumis sativus]	1077	0	2608	20	pfam02491, PRK00290
10	gi|1143427|emb|CAA52149.1| heat shock protein 70 [Cucumis sativus]	830	0	1928	21	pfam02491, PRK00290
11	gi|15240308|ref|NP_198583.1| unknown protein [Arabidopsis thaliana]	143	4.00E-33	611	2	cd00298
12	gi|1143427|emb|CAA52149.1| heat shock protein 70 [Cucumis sativus]	375	0	1106	7	pfam02491, PRK00290
13	gi|710434|gb|AAB03097.1| Hsp22.3 [Glycine max]	159	3.00E-38	712	4	cd00298, COG0071
14	gi|76057837|emb|CAH55766.1| peptidyl prolyl cis-trans isomerase [Oryza sativa]	247	3.00E-64	806	2	pfam00254
15	gi|6969976|gb|AAF34134.1| high molecular weight heat shock protein [MMalus x domestica]	351	9.00E-96	640	2	pfam00012
16	gi|14581677|gb|AAK64512.1| Hsp70 interacting protein/thioredoxin chimera [Vitis labrusca]	324	2.00E-87	784	2	cd02947, cd00189
17	gi|5257560|gb|AAD41409.1| cytosolic class II low molecular weight heat shock protein [Prunus dulcis]	214	3.00E-54	776	5	cd00298
18	gi|5302797|emb|CAB46039.1| HSP like protein [Arabidopsis thaliana]	239	7.00E-62	832	3	pfam00012
19	gi|558606|emb|CAA50022.1| Nthsp18p [Nicotiana tabacum]	222	1.00E-56	817	18	cd00298
20	gi|462013|sp|P35016|ENPL_CATRO Endoplasmin homolog precursor (GRP94 homolog) [Catharanthus roseus]	581	0	1244	5	pfam00183, cd00075
21	gi|50923325|ref|XP_472023.1| OSJNBa0091C07.4 [Oryza sativa (japonica cultivar-group)]	518	0	1484	8	pfam00254, cd00189
22	gi|77556324|gb|ABA99120.1| AC009176 putative heat-shock protein [Oryza sativa]	362	9.00E-99	832	2	pfam00183, PRK05218
23	gi|558606|emb|CAA50022.1| Nthsp18p [Nicotiana tabacum]	205	9.00E-52	683	10	cd00298
24	gi|123538|sp|P04793|HSP13_SOYBN 17.5 kDa class I heat shock protein (HSP 17.5-M) [Glycine max]	243	4.00E-63	757	3	cd00298
25	gi|30172151|emb|CAD89783.1| peptidylprolyl cis-trans isomerase [Oryza sativa]	243	5.00E-63	883	4	pfam00254
27	gi|558606|emb|CAA50022.1| Nthsp18p [Nicotiana tabacum]	219	9.00E-56	846	11	cd00298
28	gi|77380807|gb|ABA72320.1| heat shock protein YegD [Pseudomonas fluorenscens]	418	0	749	3	PRK11678
29	gi|710434|gb|AAB03097.1| Hsp22.3 [Glycine max]	159	2.00E-38	670	2	cd00298, COG0071
30	gi|19618|emb|CAA41547.1| heat shock protein [Medicago sativa]	129	9.00E-29	796	2	cd00298
31	gi|51091143|dbj|BAD35839.1| putative cyclophilin-40 [Oryza sativa (japonica cultivar-group)]	207	2.00E-52	729	3	cd01926, smart00028, cd00189

Table S9: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Thaumatin, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Thaumatin						
Contig	**BLAST NR**	**Score**	**e-value**	**Length**	**Reads**	**Conserved Domains**
1	gi|21592749|gb|AAM64698.1| putative thaumatin-like protein [Arabidopsis thaliana]	325	2.00E-87	1592	4	smart00205
2	gi|21592749|gb|AAM64698.1| putative thaumatin-like protein [Arabidopsis thaliana]	304	3.00E-81	911	7	smart00205
3	gi|4586372|dbj|BAA74546.2| thaumatin-like protein SE39b [Nicotiana tabacum]	313	5.00E-84	1011	4	smart00205
4	gi|71057064|emb|CAI38795.2| thaumatin-like protein [Actinidia deliciosa]	400	0	1017	41	pfam00314
5	gi|53830843|gb|AAU95244.1| putative thaumatin-like protein [Solanum tuberosum]	371	0	975	56	pfam00314
6	gi|71057064|emb|CAI38795.2| thaumatin-like protein [Actinidia deliciosa]	404	0	892	15	pfam00314
7	gi|33329390|gb|AAQ10092.1| thaumatin-like protein [Vitis vinifera]	381	0	1125	117	pfam00314
8	gi|14290153|gb|AAK59278.1| thaumatin-like protein [Sambucus nigra]	368	0	880	21	pfam00314
9	gi|6273385|gb|AAF06347.1| SCUTL2 [Vitis vinifera]	228	5.00E-84	925	10	smart00205
10	gi|4586372|dbj|BAA74546.2| thaumatin-like protein SE39b [Nicotiana tabacum]	281	2.00E-79	1037	3	smart00205
11	gi|15221033|ref|NP_173261.1| thaumatin, putative [Arabidopsis thaliana]	399	0	1147	6	smart00205
12	gi|12324220|gb|AAG52086.1| thaumatin-like protein; 9376-10898 [Arabidopsis thaliana]	314	0	858	7	smart00205
13	gi|2501182|sp|Q41350|OLP1_LYCES Osmotin-like protein precursor [Lycopersicon esculentum]	413	0	1135	10	smart00205
14	gi|38603816|gb|AAR24653.1| At5g40020 [Arabidopsis thaliana]	148	3.00E-34	890	2	smart00205
15	gi|7270553|emb|CAB81510.1| thaumatin-like protein [Arabidopsis thaliana]	343	6.00E-93	870	4	smart00205
16	gi|12323299|gb|AAG51631.1| thaumatin-like protein; 12104-13574 [Arabidopsis thaliana]	261	2.00E-68	765	2	smart00205

Table S10: EST-Contigs with e-value $< e^{-20}$ and score > 100 obtained in the Project Phytoalexin, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Phytoalexin						
Contig	**BLAST NR**	**Score**	**e-value**	**Length**	**Reads**	**Conserved Domains**
1	gi|58826317|gb|AAW82883.1| phytoalexin-deficient 4-1 protein [Solanum tuberosum]	123	7.00E-27	759	2	cd00519
2	gi|58826317|gb|AAW82883.1| phytoalexin-deficient 4-1 protein [Solanum tuberosum]	126	9.00E-28	683	2	cd00519
3	gi|58826317|gb|AAW82883.1| phytoalexin-deficient 4-1 protein [Solanum tuberosum]	528	0	1507	5	cd00519

Table S11: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Chalconesynthase, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Chalconesynthase						
Contig	BLAST NR	Score	e-value	Length	Reads	Conserved Domains
1	gi|50380153|gb|AAT76306.1| aldo-keto reductase [Fragaria x ananassa]	202	6.00E-51	657	2	pfam00248
2	gi|77381754|gb|ABA73267.1| 3-oxoacyl-(acyl-carrier-protein) synthase III, putative [Pseudomonas fluorescens PfO-1]	271	1.00E-71	759	3	cd00830, PRK09352
3	gi|1345787|sp|P48387|CHS2_CAMSI Chalcone synthase 2 (Naringenin-chalcone synthase 2) [Camellia sinensis]	691	0	1775	15	cd00831, COG3424
4	gi|1345787|sp|P48387|CHS2_CAMSI Chalcone synthase 2 (Naringenin-chalcone synthase 2) [Camellia sinensis]	660	0	1514	52	cd00831, COG3424
5	gi|1345787|sp|P48387|CHS2_CAMSI Chalcone synthase 2 (Naringenin-chalcone synthase 2) [Camellia sinensis]	713	0	1509	73	cd00831, COG3424

Table S12: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Polyphenoloxydase, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

<table>
<tr><th colspan="7">Polyphenoloxydase</th></tr>
<tr><th>Contig</th><th>BLAST NR</th><th>Score</th><th>e-value</th><th>Length</th><th>Reads</th><th>Conserved Domains</th></tr>
<tr><td>1</td><td>gi|51872305|gb|AAU12257.1| polyphenol oxidase [Populus balsamifera subsp. trichocarpa x Populus deltoides]</td><td>616</td><td>0</td><td>1851</td><td>8</td><td>pfam00264</td></tr>
<tr><td>2</td><td>gi|1172584|sp|P43309|PPO_MALDO Polyphenol oxidase, chloroplast precursor (PPO) (Catechol oxidase) [Malus x domestica]</td><td>510</td><td>0</td><td>2127</td><td>11</td><td>pfam00264</td></tr>
<tr><td>3</td><td>gi|4158170|emb|CAA06855.1| catechol oxidase; polyphenol oxidase [Ipomoema batatas]</td><td>608</td><td>0</td><td>2070</td><td>30</td><td>pfam00264</td></tr>
<tr><td>4</td><td>gi|6137618|pdb|1BT3|A Chain A, Catechol Oxidase [Ipomoea Batatas]</td><td>315</td><td>2.00E-84</td><td>1000</td><td>14</td><td>pfam00264</td></tr>
</table>

Table S13: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Importin, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

<table>
<tr><th colspan="7">Importin</th></tr>
<tr><th>Contig</th><th>BLAST NR</th><th>Score</th><th>e-value</th><th>Length</th><th>Reads</th><th>Conserved Domains</th></tr>
<tr><td>1</td><td>gi|15238758|ref|NP_200160.1| protein transporter [Arabidopsis thaliana]</td><td>668</td><td>0</td><td>1309</td><td>7</td><td>pfam03810, cd00020, COG5215</td></tr>
<tr><td>2</td><td>gi|13752562|gb|AAK38727.1|AF369707_1 importin alpha 2 [Capsicum annuum]</td><td>299</td><td>1E-79</td><td>807</td><td>3</td><td>cd00020, pfam01749, COG5064</td></tr>
<tr><td>4</td><td>gi|24940138|emb|CAD11990.1| rubisco small subunit [Coffea arabica]</td><td>163</td><td>8E-39</td><td>833</td><td>2</td><td>cd03527</td></tr>
<tr><td>5</td><td>gi|13752562|gb|AAK38727.1|AF369707_1 importin alpha 2 [Capsicum annuum]</td><td>555</td><td>0</td><td>1238</td><td>6</td><td>cd00020, pfam01749, COG5064</td></tr>
<tr><td>6</td><td>gi|84453224|dbj|BAE71209.1| putative importin alpha [Trifolium pratense]</td><td>223</td><td>3E-59</td><td>873</td><td>5</td><td>cd00020, pfam01749, COG5064</td></tr>
<tr><td>7</td><td>gi|13752562|gb|AAK38727.1|AF369707_1 importin alpha 2 [Capsicum annuum]</td><td>369</td><td>0</td><td>1029</td><td>4</td><td>cd00020, pfam01749, COG5064</td></tr>
<tr><td>8</td><td>gi|84453224|dbj|BAE71209.1| putative importin alpha [Trifolium pratense]</td><td>492</td><td>0</td><td>1482</td><td>12</td><td>cd00020, pfam01749, COG5064</td></tr>
<tr><td>9</td><td>gi|23954312|emb|CAC79691.1| Importin beta-like protein [Oryza sativa]</td><td>431</td><td>0</td><td>954</td><td>2</td><td>pfam03810, cd00020, COG5215</td></tr>
<tr><td>10</td><td>gi|84453224|dbj|BAE71209.1| putative importin alpha [Trifolium pratense]</td><td>128</td><td>4E-36</td><td>1147</td><td>2</td><td>cd00020, pfam01749, COG5064</td></tr>
<tr><td>11</td><td>gi|26449534|dbj|BAC41893.1| putative importin beta [Arabidopsis thaliana]</td><td>516</td><td>0</td><td>950</td><td>2</td><td>pfam03810, cd00020, COG5215</td></tr>
</table>

Table S14: EST-Contigs with e-value < e^{-20} and score > 100 obtained in the Project Glucosyltransferase, and their blast hit, score, e-value, size, number of reads, and conserved domains from putative proteins.

