Genetics and Molecular Biology, 30, 3 (suppl), 931-942 (2007)

www.sbg.org.br

Research Article

Putative resistance genes in the CitEST database

Simone Guidetti-Gonzalez and Helaine Carrer

Departamento de Ciências Biológicas, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, SP, Brazil.

Abstract

Disease resistance in plants is usually associated with the activation of a wide variety of defense responses to prevent pathogen replication and/or movement. The ability of the host plant to recognize the pathogen and to activate defense responses is regulated by direct or indirect interaction between the products of plant resistance (*R*) and pathogen avirulence (*Avr*) genes. Attempted infection of plants by avirulent pathogens elicits a battery of defenses often followed by the collapse of the challenged host cells. Localized host cell death may help to prevent the pathogen from spreading to uninfected tissues, known as hypersensitive response (HR). When either the plant or the pathogen lacks its cognate gene, activation of the plant's defense responses fails to occur or is delayed and does not prevent pathogen colonization. In the CitEST database, we identified 1,300 reads related to *R* genes in *Citrus* which have been reported in other plant species. These reads were translated *in silico*, and alignments of their amino acid sequences revealed the presence of characteristic domains and motifs that are specific to *R* gene classes. The description of the reads identified suggests that they function as resistance genes in citrus.

Key words: hypersensitive response (HR), plant disease resistance, *Citrus,* EST sequences.

Received: July 21, 2006; Accepted: February 28, 2007.

Introduction

During their life cycle, plants are subjected to numerous and diverse threats from the outside environment. Infections by pathogenic fungi, bacteria, and viruses are among the most serious menaces that plants have to cope with (Wojtaszek, 1997). Since plants are sessile, they have developed a broad range of strategies, including genetic mechanisms to react and to protect themselves against biotic and abiotic stresses.

Genetic control of plant disease resistance often relies on the simultaneous occurrence of a resistance (*R*) gene in the plant genome and a specific corresponding avirulence (*Avr*) gene in the pathogen genome (Flor, 1971). The resistance provided by these genes is highly specific and effective only against pathogens expressing a corresponding avirulence gene. These observations are consistent with *R* genes encoding receptors that detect, directly or indirectly, the products of the pathogen *Avr* genes (Dangl and Jones, 2001). Upon pathogen recognition, the R proteins trigger defenses that often result in a hypersensitive response (HR), which leads to a rapid induction of host cell death at the site of the pathogen invasion (Noutoshi *et al.*, 2005). This plant response is associated with massive cellular ion fluxes, generation of reactive oxygen species, cell wall strengthening, and also with the expression of many defense proteins, including pathogenesis-related (PR) proteins (Heath, 2000).

The largest group of *R* genes encodes cytoplasmic proteins containing a central nucleotide binding (NBS) and a carboxyl Leu-rich repeat (LRR) domain, encoded by the *NBS-LRR* genes (Tör *et al.*, 2004). This protein group is subdivided into two major subclasses: (1) those containing an amino-terminal coiled-coil (CC) domain (CC-NBS-LRR), such as RPS2, RPM1, RPS5, RPP13 (Bittner-Eddy *et al.*, 2000) and RPP8 (McDowell *et al.*, 1998), and (2) those containing an amino-terminal domain resembling the cytoplasmic signaling domain of the Toll and Interleukin-1 (TIR) transmembrane receptors (TIR-NBS-LRR), such as RPS4, RPP1, RPP5 and N (Whitham *et al.*, 1994; Dangl and Jones, 2001; Tör *et al.*, 2004). In addition, the TIR-NBS-LRR group can be divided into two subgroups depending on the presence of a C-terminal non-LRR (CNL) domain (Dodds *et al.*, 2001).

The TIR and non-TIR NBS-LRR sequences can be distinguished by motifs internal to their NBS-domains or by a single amino acid residue in the final portion of the Kinase-2 motif, which invariably is an aspartic acid in the TIR NBS-LRR subclass, and a tryptophan in the non-TIR NBS-LRR subclass (Meyers *et al.*, 1999). The TIR se-

Send correspondence to Helaine Carrer. Departamento de Ciências Biológicas, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Av. Pádua Dias 11, 13418-900 Piracicaba, SP, Brazil. E-mail: hecarrer@esalq.usp.br.

quences are present in dicot species, whereas the non-TIR sequences have been reported throughout the angiosperms (Meyers *et al.*, 1999; Pan *et al.*, 2000). Studies have emphasized the importance of these specific domains for resistance. The LRR-kinase receptor proteins have been assigned functions in normal plant development and hormone perception as well as *R* function (Trotochaud *et al.*, 1999; Wang *et al.*, 2001). In contrast, the NBS-LRR class has been genetically linked to disease-resistance function (Nimchuk *et al.*, 2003). Other R structures may have been derived from protein families with pleiotropic functions in plant growth and development (Nimchuk *et al.*, 2003).