Glucosyltransferase						
Contig	BLAST NR	Score	e-value	Length	Reads	Conserved Domains
1	gi|147795324|emb|CAN67249.1| hypothetical protein [Vitis vinifera]	271	3.00E-71	671	2	pfam00201, COG1819
2	gi|147767625|emb|CAN60198.1| hypothetical protein [Vitis vinifera]	347	7.00E-94	1093	4	pfam00201, COG1819
3	gi|50284480|dbj|BAD29721.1| UDP-glucose glucosyltransferase [Catharanthus roseus]	476	0	1248	3	pfam00201
4	gi|147861933|emb|CAN78769.1| hypothetical protein [Vitis vinifera]	481	0	1128	5	pfam00201, COG1819
5	gi|119640513|gb|ABL85473.1| gylcosyltransferase UGT71A13 [Maclura pomifera]	120	1.00E-25	827	2	pfam00201, COG1819
6	gi|28380078|sp|Q9AR73|HQGT_RAUSE Hydroquinone glucosyltransferase (Arbutin synthase) [Rauvolfia serpentina]	580	0	1345	10	pfam00201, COG1819
7	gi|147821100|emb|CAN70963.1| hypothetical protein [Vitis vinifera]	358	1.00E-97	681	2	pfam00201, COG1819
9	gi|15228063|ref|NP_181234.1| UDP-glucoronosyl/UDP-glucosyl transferase family protein [Arabidopsis thaliana]	385	0	1184	8	pfam00201, COG1819
10	gi|147854062|emb|CAN83394.1| hypothetical protein [Vitis vinifera]	315	1.00E-84	710	2	pfam00201, COG1819
11	gi|125541419|gb|EAY87814.1| hypothetical protein OsI_009047 [Oryza sativa]	392	0	1124	5	COG0380, pfam02358, pfam00982
12	gi|454245|emb|CAA50376.1| anthocyanin 3 glucoside: rhamnosyltransferase [Petunia x hybrida]	422	0	1879	8	COG1819
13	gi|58430500|dbj|BAD89044.1| putative glycosyltransferase [Solanum aculeatissimum]	316	1.00E-84	906	3	pfam00201
14	gi|28380078|sp|Q9AR73|HQGT_RAUSE Hydroquinone glucosyltransferase (Arbutin synthase) [Rauvolfia serpentina]	459	0	1208	11	pfam00201, COG1819
15	gi|146148629|gb|ABQ02257.1| O-glucosyltransferase 2 [Vitis labrusca]	365	5.00E-99	1415	8	pfam00201, COG1819
17	gi|147855978|emb|CAN80742.1| hypothetical protein [Vitis vinifera]	473	0	1600	5	pfam00201, COG1819
18	gi|147846163|emb|CAN81633.1| hypothetical protein [Vitis vinifera]	310	8.00E-83	817	2	pfam00201, COG1819
19	gi|147767625|emb|CAN60198.1| hypothetical protein [Vitis vinifera]	257	4.00E-67	790	2	pfam00201, COG1819
20	gi|62241067|dbj|BAD93690.1| glycosyltransferase NTGT5b [Nicotiana tabacum]	299	1.00E-79	702	3	pfam00201
21	gi|147812173|emb|CAN61517.1| hypothetical protein [Vitis vinifera]	369	0	717	2	pfam00201, COG1819
22	gi|145329987|ref|NP_001077979.1| UDP-glucoronosyl/UDP-glucosyl transferase family protein [Arabidopsis thaliana]	199	1.00E-49	770	6	pfam00201, COG1819
23	gi|38532106|gb|AAR23313.1| beta-1,4-mannan synthase [Cyamopsis tetragonoloba]	290	7.00E-77	778	3	pfam00535, COG1215
24	gi|115467628|ref|NP_001057413.1| Os06g0288300 [Oryza sativa]	145	1.00E-40	972	6	pfam00201
25	gi|15236546|ref|NP_194913.1| PGR3 (PROTON GRADIENT REGULATION 3); binding [Arabidopsis thaliana]	291	4.00E-77	835	3	pfam01535, COG0457, pfam06239
26	gi|15242557|ref|NP_195906.1| pentatricopeptide (PPR) repeat-containing protein [Arabidopsis thaliana]	346	8.00E-94	743	2	pfam01535
27	gi|70734017|ref|YP_257657.1| lipopolysaccharide core biosynthesis protein RfaG [Pseudomonas fluorescens Pf-5]	405	0	817	2	pfam00534, COG0438
28	gi|147806166|emb|CAN70002.1| hypothetical protein [Vitis vinifera]	205	2.00E-51	722	4	pfam00201, COG1819
29	gi|4455123|gb|AAD21086.1| flavonoid 3-O-glucosyltransferase [Forsythia x intermedia]	273	0	1137	2	pfam00201
30	gi|34148023|gb|AAQ62571.1| glycosyltransferase 5 [Ipomoea trifida]	306	1.00E-81	901	3	pfam00535,COG1215
31	gi|21435782|gb|AAM53963.1|AF515727_1 UDP-glucosyltransferase [Stevia rebaudiana]	229	1.00E-58	714	2	pfam04101, pfam00201, COG1819
32	gi|70734017|ref|YP_257657.1| lipopolysaccharide core biosynthesisprotein RfaG [Pseudomonas fluorescens Pf-5]	438	0	858	7	pfam00534, COG0438
33	gi|629669|pir||S39507 glucuronosyl transferase homolog, ripening-related [Solanum lycopersicon]	180	7.00E-44	819	2	pfam00201
34	gi|62241065|dbj|BAD93689.1| glycosyltransferase NTGT5a [Nicotiana tabacum]	496	0	1523	13	pfam00201
35	gi|20149064|gb|AAM12787.1| putative anthocyanidine rhamnosyl-transferase [Capsicum annuum]	380	0	1696	4	COG1819
36	gi|19568935|gb|AAL91978.1|AF483209_1 putative trehalose synthase [Solanum tuberosum]	599	0	1103	4	COG0380, pfam02358, pfam00982
37	gi|22759895|dbj|BAC10994.1| rhamnosyl transferase [Nierembergia sp. NB17]	216	1.00E-54	693	2	pfam00201, COG1819
38	gi|15224213|ref|NP_179460.1| ATTPS11 [Arabidopsis thaliana]	304	3.00E-81	755	2	COG0380, pfam02358, pfam00982
39	gi|147795873|emb|CAN74227.1| hypothetical protein [Vitis vinifera]	385	0	1078	4	pfam00201, COG1819
40	gi|50284482|dbj|BAD29722.1| UDP-glucose glucosyltransferase [Catharanthus roseus]	630	0	1588	6	pfam00201
41	gi|126668372|ref|ZP_01739330.1| glycosyl transferases-like protein [Marinobacter sp. ELB17]	296	5.00E-79	729	3	pfam00535, PRK10714
42	gi|51705431|gb|AAU09445.1| putative UDP-rhamnose:rhamnosyltransferase [Fragaria x ananassa]	217	9.00E-66	1114	5	pfam00201
43	gi|119640480|gb|ABL85472.1| glycosyltransferase UGT72B9 [Maclura pomifera]	114	2.00E-24	516	3	pfam00201
44	gi|147768453|emb|CAN78332.1| hypothetical protein [Vitis vinifera]	243	9.00E-63	842	2	pfam00201, COG1819
45	gi|77457186|ref|YP_346691.1| alginate biosynthesis protein Alg8 [Pseudomonas fluorescens PfO-1]	400	0	869	3	No CD has been identified
46	gi|147767625|emb|CAN60198.1| hypothetical protein [Vitis vinifera]	217	5.00E-55	685	3	pfam00201, COG1819
47	gi|92877343|gb|ABE84479.1| Pentatricopeptide repeat [Medicago truncatula]	167	6.00E-40	719	2	pfam01535
48	gi|21435782|gb|AAM53963.1|AF515727_1 UDP-glucosyltransferase [Stevia rebaudiana]	410	0	1639	6	pfam04101, pfam00201, COG1819
49	gi|51705413|gb|AAU09444.1| UDP-glucose glucosyltransferase [Fragaria anassa]	375	0	1280	4	pfam00201
50	gi|147815573|emb|CAN68287.1| hypothetical protein [Vitis vinifera]	266	8.00E-70	750	3	pfam00201, COG1819
51	gi|147781122|emb|CAN71907.1| hypothetical protein [Vitis vinifera]	255	2.00E-66	788	2	pfam00201, COG1819
52	gi|20146093|dbj|BAB88935.1| glucosyltransferase NTGT2 [Nicotiana tabacum]	655	0	1831	18	pfam03033, COG1819
53	gi|147811099|emb|CAN70169.1| hypothetical protein [Vitis vinifera]	362	0	1517	11	pfam00201, COG1819
54	gi|147846160|emb|CAN79485.1| hypothetical protein [Vitis vinifera]	461	0	1037	2	pfam00201, COG1819
55	gi|15227717|ref|NP_180576.1| UDP-glucoronosyl/UDP-glucosyl transferase family protein [Arabidopsis haliana]	408	0	1532	5	COG1819
56	gi|115473997|ref|NP_001060597.1| Os07g0671200 [Oryza sativa]	160	3.00E-38	571	2	pfam01535
57	gi|24459979|dbj|BAC22617.1| UDP-glucose:sterol 3-O-glucosyltransferase [Panax ginseng]	511	0	809	3	pfam03033, COG1819
58	gi|147855978|emb|CAN80742.1| hypothetical protein [Vitis vinifera]	310	8.00E-83	961	5	pfam00201, COG1819
59	gi|14349251|dbj|BAB60720.1| glucosyltransferase [Nicotiana tabacum]	544	0	1366	7	pfam00201
60	gi|147815573|emb|CAN68287.1| hypothetical protein [Vitis vinifera]	226	3.00E-60	812	2	pfam00201, COG1819
61	gi|14192682|gb|AAK54465.1| cold-induced glucosyl transferase [Solanum tuberosum]	192	2.00E-52	768	2	pfam03033, pfam00201, COG1819
62	gi|116310407|emb|CAH67416.1| OSIGBa0143N19.10 [Oryza sativa]	290	2.00E-78	1123	2	pfam01501
63	gi|115310620|emb|CAJ32597.1| sucrose synthase [Coffea arabica]	483	0	871	3	pfam00534, pfam00862
64	gi|147787516|emb|CAN77812.1| hypothetical protein [Vitis vinifera]	151	3.00E-35	714	3	pfam00201, COG1819
65	gi|147864250|emb|CAN83017.1| hypothetical protein [Vitis vinifera]	229	5.00E-60	969	2	pfam00201, COG1819
66	gi|147811099|emb|CAN70169.1| hypothetical protein [Vitis vinifera]	106	1.00E-37	968	4	pfam00201, COG1819
67	gi|147795324|emb|CAN67249.1| hypothetical protein [Vitis vinifera]	232	2.00E-59	709	3	pfam00201, COG1819
68	gi|147800590|emb|CAN77507.1| hypothetical protein [Vitis vinifera]	148	1.00E-34	522	2	pfam00201, COG1819
69	gi|147818360|emb|CAN62624.1| hypothetical protein [Vitis vinifera]	202	3.00E-68	804	2	pfam00201, COG1819