The second group of *R* genes contains the cytoplasmic serine-threonine kinase represented by the *Pto* genes. These genes confer resistance to the bacterial pathogen *Pseudomonas syringae* pv *tomato* (Martin *et al.*, 1993). The third group of *R* genes encodes the receptor-like kinases (RLKs) and they contain an extracellular LRR domain with a single transmembrane spanning region and a cytoplasmic kinase domain (Tör *et al.*, 2004). The resistance gene *Xa21*, which confers resistance to *Xanthomonas oryzae* pv. *oryzae* in rice (Song *et al.*, 1995), belongs to this group (Tör *et al.*, 2004). The *Arabidopsis* genome contains 174 sequences with homology to transmembrane kinases, but only one has an assigned role in resistance (The Arabidopsis Genome Initiative, 2000).

The *Cf* gene family and *HcrVf2* are examples in the fourth group of *R* genes, which are receptor-like proteins (RLPs) (Tör *et al.*, 2004). They are similar to the RLK genes in that they encode extracellular LRRs and a C-terminal membrane anchor but lack the cytoplasmic kinase domain (Dixon *et al.*, 1996).

In the present work, putative *R* genes were identified in Citrus from searches in the Citrus EST (CitEST) database.

Materials and Methods

The Citrus EST database (CitEST) was developed at the Centro APTA Citros 'Sylvio Moreira', in São Paulo State, Brazil. It consists of EST (expressed sequence tag) sequences, prepared from mRNAs retro-transcribed (cDNAs) from diverse citrus species, grown under different conditions (details of cDNA sequences are described in Targon *et al.*, this issue). The CitEST database was surveyed aiming to identify putative *R* genes. Reads were searched by keyword and tBLASTn against the CitEST database using characterized gene product sequences from citrus as query, when available and other species, such as *Arabidopsis thaliana*, tomato (*Lycopersicon esculentum),* rice (*Oryza sativa)* and tobacco (*Nicotiana tabacum)* when no citrus sequences were available in the GenBank. Reads not related to the resistance proteins or those that exhibited E-values greater than 10^{-10} were excluded from the analyses. The remaining ones were clustered using the CAP3 (Huang and Madan, 1999) according to bioinformatic parameters established for all analyses of the CitEST database (Reis *et al.*, this issue).

The deduced amino acid sequences were aligned using ClustalW (http://www2.ebi.ac.uk/clustalw/), and the alignments were shaded using Boxshade (http://www.ch.embnet.org/software/ BOX_form.html). In this software, the name of each protein is indicated on the left of the alignment, and identical amino acids are shaded in black while conservative substitutions are shaded in gray.

Results and Discussion

Numerous *R* genes have been cloned from a wide range of plant species, including *Arabidopsis thaliana*, flax (*Linum usitatissimum*), tomato (*Lycopersicon esculentum*), tobacco (*Nicotiana tabacum*), sugar beet (*Beta vulgaris*), apple (*Malus domestica*), rice (*Oryza sativa*), barley (*Hordeum vulgare*), and maize (*Zea mays*). Sequencing of the complete *Arabidopsis* genome has revealed approximately 149 NBS-LRR genes (Meyers *et al.*, 2003), while about 600 NBS-LRR genes have been identified in the rice genome (Goff *et al.*, 2002). No other function has been ascribed to these rice genes, suggesting that all these functional members may be involved in plant defense (Ayliffe and Lagudah, 2004).

The CitEST database search resulted in 1,300 identified reads related to *R* genes. They formed 259 contigs and 332 singletons after clusterization, and show similarities with a total of 137 *R* genes of different classes: 101 NBS-LRR (including NBS-LRR, CC-NBS-LRR and TIR-NBS-LRR); 30 RLPs (receptor-like proteins); two RLKs (receptor-like kinases); two cytoplasmic Ser/Thr kinases; and two seven-transmembrane (7-TM) family of resistance proteins (this last one is not classified into classical groups). All genes identified are shown in Table S1. Below is a description of each class of the *R* genes studied.