70	gi|147821441|emb|CAN74580.1| hypothetical protein [Vitis vinifera]	358	1.00E-97	801	3	pfam00201, COG1819
71	gi|125553006|gb|EAY98715.1| hypothetical protein OsI_019948 [Oryza sativa]	801	0	2030	8	COG0380, pfam02358, pfam00982
72	gi|115310620|emb|CAJ32597.1| sucrose synthase [Coffea arabica]	528	0	1066	3	pfam00534, pfam00862
73	gi|27461029|gb|AAL06646.1| flavonoid 1-2 rhamnosyltransferase [Citrus maxima]	241	4.00E-62	810	2	COG1819
74	gi|15225508|ref|NP_181493.1| glycosyl transferase family 2 protein [Arabidopsis thaliana]	469	0	1434	6	pfam00535
75	gi|140049083|gb|ABO79665.1| Glycosyl transferase, family 8 [Medicago truncatula]	239	4.00E-61	1408	8	pfam01501
76	gi|147809370|emb|CAN71202.1| hypothetical protein [Vitis vinifera]	775	0	1312	2	pfam00201, COG1819
77	gi|147855978|emb|CAN80742.1| hypothetical protein [Vitis vinifera]	305	4.00E-81	1064	3	pfam00201, COG1819
78	gi|147828262|emb|CAN62185.1| hypothetical protein [Vitis vinifera]	337	4.00E-91	782	2	pfam00201, COG1819
79	gi|115467634|ref|NP_001057416.1| Os06g0289200 [Oryza sativa]	251	2.00E-84	1332	7	pfam00201
80	gi|147855615|emb|CAN83466.1| hypothetical protein [Vitis vinifera]	478	0	1008	5	pfam00201, COG1819
81	gi|147767625|emb|CAN60198.1| hypothetical protein [Vitis vinifera]	495	0	1544	5	pfam00201, COG1819
82	gi|147826555|emb|CAN61971.1| hypothetical protein [Vitis vinifera]	380	0	757	3	pfam00201, COG1819
83	gi|62241063|dbj|BAD93688.1| glucosyltransferase NTGT4 [Nicotiana tabacum]	247	4.00E-64	652	2	pfam00201
84	gi|102139944|gb|ABF70084.1| trehalose-6-phosphate synthase, putative [Musa balbisiana]	128	4.00E-28	818	2	COG0380, pfam02358, pfam00982
85	gi|147862190|emb|CAN82595.1| hypothetical protein [Vitis vinifera]	891	0	2586	16	pfam00201, COG1819
86	gi|110932098|gb|ABH03018.1| resveratrol/hydroxycinnamic acid O-glucosyltransferase [Vitis labrusca]	708	0	1719	15	pfam00201, COG1819
87	gi|52139814|gb|AAU29197.1| sucrose phosphate synthase [Lycopersicon esculentum]	570	0	1062	3	pfam00534, pfam05116
88	gi|147779442|emb|CAN74357.1| hypothetical protein [Vitis vinifera]	509	0	830	2	pfam00201, COG1819
89	gi|67845751|emb|CAI56307.1| sucrose synthase [Coffea canephora]	1625	0	3092	191	pfam00534, pfam00862
90	gi|147818682|emb|CAN76185.1| hypothetical protein [Vitis vinifera]	1157	0	2346	16	pfam00201, COG1819
91	gi|20149064|gb|AAM12787.1| putative anthocyanidine rhamnosyl-transferase [Capsicum annuum]	196	2.00E-49	360	2	COG1819
93	gi|15235900|ref|NP_192536.1| ATCSLC12 (Cellulose synthase-like C12); transferase, transferring glycosyl groups [Arabidopsis thaliana]	956	0	2097	7	pfam00535
94	gi|115310620|emb|CAJ32597.1| sucrose synthase [Coffea arabica]	889	0	1326	8	pfam00534, pfam00862
95	gi|115310618|emb|CAJ32596.1| sucrose synthase [Coffea arabica]	401	0	851	4	pfam00534, pfam00862
96	gi|20146093|dbj|BAB88935.1| glucosyltransferase NTGT2 [Nicotiana tabacum]	544	0	1315	8	pfam03033, COG1819
97	gi|42795466|gb|AAS46243.1| xyloglucan endotransglucosylase-hydrolase XTH7 [Lycopersicon esculentum]	392	0	889	14	cd02176, pfam06955
98	gi|147833323|emb|CAN61988.1| hypothetical protein [Vitis vinifera]	320	5.00E-86	769	2	pfam00201, COG1819
99	gi|67845751|emb|CAI56307.1| sucrose synthase [Coffea canephora]	474	0	1024	3	pfam00534, pfam00862
100	gi|147800590|emb|CAN77507.1| hypothetical protein [Vitis vinifera]	397	0	1376	5	pfam00201, COG1819
101	gi|147768350|emb|CAN60444.1| hypothetical protein [Vitis vinifera]	349	7.00E-95	726	2	pfam00201, COG1819
102	gi|147815573|emb|CAN68287.1| hypothetical protein [Vitis vinifera]	207	2.00E-63	898	3	pfam00201, COG1819
103	gi|147768350|emb|CAN60444.1| hypothetical protein [Vitis vinifera]	166	5.00E-40	541	2	pfam00201, COG1819
104	gi|115310618|emb|CAJ32596.1| sucrose synthase [Coffea arabica]	279	2.00E-88	997	2	pfam00534, pfam00862
105	gi|147815573|emb|CAN68287.1| hypothetical protein [Vitis vinifera]	309	2.00E-82	902	11	pfam00201, COG1819
106	gi|145337405|ref|NP_177278.3| UDP-glucose:glycoprotein glucosyltransferase [Arabidopsis thaliana]	572	0	1477	10	pfam06427, pfam01501
107	gi|37993657|gb|AAR06914.1| UDP-glycosyltransferase 71E1 [Stevia rebaudiana]	286	7.00E-76	735	2	pfam00201
109	gi|27461029|gb|AAL06646.1| flavonoid 1-2 rhamnosyltransferase [Citrus maxima]	198	3.00E-49	764	3	COG1819
110	gi|15226332|ref|NP_180375.1| glycosyltransferase family protein [Arabidopsis thaliana]	434	0	1178	2	pfam00201, COG1819
111	gi|19743740|gb|AAL92461.1| putative glucosyltransferase [Lycopersicon esculentum]	361	0	1183	3	pfam00201
112	gi|147811099|emb|CAN70169.1| hypothetical protein [Vitis vinifera]	108	3.00E-22	844	2	pfam00201, COG1819
113	gi|147790314|emb|CAN74373.1| hypothetical protein [Vitis vinifera]	206	9.00E-52	824	2	pfam00201, COG1819
114	gi|147787516|emb|CAN77812.1| hypothetical protein [Vitis vinifera]	203	9.00E-51	783	5	pfam00201, COG1819
115	gi|147790314|emb|CAN74373.1| hypothetical protein [Vitis vinifera]	255	2.00E-66	754	2	pfam00201, COG1819
116	gi|147768688|emb|CAN76057.1| hypothetical protein [Vitis vinifera]	462	0	1759	6	pfam00201, COG1819
117	gi|15228063|ref|NP_181234.1| UDP-glucoronosyl/UDP-glucosyl transferase family protein [Arabidopsis thaliana]	560	0	1316	2	pfam06427, pfam01501
118	gi|22326571|ref|NP_195861.2| DIE2/ALG10 family [Arabidopsis thaliana]	333	0	1670	7	pfam04922
119	gi|147815573|emb|CAN68287.1| hypothetical protein [Vitis vinifera]	301	2.00E-80	651	3	pfam00201, COG1819
121	gi|20146093|dbj|BAB88935.1| glucosyltransferase NTGT2 [Nicotiana tabacum]	280	8.00E-74	967	7	pfam03033, COG1819
122	gi|62241067|dbj|BAD93690.1| glycosyltransferase NTGT5b [Nicotiana tabacum]	307	4.00E-82	769	5	pfam00201
123	gi|146148629|gb|ABQ02257.1| O-glucosyltransferase 2 [Vitis labrusca]	342	3.00E-92	1235	7	pfam00201, COG1819
124	gi|147855978|emb|CAN80742.1| hypothetical protein [Vitis vinifera]	302	4.00E-82	1097	2	pfam00201, COG1819
125	gi|19743740|gb|AAL92461.1| putative glucosyltransferase [Lycopersicon esculentum]	451	0	1314	5	pfam00201
126	gi|77176831|gb|ABA64521.1| sucrose-phosphate synthase isoform B [Nicotiana tabacum]	579	0	1022	3	pfam00534, pfam05116, pfam00862
127	gi|147790314|emb|CAN74373.1| hypothetical protein [Vitis vinifera]	514	0	1819	16	pfam00201, COG1819
128	gi|629669|pir||S39507 glucuronosyl transferase homolog, ripening-related [Solanum lycopersicon]	455	0	1256	2	pfam00201
129	gi|147790314|emb|CAN74373.1| hypothetical protein [Vitis vinifera]	228	6.00E-58	1034	3	pfam00201, COG1819
130	gi|92884230|gb|ABE87250.1| UDP-glucuronosyl/UDP-glucosyltransferase [Medicago truncatula]	137	6.00E-62	841	2	pfam00201

Table S15: ESTs from NS1, RM1, RX1 and SS1 libraries, shown in Figure 1.

	Project NBS-LRR		
ESTs NS1	**BLAST X**	**E-value**	**Size (bp)**
1	P0039H02.32 [Oryza sativa (japonica cultivar-group)]	3.00E-11	797
2	rad-51 (Fifty one) like, Short RFS-1, f-box protein Fbl2 (51.6 kD) (rfs-1Co) [Caenorhabditis elegans]	0.005	871
3	NBS-LRR protein [Solanum acaule]	7.00E-13	889
4	F-box and leucine-rich repeat protein 2; F-box protein containing leucine-rich repeats [Homo sapiens]	3.00E-09	919
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	NBS-LRR-like protein [Oryza sativa (japonica cultivar-group)]	9.00E-26	908
2	NBS-LRR resistance-like protein J71 [Phaseolus vulgaris]	1.00E-26	908
3	hypothetical protein F6H11.60 - Arabidopsis thaliana	4.00E-33	987
4	hypothetical protein F6H11.60 - Arabidopsis thaliana	0.001	981
5	elicitor-inducible LRR receptor-like protein EILP [Nicotiana tabacum]	5.00E-34	878
6	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	4.00E-13	872
7	LRR receptor-like protein kinase [Nicotiana tabacum]	5.00E-90	914
8	putative TIR/NBS/LRR disease resistance protein [Pinus taeda]	6.00E-08	889
9	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	1.00E-43	930
10	leucine-rich repeat transmembrane protein kinase, putative [Arabidopsis thaliana]	4.00E-21	901
11	disease resistance protein (TIR-NBS-LRR class), putative [Arabidopsis thaliana]	3.00E-07	925
12	hypernodulation aberrant root formation protein [Lotus japonicus]	8.00E-26	809
13	LRR receptor-like kinase 2 [Arabidopsis thaliana]	1.00E-07	871
14	similar to NBS-LRR type resistance gene [Oryza sativa (japonica cultivar-group)]	3.00E-46	923
15	leucine-rich repeat transmembrane protein kinase, putative [Arabidopsis thaliana]	7.00E-31	922
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	disease resistance protein family (LRR) [Arabidopsis thaliana]	3.00E-05	957
2	LRR receptor protein kinase -related [Arabidopsis thaliana]	1.00E-16	972
3	NBS-LRR protein [Solanum acaule]	1.00E-17	837
4	Phytosulfokine receptor precursor (Phytosulfokine LRR receptor kinase)	2.00E-08	966
5	leucine-rich repeat protein LRP - tomato	5.00E-56	932
6	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	2.00E-06	960
7	disease resistance protein (NBS-LRR class), putative [Arabidopsis thaliana]	0.0004	1008
8	NBS-LRR disease resistance protein homologue [Hordeum vulgare]	5.00E-29	988
9	LRR receptor-like kinase 1 [Arabidopsis thaliana]	0.004	967
10	NBS-LRR protein [Solanum acaule]	5.00E-17	1017
11	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	2.00E-08	889
12	coronatine-insensitive 1 (COI1), AtFBL2 [Arabidopsis thaliana]	7.00E-05	983
13	coronatine-insensitive 1 (COI1), AtFBL2 [Arabidopsis thaliana]	3.00E-07	976
14	NBS-LRR protein [Solanum acaule]	5.00E-13	857
15	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	9.00E-07	882
16	putative NBS-LRR type resistance protein, 3' partial [Oryza sativa]	0.0004	908
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	NBS-LRR protein [Solanum tuberosum]	3.00E-24	888
2	GPA2-like NBS-LRR protein [Solanum nigrum]	2.00E-15	878
3	P0039H02.32 [Oryza sativa (japonica cultivar-group)]	0.0002	874
	Project Resistance		
ESTs NS1	**BLAST X**	**E-value**	**Size (bp)**
1	disease resistance protein Prf - tomato	2.00E-14	889
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	hypothetical protein Rv1258c [Mycobacterium tuberculosis H37Rv]	1.00E-05	907
2	fusaric acid resistance protein [Xanthomonas campestris pv. campestris str. ATCC 33913]	1.00E-15	881
3	leucine rich repeat protein family [Arabidopsis thaliana]	9.00E-38	876
4	pleiotropic drug resistance like protein [Nicotiana tabacum]	7.00E-09	905
5	Putative disease resistance protein At1g58400	4.00E-13	872
6	daunorubicin resistance protein [Enterococcus faecalis V583]	1.00E-17	854
7	disease resistance protein family [Arabidopsis thaliana]	3.00E-34	878
8	orf, hypothetical protein [Escherichia coli K12]	9.00E-52	884
9	orf, hypothetical protein [Escherichia coli K12]	3.00E-42	845
10	Acriflavin resistance protein F (EnvD protein). [Escherichia coli]	0.002	976
11	dTDP-glucose 4-6-dehydratase homolog D18 - Arabidopsis thaliana	1.00E-35	925
12	[Segment 1 of 2] Mercuric reductase (Hg(II) reductase)	3.00E-29	933
13	Glyoxalase, Glyoxalase/Bleomycin resistance protein/Dioxygenase superfamily [Bacillus anthracis A2012]	3.00E-19	973
14	putative Cf2/Cf5 disease resistance protein [Oryza sativa (japonica cultivar-group)]	1.00E-31	941
15	leucine rich repeat protein family [Arabidopsis thaliana]	0.003	963
16	leucine rich repeat protein family [Arabidopsis thaliana]	6.00E-11	939
17	Multidrug resistance ABC transporter ATP-binding and permease protein [Bacillus cereus ATCC 14579]	5.00E-09	945
18	D-alanyl-D-alanine dipeptidase (D-Ala-D-Ala dipeptidase) (Vancomycin B-type resistance protein vanX)	6.00E-15	943
19	membrane protein OpdE [Pseudomonas aeruginosa PA01]	6.00E-24	902
20	integral membrane lipid kinase-like protein [Mesorhizobium loti]	3.00E-17	947
21	expressed protein [Arabidopsis thaliana]	8.00E-40	921
22	orf, hypothetical protein [Escherichia coli K12]	3.00E-53	635
23	pleiotropic drug resistance like protein [Nicotiana tabacum]	4.00E-34	809
24	FUSARIC ACID RESISTANCE PROTEIN FUSE [Brucella melitensis]	2.00E-05	947
25	fusaric acid resistance protein, putative [Pseudomonas putida KT2440]	9.00E-08	943
26	pleiotropic drug resistance like protein [Nicotiana tabacum]	4.00E-43	908
27	disease resistance-like protein [Coffea arabica]	2.00E-71	908
28	dTDP-glucose 4-6-dehydratase homolog D18 - Arabidopsis thaliana	2.00E-31	923
29	disease resistance-like protein [Coffea arabica]	3.00E-68	908

30	senescence-associated protein [Arabidopsis thaliana]	2.00E-33	944
31	putative TIR/NBS/LRR disease resistance protein [Pinus taeda]	6.00E-08	889
32	Tetracenomycin C resistance and export protein	3.00E-18	972
33	FUSARIC ACID RESISTANCE PROTEIN FUSB / FUSARIC ACID RESISTANCE PROTEIN FUSC [Brucella melitensis]	3.00E-05	1040
34	dTDP-glucose 4-6-dehydratase homolog D18 - Arabidopsis thaliana	2.00E-15	924
35	ACRIFLAVIN RESISTANCE PROTEIN D [Brucella melitensis]	8.00E-11	929
36	similar to Pib(rice blast resistance gene) [Oryza sativa (japonica cultivar-group)]	8.00E-49	923
37	natural resistance-associated macrophage protein [Burkholderia cepacia genomovar III]	3.00E-28	939
38	golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1 (3M842) [Caenorhabditis elegans]	1.00E-08	956
39	multidrug resistance P-glycoprotein, putative [Arabidopsis thaliana]	5.00E-54	1022
40	outer membrane protein OprM precursor [Pseudomonas aeruginosa PA01]	0.003	985
41	PROBABLE RESPONSE REGULATOR FOR COBALT ZINC CADMIUM RESISTANCE TRANSCRIPTION REGULATOR PROTEIN [Ralstonia solanacearum]	4.00E-71	921
42	putative RSH, disease resistance-related protein [Oryza sativa (japonica cultivar-group)]	6.00E-94	895
43	DAUNORUBICIN RESISTANCE TRANSMEMBRANE PROTEIN [Brucella melitensis]	7.00E-19	999
44	cobalt-zinc-cadmium resistance protein [Xylella fastidiosa Temecula1]	5.00E-05	921
45	multidrug resistance protein [Bradyrhizobium japonicum]	3.00E-43	912
46	fusaric acid resistance protein [Xanthomonas campestris pv. campestris str. ATCC 33913]	2.00E-16	857
47	multidrug resistance ABC transporter [Clostridium tetani E88]	1.00E-09	884
48	multidrug resistance ABC transporter [Clostridium tetani E88]	1.00E-07	912
49	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]	1.00E-43	930
50	PROBABLE RESPONSE REGULATOR FOR COBALT ZINC CADMIUM RESISTANCE TRANSCRIPTION REGULATOR PROTEIN [Ralstonia solanacearum]	5.00E-69	909
51	dTDP-glucose 4-6-dehydratase homolog D18 - Arabidopsis thaliana	4.00E-50	933
52	disease resistance protein (TIR-NBS-LRR class), putative [Arabidopsis thaliana]	3.00E-07	925
53	Tetracycline resistance protein from transposon Tn4351/Tn4400	5.00E-14	926
54	florfenicol resistance protein-like [Oryza sativa (japonica cultivar-group)]	2.00E-16	913
55	pleiotropic drug resistance like protein [Nicotiana tabacum]	3.00E-14	950
56	tellurium resistance protein TerC [Rickettsia conorii]	2.00E-07	948
57	PROBABLE MULTIDRUG RESISTANCE TRANSMEMBRANE PROTEIN [Ralstonia solanacearum]	4.00E-30	976
58	Tetracycline resistance protein from transposon Tn4351/Tn4400	4.00E-20	903
59	bacteriocin resistance protein, putative [Neisseria meningitidis MC58]	1.00E-07	906
60	fusaric acid resistance protein, putative [Pseudomonas putida KT2440]	0	875
61	bacteriocin resistance protein, putative [Neisseria meningitidis MC58]	4.00E-07	909
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	disease resistance protein [Lycopersicon esculentum]	5.00E-11	966
2	disease resistance protein BS2 [Capsicum chacoense]	3.00E-16	837
3	Putative disease resistance protein Hcr2-0B [Oryza sativa]	2.00E-12	952
4	multidrug resistance protein, putative [Vibrio cholerae]	9.00E-52	890
5	isease resistance E - tomato	4.00E-49	929
6	disease resistance protein -related [Arabidopsis thaliana]	1.00E-05	957
7	verticillium wilt disease resistance protein Ve2 [Lycopersicon esculentum]	1.00E-29	949
8	putative Cf2/Cf5 disease resistance protein [Oryza sativa (japonica cultivar-group)]	5.00E-22	940
9	probable disease resistance protein [imported] - Arabidopsis thaliana	2.00E-08	889
10	putative Mlo (pathogen resistance) protein [Oryza sativa]	1.00E-15	915
11	hypothetical protein T3F20.24 [imported] - Arabidopsis thaliana	2.00E-17	934
12	dTDP-glucose 4-6-dehydratase homolog D18 - Arabidopsis thaliana	3.00E-05	944
13	disease resistance protein D - tomato	1.00E-48	941
14	putative resistance protein [Lycopersicon esculentum]	6.00E-72	948
15	fusaric acid resistance protein, putative [Pseudomonas putida KT2440]	1.00E-13	959
16	disease resistance-like protein [Coffea canephora]	4.00E-22	1017
17	disease resistance protein Cf-2.1 - currant tomato	1.00E-18	992
18	ribosomal protein S14; 40S ribosomal protein S14; emetine resistance [Homo sapiens]	3.00E-52	962
19	Probable WRKY transcription factor 19 (WRKY DNA-binding protein 19)	5.00E-13	927
20	similar to Pib(rice blast resistance gene) [Oryza sativa (japonica cultivar-group)]	3.00E-29	988
21	disease resistance protein (NBS-LRR class), putative [Arabidopsis thaliana]	0.0004	1008
22	dTDP-glucose 4-6-dehydratase homolog D18 - Arabidopsis thaliana	2.00E-10	967
23	disease resistance protein Hcr2-5D - tomato	4.00E-12	971
24	disease resistance protein Prf - tomato	1.00E-14	960
25	hypothetical protein T3F20.24 [imported] - Arabidopsis thaliana	5.00E-22	955
26	dTDP-glucose 4-6-dehydratase homolog D18 - Arabidopsis thaliana	2.00E-96	1036
27	dTDP-glucose 4-6-dehydratase homolog D18 - Arabidopsis thaliana	1.00E-37	945
28	pleiotropic drug resistance like protein [Nicotiana tabacum]	3.00E-07	977
29	disease resistance response protein-related/ dirigent protein-related [Arabidopsis thaliana]	8.00E-32	912
30	probable fusaric acid resistance protein FusE II [Chromobacterium violaceum ATCC 12472]	2.00E-16	926
31	ribosomal protein S14; 40S ribosomal protein S14; emetine resistance [Homo sapiens]	9.00E-53	921
32	probable multidrug resistance efflux pump [Pseudomonas aeruginosa PA01]	1.00E-47	938
33	putative disease resistance protein [Oryza sativa]	6.00E-28	919
34	probable oxidative stress resistance two-component transmembrane sensor histidine kinase transcription regulator protein [Ralstonia solanacearum]	2.00E-14	916
35	barley stem rust resistance protein [Hordeum vulgare subsp. vulgare]	4.00E-06	891
36	SUMO-1 activating enzyme subunit 2 [Homo sapiens]	2.00E-25	901
37	SUMO-1 activating enzyme subunit 2 [Homo sapiens]	3.00E-58	883
38	multidrug resistance P-glycoprotein (pgp1) [Arabidopsis thaliana]	5.00E-18	837
39	disease resistance protein Prf - tomato	5.00E-23	857
40	disease resistance protein Prf - tomato	7.00E-33	882
41	putative NBS-LRR type resistance protein, 3' partial [Oryza sativa]	0.0004	908
42	drug resistance transporter, Bcr/CflA family protein [Pseudomonas syringae pv. tomato str. DC3000]	6.00E-47	923
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	pleiotropic drug resistance like protein [Nicotiana tabacum]	5.00E-27	890

2	disease resistance protein BS2 [Capsicum chacoense]	7.00E-28	878
3	disease resistance protein BS2 [Capsicum chacoense]	1.00E-31	888
4	disease resistance protein (TIR class), putative [Arabidopsis thaliana]	1.00E-13	895
5	senescence-associated protein [Arabidopsis thaliana]	2.00E-32	861
6	pleiotropic drug resistance like protein [Nicotiana tabacum]	5.00E-40	861
7	disease resistance protein Prf - tomato	0.0002	862
Project Chitinase			
ESTs NS1	**BLAST X**	**E-value**	**Size (bp)**
1	chitinase [Hevea brasiliensis]	2.00E-41	846
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	chitinase (EC 3.2.1.14) Chib1 - soybean	4.00E-51	953
2	glycosyl hydrolase family 19 (chitinase) [Arabidopsis thaliana]	7.00E-87	783
3	chitinase 3-like protein precursor [Trichosanthes kirilowii]	3.00E-32	993
4	chitinase (EC 3.2.1.14) class III, acidic - soybean	2.00E-20	964
5	chitinase (EC 3.2.1.14) class III, acidic - soybean	5.00E-15	902
6	chitinase 3-like protein precursor [Trichosanthes kirilowii]	4.00E-49	935
7	chitinase 3-like protein precursor [Trichosanthes kirilowii]	8.00E-45	881
8	class VII chitinase precursor [Gossypium hirsutum]	6.00E-76	938
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	chitinase III [Vitis vinifera]	2.00E-65	994
2	putative class III acidic chitinase [Oryza sativa (japonica cultivar-group)]	5.00E-64	941
3	chitinase [Hevea brasiliensis]	2.00E-52	941
4	chitinase (EC 3.2.1.14) - adzuki bean	2.00E-28	936
5	chitinase (EC 3.2.1.14) Chib1 - soybean	2.00E-60	918
6	chitinase homolog [Coffea arabica]	7.00E-45	983
7	ACIDIC ENDOCHITINASE PRECURSOR	3.00E-36	963
8	chitinase 3-like protein precursor [Trichosanthes kirilowii]	1.00E-65	904
9	chitinase homolog [Coffea arabica]	5.00E-53	913
10	chitinase [Psophocarpus tetragonolobus]	2.00E-60	875
11	chitinase [Trifolium repens]	1.00E-43	924
12	ACIDIC ENDOCHITINASE PRECURSOR	6.00E-67	870
13	chitinase [Hevea brasiliensis]	1.00E-50	907
14	chitinase [Hevea brasiliensis]	2.00E-75	968
15	chitinase precursor [Petroselinum crispum]	0	974
16	chitinase [Hevea brasiliensis]	1.00E-45	930
17	class III acidic chitinase [Malus x domestica]	1.00E-64	959
18	class VII chitinase precursor [Gossypium hirsutum]	6.00E-44	967
19	ACIDIC ENDOCHITINASE SE2 PRECURSOR	2.00E-55	984
20	probable chitinase (EC 3.2.1.14) precursor - cucumber	1.00E-16	933
21	chitinase homolog [Coffea arabica]	2.00E-38	1016
22	class III chitinase [Lupinus albus]	1.00E-66	962
23	chitinase 3-like protein precursor [Trichosanthes kirilowii]	4.00E-39	958
24	chitinase [Hevea brasiliensis]	1.00E-48	1016
25	chitinase 3-like protein precursor [Trichosanthes kirilowii]	6.00E-62	916
26	ACIDIC ENDOCHITINASE PRECURSOR	2.00E-57	952
27	chitinase (EC 3.2.1.14) class III, acidic - soybean	9.00E-18	927
28	chitinase 1 [Cucumis melo]	4.00E-18	917
29	class III chitinase [Lupinus albus]	1.00E-27	945
30	chitinase homolog [Coffea arabica]	3.00E-06	827
31	ACIDIC ENDOCHITINASE PRECURSOR	1.00E-38	925
32	chitinase [Psophocarpus tetragonolobus]	2.00E-63	959
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	chitinase 3-like protein precursor [Trichosanthes kirilowii]	1.00E-35	866
2	chitinase 3-like protein precursor [Trichosanthes kirilowii]	4.00E-25	869
3	class III chitinase-like protein [Sesbania rostrata]	0.0006	872
4	chitinase 3-like protein precursor [Trichosanthes kirilowii]	4.00E-12	862
5	probable chitinase (EC 3.2.1.14) precursor - cucumber	5.00E-14	856
6	chitinase (EC 3.2.1.14) class III, acidic - soybean	2.00E-18	861
7	ACIDIC ENDOCHITINASE PRECURSOR	1.00E-83	868
8	chitinase 3-like protein precursor [Trichosanthes kirilowii]	3.00E-26	860
9	chitinase 3-like protein precursor [Trichosanthes kirilowii]	5.00E-20	874
10	chitinase 3-like protein precursor [Trichosanthes kirilowii]	4.00E-27	875
11	probable chitinase (EC 3.2.1.14) precursor - cucumber	7.00E-13	890
12	probable chitinase (EC 3.2.1.14) precursor - cucumber	5.00E-16	865
13	chitinase 3-like protein precursor [Trichosanthes kirilowii]	5.00E-53	854
14	chitinase (EC 3.2.1.14) class III, acidic - soybean	4.00E-13	835
15	probable chitinase (EC 3.2.1.14) precursor - cucumber	3.00E-14	875
16	ACIDIC ENDOCHITINASE PRECURSOR	6.00E-74	883
17	Glycosidase, Chitin Degradation, Multifunctional Enzyme	1.00E-53	911
18	chitinase homolog [Coffea arabica]	3.00E-53	894
19	chitinase 3-like protein precursor [Trichosanthes kirilowii]	7.00E-34	870
20	chitinase [Psophocarpus tetragonolobus]	4.00E-66	865
21	ACIDIC ENDOCHITINASE PRECURSOR	3.00E-68	878
22	chitinase [Hevea brasiliensis]	4.00E-70	881
23	chitinase [Hevea brasiliensis]	2.00E-62	884
24	ACIDIC ENDOCHITINASE PRECURSOR	1.00E-62	873
25	class III chitinase-like protein [Sesbania rostrata]	0.0002	881