NBS-LRR class

ADR1 belongs to the CC-NBS-LRR class. The over-expression of this protein produces constitutive activation of salicylic acid-dependent defense genes and conveys broad-spectrum disease resistance in *A. thaliana* (Grant *et al.*, 2003). This mutant line also exhibited enhanced drought tolerance suggesting significant overlap between biotic and abiotic stress signaling networks (Chini *et al.*, 2004). Two contigs were identified in the CitEST database, C219 and C216, which are closely related to this protein (Table S1). The contig C219 (e-116) is composed of five reads from *Poncirus trifoliata* bark libraries and three reads from *Citrus reticulata* fruit libraries, and one read from leaf libraries. The contig C216 (e = 0.0) presents 12 reads in which four are from *C. reticulata* and *C. sinensis* fruit libraries and six are from *C. reticulata* and *P. trifoliata* leaf libraries. These results show a non-specific pattern of expression of this gene in citrus plants. The alignment of these contigs with the *Arabidopsis*

ADR1 protein (Genbank Accession Number NP_195056) shows conservative motifs of the NBS region (P-loop, kinase2, RNBS-A, GLPL, RNBS-D, QHDV, TVS and PKAE) (Meyers *et al.*, 1999; Cannon *et al.*, 2002; Grant *et al.*, 2003; Chini and Loake, 2005) as shown in Figure 1. Only the ADR1 protein family contains a glutamine (Q) instead of a methionine (M) as the third residue of the MHDV motif, and is referred to as QHDV in this family (Chini and Loake, 2005). TVS and PKAE motifs also seem to belong to the ADR1 protein family (Chini and Loake, 2005), and are also present in the citrus contigs C216 and C219.