26	ACIDIC ENDOCHITINASE PRECURSOR	6.00E-74	881
27	probable chitinase (EC 3.2.1.14) precursor - cucumber	4.00E-16	883
28	chitinase (EC 3.2.1.14) class III, acidic - soybean	7.00E-18	875
29	ACIDIC ENDOCHITINASE PRECURSOR	9.00E-72	887
30	class III acidic chitinase [Musa acuminata]	2.00E-47	879
31	ACIDIC ENDOCHITINASE PRECURSOR	7.00E-65	945
32	ACIDIC ENDOCHITINASE PRECURSOR	9.00E-45	870
33	ACIDIC ENDOCHITINASE PRECURSOR	9.00E-76	881
34	ACIDIC ENDOCHITINASE PRECURSOR	2.00E-62	881
35	ACIDIC ENDOCHITINASE PRECURSOR	2.00E-63	886
36	ACIDIC ENDOCHITINASE PRECURSOR	9.00E-72	878
37	chitinase [Hevea brasiliensis]	5.00E-40	906
38	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-73	902
39	chitinase 3-like protein precursor [Trichosanthes kirilowii]	3.00E-80	879
40	ACIDIC ENDOCHITINASE PRECURSOR	6.00E-70	873
41	chitinase [Hevea brasiliensis]	1.00E-48	882
42	ACIDIC ENDOCHITINASE PRECURSOR	2.00E-67	872
43	ACIDIC ENDOCHITINASE PRECURSOR	9.00E-72	890
44	class III chitinase-like protein [Sesbania rostrata]	0.0002	888
45	ACIDIC ENDOCHITINASE PRECURSOR	1.00E-22	892
46	chitinase [Hevea brasiliensis]	2.00E-43	875
47	chitinase 3-like protein precursor [Trichosanthes kirilowii]	4.00E-39	870
48	probable chitinase (EC 3.2.1.14) precursor - cucumber	2.00E-18	863
49	chitinase 3-like protein precursor [Trichosanthes kirilowii]	1.00E-07	875
50	chitinase [Hevea brasiliensis]	1.00E-45	865
51	chitinase 3-like protein precursor [Trichosanthes kirilowii]	6.00E-34	900
52	chitinase [Hevea brasiliensis]	3.00E-67	903
53	chitinase [Psophocarpus tetragonolobus]	2.00E-63	875
54	ACIDIC ENDOCHITINASE PRECURSOR	9.00E-82	864
55	probable chitinase (EC 3.2.1.14) precursor - cucumber	5.00E-14	898
56	chitinase 3-like protein precursor [Trichosanthes kirilowii]	5.00E-69	886
57	chitinase 3-like protein precursor [Trichosanthes kirilowii]	3.00E-39	885
58	ACIDIC ENDOCHITINASE PRECURSOR	4.00E-76	886
59	probable chitinase (EC 3.2.1.14) precursor - cucumber	5.00E-14	902
60	Glycosidase, Chitin Degradation, Multifunctional Enzyme	5.00E-70	900
61	ACIDIC ENDOCHITINASE PRECURSOR	4.00E-75	878
62	ACIDIC ENDOCHITINASE PRECURSOR	1.00E-28	987
63	chitinase [Hevea brasiliensis]	7.00E-40	865
64	ACIDIC ENDOCHITINASE PRECURSOR	2.00E-73	873
65	chitinase [Hevea brasiliensis]	2.00E-50	896
66	probable chitinase (EC 3.2.1.14) precursor - cucumber	1.00E-15	897
67	chitinase [Psophocarpus tetragonolobus]	7.00E-67	889
68	probable chitinase (EC 3.2.1.14) precursor - cucumber	3.00E-14	883
69	chitinase 3-like protein precursor [Trichosanthes kirilowii]	9.00E-40	894
70	chitinase 3-like protein precursor [Trichosanthes kirilowii]	6.00E-09	878
71	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-25	883
72	probable chitinase (EC 3.2.1.14) precursor - cucumber	5.00E-14	874
73	chitinase [Hevea brasiliensis]	7.00E-52	854
74	chitinase 3-like protein precursor [Trichosanthes kirilowii]	9.00E-40	871
75	chitinase 3-like protein precursor [Trichosanthes kirilowii]	8.00E-40	852
76	probable chitinase (EC 3.2.1.14) precursor - cucumber	2.00E-12	862
77	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-09	851
78	chitinase [Hevea brasiliensis]	7.00E-42	851
79	chitinase [Hevea brasiliensis]	2.00E-41	847
80	chitinase [Hevea brasiliensis]	4.00E-45	868
81	class III chitinase-like protein [Sesbania rostrata]	0.0004	815
82	chitinase 3-like protein precursor [Trichosanthes kirilowii]	1.00E-19	835
83	chitinase [Cucurbita moschata]	8.00E-07	888
84	probable chitinase (EC 3.2.1.14) precursor - cucumber	3.00E-13	857
85	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-38	868
86	class III chitinase-like protein [Sesbania rostrata]	7.00E-07	850
87	probable chitinase (EC 3.2.1.14) precursor - cucumber	4.00E-16	867
88	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-31	854
89	chitinase 3-like protein precursor [Trichosanthes kirilowii]	9.00E-40	859
90	chitinase 3-like protein precursor [Trichosanthes kirilowii]	9.00E-72	861
91	chitinase 3-like protein precursor [Trichosanthes kirilowii]	1.00E-38	873
92	chitinase homolog [Coffea arabica]	4.00E-45	898
93	probable chitinase (EC 3.2.1.14) precursor - cucumber	2.00E-09	902
94	chitinase 3-like protein precursor [Trichosanthes kirilowii]	1.00E-12	851
95	chitinase (EC 3.2.1.14) class III, acidic - soybean	0.0003	849
96	chitinase 3-like protein precursor [Trichosanthes kirilowii]	9.00E-21	859
97	chitinase [Psophocarpus tetragonolobus]	7.00E-79	840
98	probable chitinase (EC 3.2.1.14) precursor - cucumber	5.00E-16	843
99	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-30	833
100	chitinase [Hevea brasiliensis]	2.00E-50	825
101	probable chitinase (EC 3.2.1.14) precursor - cucumber	2.00E-14	864
102	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-31	831

103	chitinase 3-like protein precursor [Trichosanthes kirilowii]	1.00E-27	860
104	chitinase [Hevea brasiliensis]	1.00E-58	844
105	chitinase homolog [Coffea arabica]	2.00E-35	879
106	ACIDIC ENDOCHITINASE PRECURSOR	3.00E-75	869
107	ACIDIC ENDOCHITINASE PRECURSOR	1.00E-74	868
108	chitinase [Hevea brasiliensis]	1.00E-43	875
109	chitinase homolog [Coffea arabica]	6.00E-23	867
110	ACIDIC ENDOCHITINASE PRECURSOR	4.00E-68	865
111	probable chitinase (EC 3.2.1.14) precursor - cucumber	1.00E-24	867
112	chitinase 3-like protein precursor [Trichosanthes kirilowii]	3.00E-20	892
113	ACIDIC ENDOCHITINASE PRECURSOR	2.00E-73	861
114	ACIDIC ENDOCHITINASE PRECURSOR	1.00E-71	851
115	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-74	857
116	chitinase 3-like protein precursor [Trichosanthes kirilowii]	6.00E-60	868
117	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-62	850
118	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-34	903
119	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-32	859
120	chitinase 3-like protein precursor [Trichosanthes kirilowii]	4.00E-38	872
121	chitinase 3-like protein precursor [Trichosanthes kirilowii]	2.00E-20	853
122	chitinase 3-like protein precursor [Trichosanthes kirilowii]	1.00E-25	900
123	chitinase [Hevea brasiliensis]	2.00E-42	860
124	putative chitinase III [Oryza sativa (japonica cultivar-group)]	2.00E-15	858
125	chitinase homologue [Sesbania rostrata]	3.00E-16	894
126	probable chitinase (EC 3.2.1.14) precursor - cucumber	5.00E-11	847
127	chitinase (EC 3.2.1.14) Chib1 - soybean	1.00E-68	863
128	ACIDIC ENDOCHITINASE PRECURSOR	3.00E-57	848
129	chitinase [Hevea brasiliensis]	3.00E-42	886
130	chitinase [Hevea brasiliensis]	1.00E-58	853
Project Cytochrome p450			
ESTs NS1	**BLAST X**	**E-value**	**Size (bp)**
1	Trans-cinnamate 4-monooxygenase (Cinnamic acid 4-hydroxylase) (CA4H) (C4H) (P450C4H) (Cytochrome P450 73)	2.00E-52	881
2	cytochrome b5 domain-containing protein [Arabidopsis thaliana]	1.00E-10	855
3	cytochrome oxidase deficient homolog 1 [Homo sapiens]	3.00E-44	877
4	cytochrome P450-like protein [Arabidopsis thaliana]	3.00E-76	900
5	desaturase/cytochrome b5 protein [Ricinus communis]	2.00E-60	830
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	pectinesterase family [Arabidopsis thaliana]	2.00E-37	954
2	hypothetical protein At2g36300 [imported] - Arabidopsis thaliana	1.00E-69	952
3	hypothetical protein [Pseudomonas fluorescens PfO-1]	0	870
4	wound induced protein kinase [Nicotiana tabacum]	0	880
5	tsh protein - Escherichia coli	1.00E-40	929
6	PROBABLE TRANSMEMBRANE PROTEIN [Ralstonia solanacearum]	1.00E-76	918
7	hypothetical protein [Ralstonia metallidurans]	0.0001	948
8	DnaJ protein family [Arabidopsis thaliana]	2.00E-89	860
9	hypothetical protein [Pseudomonas syringae pv. syringae B728a]	3.00E-76	941
10	CONSERVED HYPOTHETICAL PROTEIN [Ralstonia solanacearum]	1.00E-52	932
11	hypothetical protein [Burkholderia fungorum]	9.00E-58	958
12	hypothetical protein [Burkholderia fungorum]	8.00E-10	885
13	Photosystem I reaction center subunit II, chloroplast precursor (Photosystem I 20 kDa subunit) (PSI-D) (PS I subunit 5)	1.00E-84	988
14	sensory box histidine kinase/response regulator [Pseudomonas syringae pv. tomato str. DC3000]	1.00E-86	861
15	unknown [Arabidopsis thaliana]	3.00E-69	944
16	putative cytochrome P450 [Arabidopsis thaliana]	2.00E-23	964
17	cytochrome b561 family protein [Pseudomonas syringae pv. tomato str. DC3000]	2.00E-09	945
18	cytochrome b5 domain-containing protein [Arabidopsis thaliana]	3.00E-25	837
19	Thiamine pyrophosphate-requiring enzymes [Thermoanaerobacter tengcongensis]	3.00E-21	901
20	Thiamine pyrophosphate-requiring enzymes [Thermoanaerobacter tengcongensis]	3.00E-21	911
21	cytochrome P450, putative [Arabidopsis thaliana]	2.00E-79	944
22	cytochrome P-450 [Lithospermum erythrorhizon]	3.00E-95	865
23	elicitor-inducible cytochrome P450 [Nicotiana tabacum]	2.00E-73	890
24	NADPH-cytochrome P-450 reductase [Ophiorrhiza pumila]	6.00E-90	949
25	cytochrome BD2, subunit II [Vibrio parahaemolyticus RIMD 2210633]	5.00E-21	1037
26	cytochrome BD2 subunit II [Salmonella typhimurium LT2]	3.00E-35	949
27	cytochrome BD2 subunit II [Salmonella typhimurium LT2]	2.00E-33	951
28	Trans-cinnamate 4-monooxygenase (Cinnamic acid 4-hydroxylase) (CA4H) (C4H) (P450C4H) (Cytochrome P450 73)	3.00E-98	824
29	cytochrome oxidase subunit I [Rhodothermus marinus]	1.00E-35	802
30	cytochrome P450, putative [Arabidopsis thaliana]	4.00E-72	968
31	Cytochrome c-type biogenesis protein [Methanosarcina mazei Goe1]	0.002	981
32	cytochrome b561 family protein [Pseudomonas syringae pv. tomato str. DC3000]	6.00E-08	910
33	cytochrome b561 family protein [Pseudomonas syringae pv. tomato str. DC3000]	3.00E-07	1062
34	Allene oxide synthase, chloroplast precursor (Hydroperoxide dehydrase) (Cytochrome P450 74A)	3.00E-68	928
35	L-lactate dehydrogenase [Escherichia coli K12]	3.00E-52	940
36	L-lactate dehydrogenase [Escherichia coli K12]	2.00E-52	918
37	Cytochrome P450 77A2 (CYPLXXVIIA2) (P-450EG5)	0	953
38	cytochrome P450 [Solanum tuberosum]	2.00E-21	903
39	cytochrome P450 [Solanum tuberosum]	1.00E-52	923
40	cytochrome b561 family protein [Pseudomonas syringae pv. tomato str. DC3000]	9.00E-10	947
41	NADPH-cytochrome P-450 reductase [Ophiorrhiza pumila]	8.00E-06	1090

42	NADPH-cytochrome P-450 reductase [Ophiorrhiza pumila]	2.00E-58	882
43	Cytochrome c-type biogenesis protein [Methanosarcina mazei Goe1]	0.002	895
44	nodulin MtN3 family protein [Arabidopsis thaliana]	2.00E-29	900
45	Cytochrome P450 77A2 (CYPLXXVIIA2) (P-450EG5)	1.00E-74	907
46	putative cytochrome P450 [Arabidopsis thaliana]	8.00E-21	893
47	cytochrome c family protein [Pseudomonas putida KT2440]	1.00E-31	908
48	cytochrome c oxidase subunit (cbb3-type) [Pseudomonas aeruginosa PA01]	3.00E-92	853
49	nitrate reductase 1, cytochrome b(NR), gamma subunit [Shigella flexneri 2a str. 2457T]	6.00E-16	905
50	Cytochrome P450 82A2 (P450 CP4)	4.00E-14	877
51	cytochrome P450 [Solanum tuberosum]	5.00E-64	920
52	cytochrome b561 family protein [Pseudomonas syringae pv. tomato str. DC3000]	4.00E-08	982
53	L-lactate dehydrogenase [Escherichia coli K12]	2.00E-55	866
54	cytochrome P450, putative [Arabidopsis thaliana]	6.00E-65	962
55	putative cytochrome P450 [Solanum tuberosum]	2.00E-06	853
56	60S RIBOSOMAL PROTEIN L10-3 (QM/R22)	0	948
57	cytochrome b5 domain-containing protein [Arabidopsis thaliana]	1.00E-36	942
58	NADPH-ferrihemoprotein reductase (EC 1.6.2.4) 1 - parsley	3.00E-17	1092
59	cytochrome P450 [Petunia x hybrida]	3.00E-62	938
60	heme exporter protein C; cytochrome C-type biogenesis protein [Xanthomonas axonopodis pv. citri str. 306]	6.00E-44	930
61	Allene oxide synthase, chloroplast precursor (Hydroperoxide dehydrase) (Cytochrome P450 74A)	5.00E-64	933
62	NADPH-cytochrome P-450 reductase [Ophiorrhiza pumila]	5.00E-53	955
63	probable thiol:disulfide interchange protein [Pseudomonas aeruginosa PA01]	2.00E-52	952
64	Cytochrome c-type biogenesis protein [Methanosarcina mazei Goe1]	0.002	963
65	allene oxide synthase / cytochrome P450 74A [Arabidopsis thaliana]	3.00E-81	934
66	putative cytochrome B 561 [Craterostigma plantagineum]	4.00E-38	829
67	probable cytochrome c [Nitrosomonas europaea ATCC 19718]	2.00E-42	859
68	allene oxide synthase / cytochrome P450 74A [Arabidopsis thaliana]	3.00E-06	908
69	NADPH-cytochrome P450 oxidoreductase (EC 1.-.-.-) - common tobacco	3.00E-85	1000
70	expressed protein [Arabidopsis thaliana]	3.00E-23	936
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	No Hits Found	0	894
2	No Hits Found	0	902
3	No Hits Found	0	894
4	No Hits Found	0	870
5	No Hits Found	0	930
6	No Hits Found	0	962
7	No Hits Found	0	988
8	No Hits Found	0	955
9	No Hits Found	0	923
10	No Hits Found	0	919
11	No Hits Found	0	897
12	No Hits Found	0	847
13	No Hits Found	0	965
14	No Hits Found	0	891
15	No Hits Found	0	963
16	60S RIBOSOMAL PROTEIN L10-3 (QM/R22)	0	968
17	NADPH-cytochrome P450 reductase (CPR) (P450R)	1.00E-22	945
18	cytochrome P450 -related [Arabidopsis thaliana]	4.00E-06	968
19	cytochrome b5 domain-containing protein [Arabidopsis thaliana]	6.00E-18	989
20	putative cytochrome c oxidoreductase [Arabidopsis thaliana]	5.00E-23	956
21	cytochrome P450 [Solanum tuberosum]	5.00E-62	959
22	cytochrome P450 monooxygenase CYP72A5 [Zea mays subsp. mays]	5.00E-33	958
23	cytochrome P450 [Solanum tuberosum]	3.00E-28	981
24	cytochrome P450, putative [Arabidopsis thaliana]	5.00E-49	978
25	cytochrome c551 peroxidase, putative [Pseudomonas putida KT2440]	6.00E-06	942
26	cytochrome b5 [Arabidopsis thaliana]	5.00E-26	891
27	cytochrome P450 [Catharanthus roseus]	2.00E-17	1017
28	NADPH-cytochrome P-450 reductase [Ophiorrhiza pumila]	3.00E-50	881
29	fatty acid desaturase/cytochrome b5 fusion protein [Arabidopsis thaliana]	5.00E-42	826
30	putative cytochrome c oxidoreductase [Arabidopsis thaliana]	4.00E-18	899
31	allene oxide synthase / cytochrome P450 74A [Arabidopsis thaliana]	6.00E-62	915
32	Ubiquinol-cytochrome C reductase complex 14 kDa protein (CR14)	9.00E-37	924
33	cytochrome c551 peroxidase, putative [Pseudomonas putida KT2440]	7.00E-06	982
34	cytochrome P450 ent-kaurene oxidase (GA3) [Arabidopsis thaliana]	8.00E-62	958
35	Allene oxide synthase, chloroplast precursor (Hydroperoxide dehydrase) (Cytochrome P450 74A)	8.00E-65	943
36	putative cytochrome P450 [Arabidopsis thaliana]	9.00E-72	938
37	Putative Cytochrome P450 [Oryza sativa]	1.00E-39	891
38	Allene oxide synthase (Rubber particle protein) (RPP)	2.00E-23	916
39	cytochrome P450 ent-kaurene oxidase (GA3) [Arabidopsis thaliana]	4.00E-79	955
40	Cytochrome c	2.00E-56	969
41	Allene oxide synthase, chloroplast precursor (Hydroperoxide dehydrase) (Cytochrome P450 74A)	3.00E-75	929
42	cytochrome c551 peroxidase, putative [Pseudomonas putida KT2440]	6.00E-06	928
43	cytochrome P450 [Solanum tuberosum]	1.00E-28	992
44	expressed protein [Arabidopsis thaliana]	9.00E-16	839
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	No Hits Found	0	889
2	Ubiquinol-cytochrome C reductase complex 14 kDa protein (CR14)	9.00E-37	854