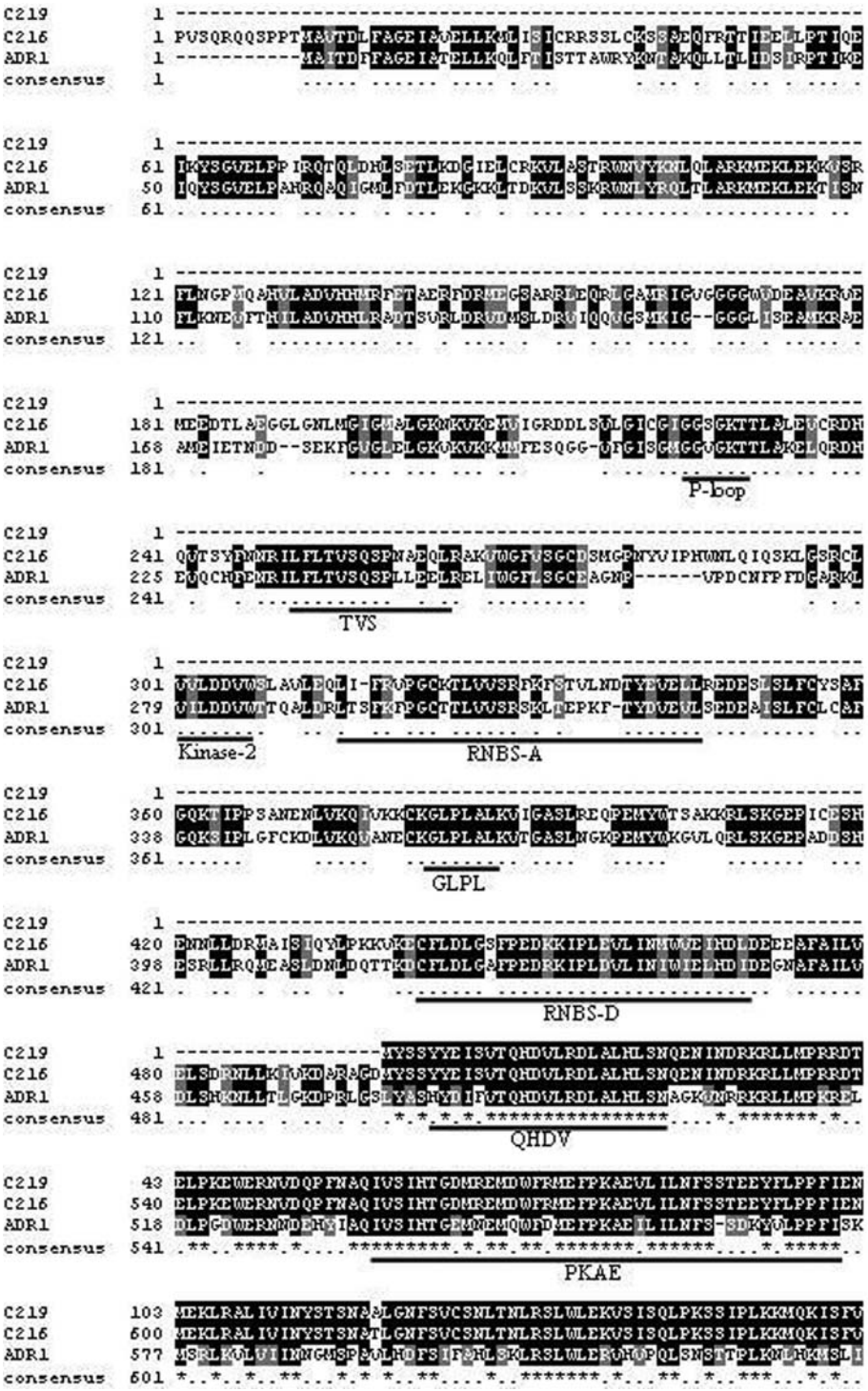

Figure 1 - Alignment (ClustalW) of the contigs C216 and C219 and *Arabidopsis* ADR1 (GenBank Accession Number NP_195056). Amino acids boxed in black and (*) are invariant, whereas residues shaded in gray and (.) are conserved in > 75% of the sequences. The motifs of the NBS domain (P-loop, kinase2, RNBS-A, GLPL, RNBS-D, QHDV, TVS and PKAE) are underlined.

The *Arabidopsis* AIG1 protein confers resistance to the *Pseudomonas syringae* pv *maculicola* strain ES4326 carrying the avrRpt2 gene and exhibiting RPS2 and avrRpt2-dependent induction early after infection (Reuber and Ausubel, 1996). Two contigs (C150 and C240) were identified as related to AIG1. The contig 150 (e-120) presents reads from *C. sinensis* (12 reads) and *C. reticulata* (five reads) libraries, while the contig 240 (e-119) contains two reads from *P. trifoliata* libraries. These results suggest species-specific expression. The singleton CR05-C1-102-032-F05-CT.F is similar (3e-36) to the B149 (CC-NBS-LRR) protein, which confers resistance to *Phytophthora infestans* in *Solanum bulbocastanum* (van der Vossen *et al.*, 2003). This singleton belongs to *C. reticulata* infected with *X. fastidiosa* library, suggesting a possible role in defense against bacterial pathogens in this plant species.

The putative citrus disease resistance proteins Pt19 and 11P31 from *Citrus grandis* x *Poncirus trifoliata* (Deng *et al.*, 2000) were also identified in the CitEST. The two singletons related to these proteins are from *C. reticulata X. fastidiosa*-infected leaf libraries, while the contig representing the 11P31 (C235) protein contains two reads, one from *C. sinensis* genome library and the other from *P. trifoliata Citrus tristeza virus*-infected leaf (Figure 2), suggesting the possible expression in infection condition. The singleton similar (4e-45) to putative disease resistance TIR-NBS R4 protein from *Malus baccata* (Table S1) is from the *P. trifoliata Citrus tristeza virus*-infected leaf library, suggesting the possible role in plant defense in this species.

Twenty-one contigs and four singletons were found in the citrus EST database similar to the *Quercus suber* resistance protein (RPc) (NBS-LRR) (Table S1). The contig 194 (1e-42) is formed by two reads: one of them is from the *C. reticulata X. fastidiosa*-infected leaf library (Figure 2). Contig 249 (2e-50) also presents two reads and one of them is from the *P. trifoliata Phytophthora*-infected bark library. Those two contigs show higher similarity to RPc genes and since they are from the libraries of pathogen infected genes, we believe that those genes are disease-resistance related.

The tomato *Bs4* resistance TIR-NBS-LRR gene specifies recognition to *Xanthomonas campestris* pv. *vesicatoria* (Ballvora *et al.*, 2001). In the CitEST database, one contig (C115) was identified as having similarity (1e-63) to this protein. This contig contains three reads from *C. sinensis* libraries: two of them are from *X. fastidiosa*-infected leaf libraries and one is from the fruit library. Since *Xylella fastidiosa* and *Xanthomonas campestris* are related bacteria (Van Sluys *et al.*, 2002), it is reasonable to infer that this gene can also provide *X. fastidiosa* resistance in citric fruits.