3	putative NADPH-cytochrome P450 reductase [Pisum sativum]	3.00E-14	838
4	Ubiquinol-cytochrome C reductase complex 14 kDa protein (CR14)	2.00E-16	875
5	putative cytochrome c oxidoreductase [Arabidopsis thaliana]	3.00E-40	871
	Project Glucanase		
ESTs NS1	**BLAST X**	**E-value**	**Size (bp)**
1	glycosyl hydrolase family 17 [Arabidopsis thaliana]	2.00E-21	875
2	putative glucanase [Oryza sativa (japonica cultivar-group)]	4.00E-14	832
3	probable glucanase - Arabidopsis thaliana	2.00E-37	893
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	xyloglucan endotransglycosylase, putative [Arabidopsis thaliana]	2.00E-68	962
2	putative beta-1, 3-glucanase [Oryza sativa (japonica cultivar-group)]	1.00E-58	910
3	Glucan endo-1, 3-beta-glucosidase, acidic isoform PR-Q' precursor ((1->3)-beta-glucan endohydrolase) ((1->3)-beta-glucanase) (Beta-1, 3-endoglucanase) (PR-35)	5.00E-80	892
4	glycosyl hydrolase family 17 (beta-1, 3-glucanase) [Arabidopsis thaliana]	3.00E-45	938
5	glucan endo-1, 3-beta-D-glucosidase (EC 3.2.1.39) [imported] - garden pea	2.00E-25	918
6	glycosyl hydrolase family 17 [Arabidopsis thaliana]	4.00E-65	854
7	glucan endo-1, 3-beta-D-glucosidase (EC 3.2.1.39) [imported] - garden pea	2.00E-76	877
8	xyloglucan endotransglycosylase, putative [Arabidopsis thaliana]	3.00E-60	941
9	putative beta-1, 3-glucanase [Oryza sativa (japonica cultivar-group)]	1.00E-71	877
10	glycosyl hydrolase family 9 [Arabidopsis thaliana]	9.00E-56	935
11	glucan endo-1, 3-beta-D-glucosidase (EC 3.2.1.39) [imported] - garden pea	3.00E-42	852
12	glycosyl hydrolase family 9 [Arabidopsis thaliana]	2.00E-56	959
13	1, 3-beta-glucanase (EC 3.2.1.-), basic - tomato	4.00E-73	1009
14	endo-beta-1, 4-D-glucanase [Lycopersicon esculentum]	3.00E-81	925
15	xyloglucan endo-1, 4-beta-D-glucanase (EC 3.2.1.-) F6H11.140 - Arabidopsis thaliana	2.00E-08	1077
16	putative beta-1, 3-glucanase [Oryza sativa (japonica cultivar-group)]	6.00E-86	935
17	xyloglucan endo-1, 4-beta-D-glucanase (EC 3.2.1.-) F6H11.140 - Arabidopsis thaliana	3.00E-07	895
18	xyloglucan endotransglycosylase, putative [Arabidopsis thaliana]	2.00E-13	853
19	xyloglucan endo-1, 4-beta-D-glucanase (EC 3.2.1.-) F6H11.140 - Arabidopsis thaliana	3.00E-50	901
20	xyloglucan endotransglycosylase, putative [Arabidopsis thaliana]	1.00E-07	801
21	putative exoglucanase precursor [Oryza sativa]	6.00E-56	879
22	xyloglucan endo-1, 4-beta-D-glucanase (EC 3.2.1.-) F6H11.140 - Arabidopsis thaliana	4.00E-34	935
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	glycosyl hydrolase family 17 [Arabidopsis thaliana]	2.00E-20	886
2	xyloglucan endo-1, 4-beta-D-glucanase (EC 3.2.1.-) (clone NXG1) - common nasturtium	3.00E-93	921
3	glucanase -related [Arabidopsis thaliana]	6.00E-19	982
4	contains similarity to endo-1, 3-1, 4-beta-D-glucanase~gene_id:MDB19.5 [Arabidopsis thaliana]	3.00E-18	999
5	1, 3-beta-glucanase (EC 3.2.1.-), acidic - tomato	7.00E-59	970
6	endo-1, 3-beta-glucanase-like protein [Pyrus pyrifolia]	1.00E-84	946
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	beta-1, 3 glucanase-like protein [Oryza sativa (japonica cultivar-group)]	2.00E-20	874
2	glucan endo-1, 3-beta-D-glucosidase (EC 3.2.1.39) [imported] - garden pea	5.00E-76	876
	Project HSP		
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	tetratricopeptide repeat (TPR)-containing protein [Arabidopsis thaliana]	1.00E-41	984
2	tetratricopeptide repeat (TPR)-containing protein [Arabidopsis thaliana]	9.00E-44	927
3	heat shock protein hsp70	7.00E-72	880
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	heat shock protein family [Arabidopsis thaliana]	4.00E-34	977
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	heat shock protein hsp70b [Arabidopsis thaliana]	0.0004	909
2	cytosolic class I small heat-shock protein HSP17.5 [Castanea sativa]	2.00E-41	892
3	tetratricopeptide repeat (TPR)-containing protein [Arabidopsis thaliana]	1.00E-21	855
4	17.6 KD CLASS I HEAT SHOCK PROTEIN (HSP 17.6-L)	7.00E-05	853
	Project Thaumatin		
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	putative thaumatin-like protein [Vitis vinifera]	2.00E-10	995
2	thaumatin-like protein [Vitis vinifera]	1.00E-89	991
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	putative thaumatin [Arabidopsis thaliana]	1.00E-95	864
	Project Chalconesynthase		
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	CHALCONE SYNTHASE 2 (NARINGENIN-CHALCONE SYNTHASE 2)	2.00E-59	915
2	CHALCONE SYNTHASE 2 (NARINGENIN-CHALCONE SYNTHASE 2)	0	838
3	CHALCONE SYNTHASE 2 (NARINGENIN-CHALCONE SYNTHASE 2)	0	936
	Project Pathogenesis		
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	subtilisin-like proteinase (EC 3.4.21.-) precursor P69B, pathogenesis-related - tomato	4.00E-09	958
2	Similar to Arabidopsis thaliana putative pathogenesis-related protein (U20347) [Oryza sativa (japonica cultivar-group)]	2.00E-16	838
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	pathogenesis-related protein 1 - parsley	9.00E-06	986
2	pathogenesis-related protein 10 [Vitis vinifera]	2.00E-36	966
3	pathogenesis-related protein 10 [Vitis vinifera]	5.00E-32	945
4	Acidic endochitinase Q precursor (Pathogenesis-related protein Q) (PR-Q)	3.00E-13	967
5	Pathogenesis-related protein R major form precursor (Thaumatin-like protein E22)	2.00E-09	995
6	pathogenesis-related protein [Oryza sativa]	2.00E-28	972
7	pathogenesis-related protein 5-1 [Helianthus annuus]	5.00E-86	991
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**

1	thaumatin family [Arabidopsis thaliana]	4.00E-48	864
Project Polyphenoloxidase			
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	catechol oxidase; polyphenol oxidase [Ipomoea batatas]	3.00E-71	893
Project Importin			
ESTs NS1	**BLAST X**	**E-value**	**Size (bp)**
1	importin alpha subunit [Arabidopsis thaliana]	7.00E-12	901
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	importin alpha 2 [Capsicum annuum]	5.00E-96	925
2	cellular apoptosis susceptibility protein (importin-alpha re-exporter), putative [Arabidopsis thaliana]	1.00E-33	1048
3	rubisco small subunit [Coffea arabica]	1.00E-93	924
4	rubisco small subunit [Coffea arabica]	1.00E-89	909
5	rubisco small subunit [Coffea arabica]	1.00E-93	888
6	rubisco small subunit [Coffea arabica]	1.00E-93	961
7	rubisco small subunit [Coffea arabica]	3.00E-91	907
8	rubisco small subunit [Coffea arabica]	4.00E-93	912
9	rubisco small subunit [Coffea arabica]	1.00E-93	989
10	rubisco small subunit [Coffea arabica]	1.00E-93	962
11	rubisco small subunit [Coffea arabica]	1.00E-93	903
12	rubisco small subunit [Coffea arabica]	1.00E-93	902
13	rubisco small subunit [Coffea arabica]	1.00E-93	974
14	rubisco small subunit [Coffea arabica]	1.00E-93	847
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	AtKAP alpha [Arabidopsis thaliana]	5.00E-94	908
2	rubisco small subunit [Coffea arabica]	7.00E-93	939
3	rubisco small subunit [Coffea arabica]	1.00E-93	957
4	No Hits Found	0	905
5	No Hits Found	0	941
6	No Hits Found	0	892
7	No Hits Found	0	914
8	rubisco small subunit [Coffea arabica]	1.00E-90	946
9	rubisco small subunit [Coffea arabica]	6.00E-78	980
10	No Hits Found	0	927
11	No Hits Found	0	883
12	No Hits Found	0	849
13	No Hits Found	0	937
14	No Hits Found	0	886
15	No Hits Found	0	902
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	probable nuclear transport factor importin alpha [imported] - Arabidopsis thaliana	2.00E-13	853
2	rubisco small subunit [Coffea arabica]	1.00E-93	846
3	rubisco small subunit [Coffea arabica]	1.00E-93	884
4	rubisco small subunit [Coffea arabica]	2.00E-54	869
Project Glucosyltransferase			
ESTs NS1	**BLAST X**	**E-value**	**Size (bp)**
1	UDP-glucose:salicylic acid glucosyltransferase [Nicotiana tabacum]	1.00E-75	881
2	putative glucosyltransferase [Lycopersicon esculentum]	9.00E-25	802
3	Sucrose synthase (Sucrose-UDP glucosyltransferase)	0	904
ESTs RM1	**BLAST X**	**E-value**	**Size (bp)**
1	Sucrose synthase (Sucrose-UDP glucosyltransferase) (SS65)	3.00E-93	926
2	glucosyltransferase-8 [Vigna angularis]	0.006	928
3	Sucrose synthase (Sucrose-UDP glucosyltransferase) (SS65)	3.00E-98	911
4	Putative indole-3-acetate beta-glucosyltransferase [Oryza sativa (japonica cultivar-group)]	1.00E-35	860
5	Sucrose synthase (Sucrose-UDP glucosyltransferase) (SS65)	0	929
6	glycosyltransferase family [Arabidopsis thaliana]	4.00E-42	871
7	glucosyl transferase [Xanthomonas axonopodis pv. citri str. 306]	6.00E-57	917
8	probable endo-xyloglucan transferase - upland cotton (fragment)	2.00E-64	941
9	alpha2-glucosyltransferase [Homo sapiens]	3.00E-29	942
10	glucosyltransferase NTGT3 [Nicotiana tabacum]	1.00E-53	960
11	probable endo-xyloglucan transferase - upland cotton (fragment)	3.00E-70	962
12	probable endo-xyloglucan transferase - upland cotton (fragment)	4.00E-37	935
13	alpha, alpha-trehalose-phosphate synthase, UDP-forming, putative [Arabidopsis thaliana]	5.00E-73	927
14	alpha, alpha-trehalose-phosphate synthase, UDP-forming, putative [Arabidopsis thaliana]	8.00E-32	1025
15	probable endo-xyloglucan transferase - upland cotton (fragment)	9.00E-12	1077
16	alpha, alpha-trehalose-phosphate synthase, UDP-forming, putative [Arabidopsis thaliana]	3.00E-90	941
17	glycosyltransferase family 2 [Arabidopsis thaliana]	0	962
18	Sucrose synthase (Sucrose-UDP glucosyltransferase) (SS65)	0	858
19	glycosyltransferase family 2 [Arabidopsis thaliana]	5.00E-22	964
20	probable endo-xyloglucan transferase - upland cotton (fragment)	0	918
21	glycosyltransferase family 20 [Arabidopsis thaliana]	2.00E-53	921
22	Limonoid UDP-glucosyltransferase (Limonoid glucosyltransferase) (Limonoid GTase) (LGTase)	6.00E-78	936
23	alpha, alpha-trehalose-phosphate synthase, UDP-forming, putative [Arabidopsis thaliana]	1.00E-68	878
24	probable endo-xyloglucan transferase - upland cotton (fragment)	4.00E-55	901
25	hypothetical protein [Pseudomonas syringae pv. syringae B728a]	4.00E-62	1012
26	ribosomal protein L7	1.00E-29	916
27	theobromine synthase [Coffea arabica]	0	924
28	expressed protein [Arabidopsis thaliana]	7.00E-64	828

29	auxin-regulated protein [Arabidopsis thaliana]	9.00E-79	940
30	putative protein [Arabidopsis thaliana]	7.00E-96	951
31	secretory peroxidase [Avicennia marina]	0	936
32	polygalacturonase (EC 3.2.1.15) 1 beta chain precursor - tomato	2.00E-69	999
33	No Hits Found	0	861
34	DEAD box RNA helicase, putative [Arabidopsis thaliana]	0	880
35	expressed protein [Arabidopsis thaliana]	4.00E-14	891
36	No Hits Found	0	855
37	GENERAL L-AMINO ACID TRANSPORT ATP-BINDING PROTEIN AAPP [Brucella melitensis]	4.00E-61	888
38	expressed protein [Arabidopsis thaliana]	3.00E-69	879
39	No Hits Found	0	947
40	No Hits Found	0	918
41	O-sialoglycoprotein endopeptidase [Pseudomonas putida KT2440]	4.00E-66	937
42	sensor histidine kinase [Pseudomonas putida KT2440]	1.00E-09	918
43	DHHC-type zinc finger domain-containing protein [Arabidopsis thaliana]	0	952
44	hypothetical protein [Pseudomonas fluorescens PfO-1]	4.00E-83	924
45	protein phosphatase 2C (PP2C), putative [Arabidopsis thaliana]	1.00E-24	966
46	expressed protein [Arabidopsis thaliana]	1.00E-49	958
47	hypothetical protein [Pseudomonas aeruginosa UCBPP-PA14]	4.00E-08	947
48	hypothetical protein [Azotobacter vinelandii]	7.00E-52	940
49	hypothetical protein [Pseudomonas fluorescens PfO-1]	9.00E-33	870
50	alpha subunit of aromatic ring hydroxylase component of chlorobenzoate 1, 2-dioxygenase [Burkholderia sp. NK8]	4.00E-94	964
ESTs RX1	**BLAST X**	**E-value**	**Size (bp)**
1	Sucrose synthase (Sucrose-UDP glucosyltransferase)	0	957
2	Sucrose synthase (Sucrose-UDP glucosyltransferase) (SS65)	0	918
3	Sucrose synthase (Sucrose-UDP glucosyltransferase) (SS65)	0	983
4	Sucrose synthase (Sucrose-UDP glucosyltransferase)	0	939
5	glucosyltransferase [Nicotiana tabacum]	6.00E-50	932
6	hypothetical protein T16E15.1 - Arabidopsis thaliana	4.00E-17	943
7	glycosyltransferase family 20 [Arabidopsis thaliana]	0	974
8	Sucrose synthase (Sucrose-UDP glucosyltransferase)	0	929
9	glucuronosyl transferase homolog, ripening-related - tomato (fragment)	4.00E-77	946
10	Sucrose synthase (Sucrose-UDP glucosyltransferase)	0	929
11	Hydroquinone glucosyltransferase (Arbutin synthase)	2.00E-48	945
12	indole-3-acetate beta-glucosyltransferase -related [Arabidopsis thaliana]	3.00E-24	890
13	Sucrose synthase (Sucrose-UDP glucosyltransferase)	2.00E-79	895
14	glucosyltransferase NTGT2 [Nicotiana tabacum]	3.00E-31	877
15	glycosyltransferase family [Arabidopsis thaliana]	2.00E-38	883
16	alpha, alpha-trehalose-phosphate synthase, UDP-forming (trehalose-6-phosphate synthase/UDP-glucose-glucosephosphate glucosyltransferase), putative [Arabidopsis thaliana]	2.00E-54	936
17	No Hits Found	0	984
18	No Hits Found	0	954
19	No Hits Found	0	910
20	No Hits Found	0	893
21	No Hits Found	0	891
22	No Hits Found	0	891
23	No Hits Found	0	869
24	No Hits Found	0	951
25	No Hits Found	0	896
26	No Hits Found	0	920
27	No Hits Found	0	855
28	No Hits Found	0	943
ESTs SS1	**BLAST X**	**E-value**	**Size (bp)**
1	putative glucosyltransferase [Arabidopsis thaliana]	2.00E-33	850
2	No Hits Found	0	903

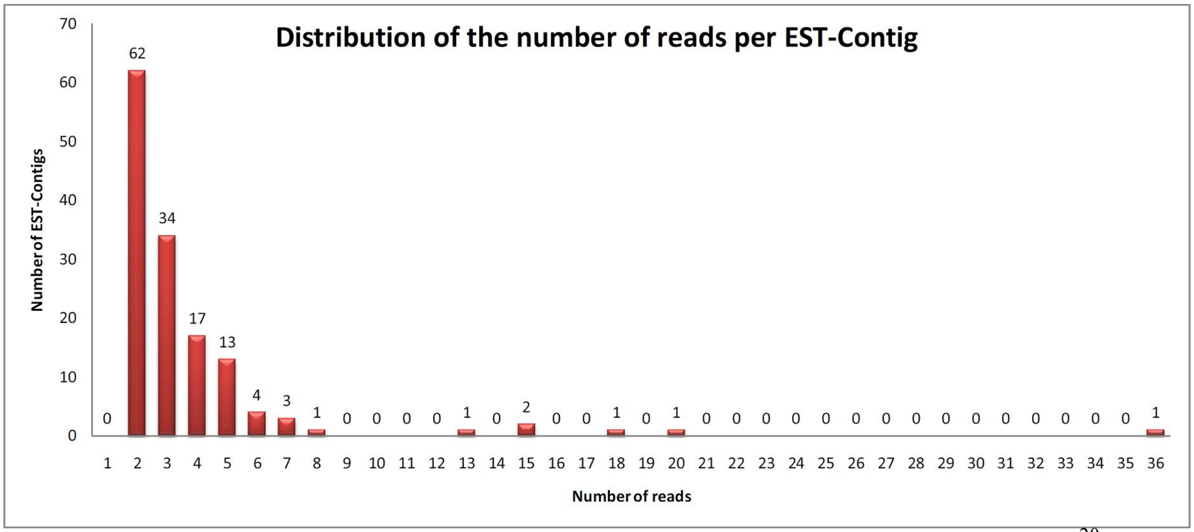


Figure S1: Distribution of the number of reads in the 140 EST-Contigs with e-value < e^{-20} and score > 100 formed after clustering.

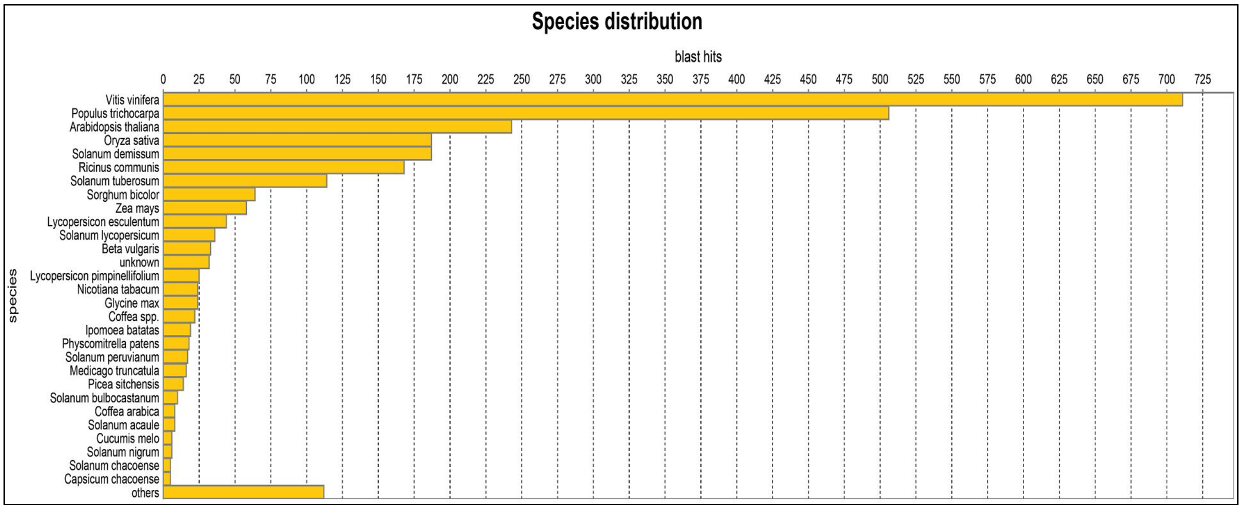


Figure S2: Species with most blast hits with the EST-Contigs

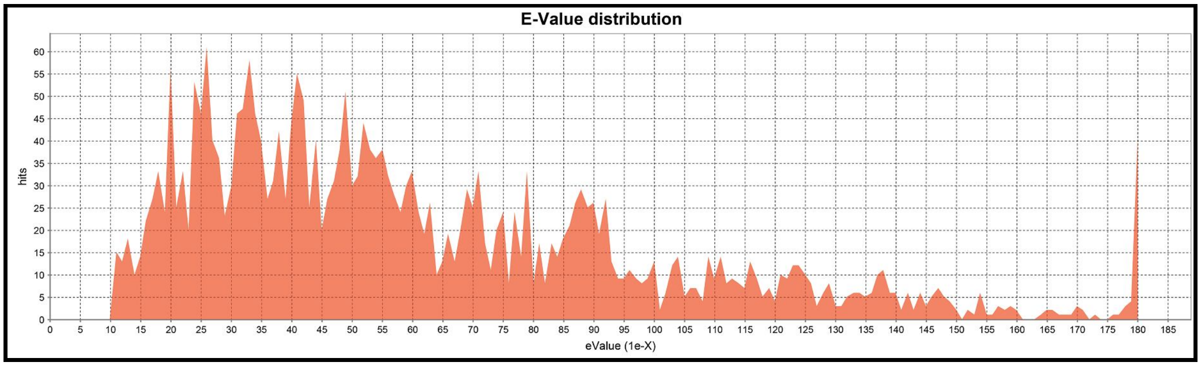


Figure S3: Average e-value distribution of the best 20 hits of the 140 EST-Contigs submitted to BlastX by Blast2GO

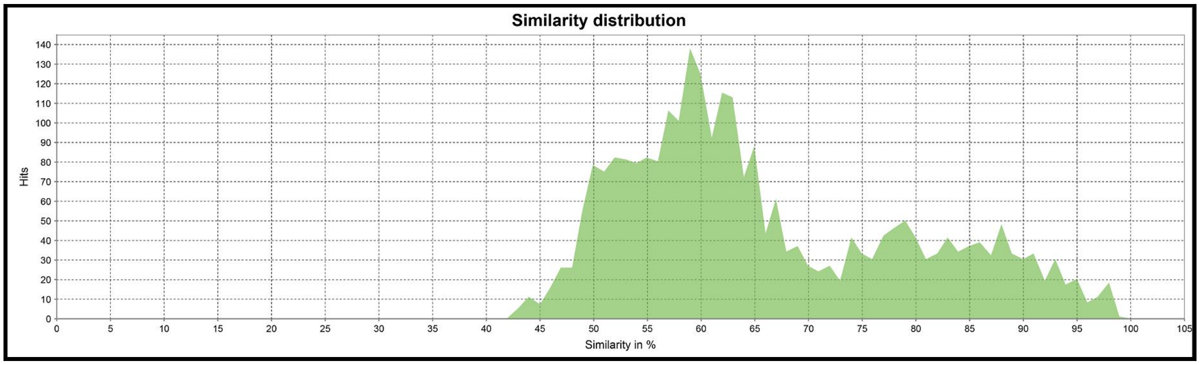


Figure S4: Average similarity value distribution of the best 20 hitsof the 140 EST-Contigs submitted to BlastX by Blast2GO

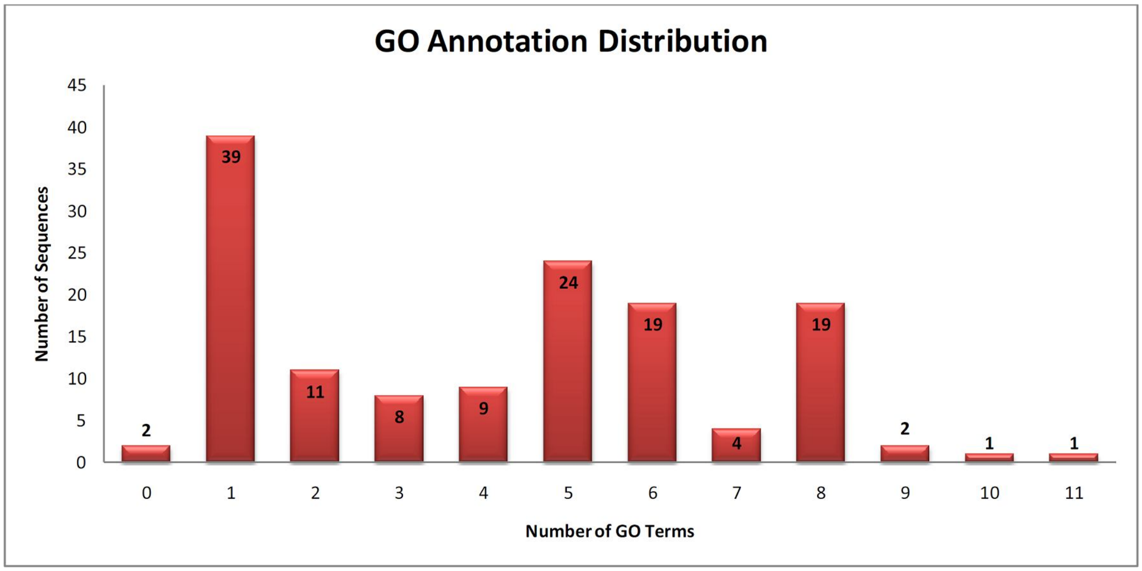


Figure S5: Distribution of the number of GO terms (Cellular Component, Molecular Function and Biological Process) from the 140 EST-Contigs analyzed by Blast2GO.

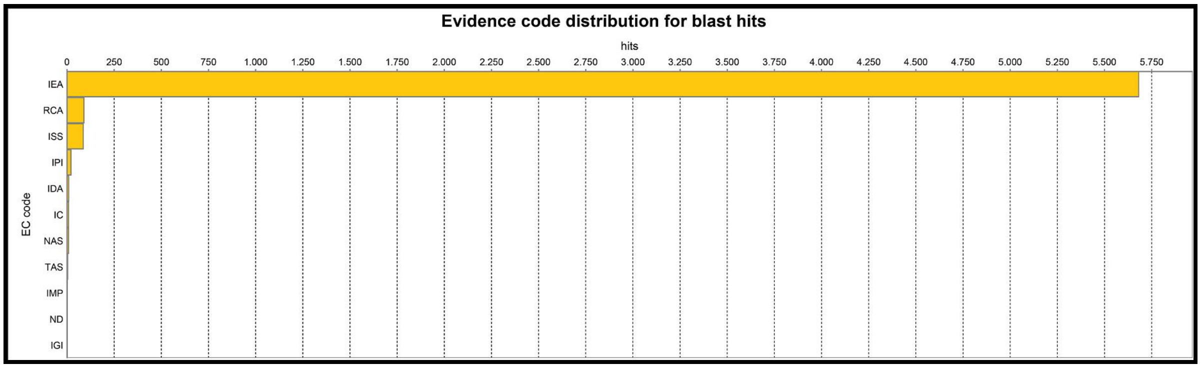


Figure S6: EC (Evidence Code) distribution for the EST-Contigs blast hits. IEA = Inferred from Electronic Annotation. ISS = Inferred from Sequence or Structural Similarity. RCA = Inferred from Reviewed Computational Analysis. IPI = Inferred from Physical Interaction. IDA = Inferred from Direct Assay. IC = Inferred by Curator. NAS = Non-traceable Author Statement. TAS = Traceable Author Statement. ND = No biological Data available. IMP = Inferred from Mutant Phenotype. IGI = Inferred from Genetic Interaction.