Sequence coding for internally conserved domains in known resistance genes have been seen in numerous plant species. The existence of conserved motifs provides opportunities for the design of degenerate primers and the isolation of disease-resistance gene analogues (RGAs) (Noir *et al.*, 2001) and disease resistance gene homologues (RGHs) (Cannon *et al.*, 2002) by PCR from plant genomes. In the CitEST database, contigs and singletons similar to RGAs and RGHs were identified (Table S1). Contig 271 shows similarity (9e-46) with RGA S-9201 (CC-NBS-LRR) from *Hordeum vulgare*. This contig presents four reads, two of them are from the library of *P. trifoliata* leaf infected by *Citrus tristeza virus*. Contig 197 presents two reads from *C. reticulata X. fastidiosa*-infected leaf library (30 days after infection) (Figure 2), and is similar (3e-19) to RGC1b (NBS-LRR) from *Lactuca sativa*. Contig 231 is also similar (2e-20) to RGC1b protein and contains only two reads: one from *P. trifoliata Citrus tristeza virus*-infected leaf library and the other is derived from the *C. sinensis X. fastidiosa*-infected leaf library. The contig 149 is highly similar (e-114) to the RGA2 (CC-NBS-LRR) protein from *Arabidopsis thaliana.* This contig contains six reads related to all species sequenced; however, the reads from *C. reticulata* and *P. trifoliata* are from the *X. fastidiosa*-infected leaf library and *Phytophthora*-infected bark library, respectively. These data suggest the disease resistance role of these reads.

The two CC-NBS-LRR class genes *RPM1* from *A. thaliana* and *RPG1*-B from *Glycine max* confer resistance to races of *Pseudomonas syringae*, the causative agent of bacterial blight, that express the avirulence gene *avr*B (Keen and Buzzel, 1991; Innes *et al.*, 1993). Searches in the CitEST identified two contigs (C95 and C210) and one singleton related to RPM1, also one contig (C10) and one sin-

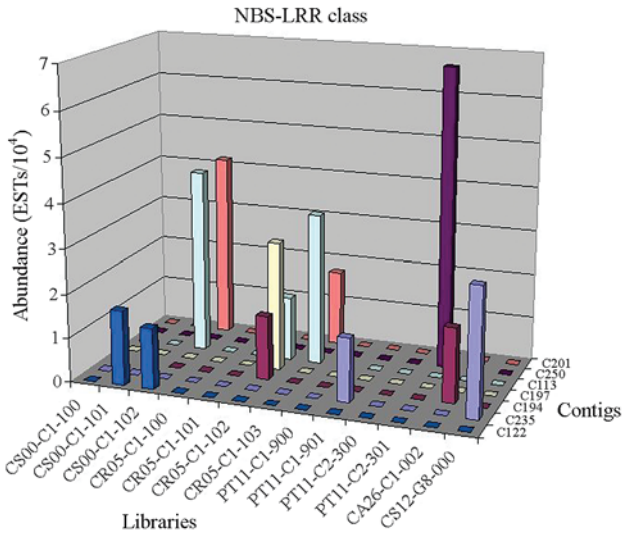

Figure 2 - Transformed data representing the relative abundance of EST by library, expressed in 10^4 reads, in contigs related to NBS-LRR class. CA: *Citrus aurantium*; CS: *Citrus sinensis*; CR: *Citrus reticulata*; PT: *Poncirus trifoliata*; C1: leaf cDNA; C2: bark cDNA; G8: shotgun genome; 100, 300 and 900: non-infected material; 101: infected with *Xylella fastidiosa*; 102: 30 days after *X. fastidiosa* infection; 103: 60 days after *X. fastidiosa* infection; 301: infected with *Phytophthora*; 901: infected with *Citrus tristeza virus*; 002: healthy material from field; 000: genome.

gleton with similarity to RPG1-B (Table S1). Contig 95 (e-102) contains most of reads from *C. sinensis* fruit libraries and contig 210 (2e-69) presents a tendency to be expressed in the fruit *C. reticulata* libraries. Contig 10 (e-108) shows a tendency to express in the *C. sinensis* and *C. reticulata* non-infected leaf libraries. The only singleton similar (4e-26) to RPG1-B is derived from *C. sinensis X. fastidiosa*-infected leaf. These sequences contain the characteristic domains and motifs related to these proteins (van der Biezen and Jones, 1998) (Figures 3 and 4).

A great number of contigs and singletons with similarities to *RPP* genes (Recognition of *Peronospora parasitica*) from *Arabidopsis thaliana* resistant to *P. parasitica* that causes downy mildew (Rehmany *et al.*,

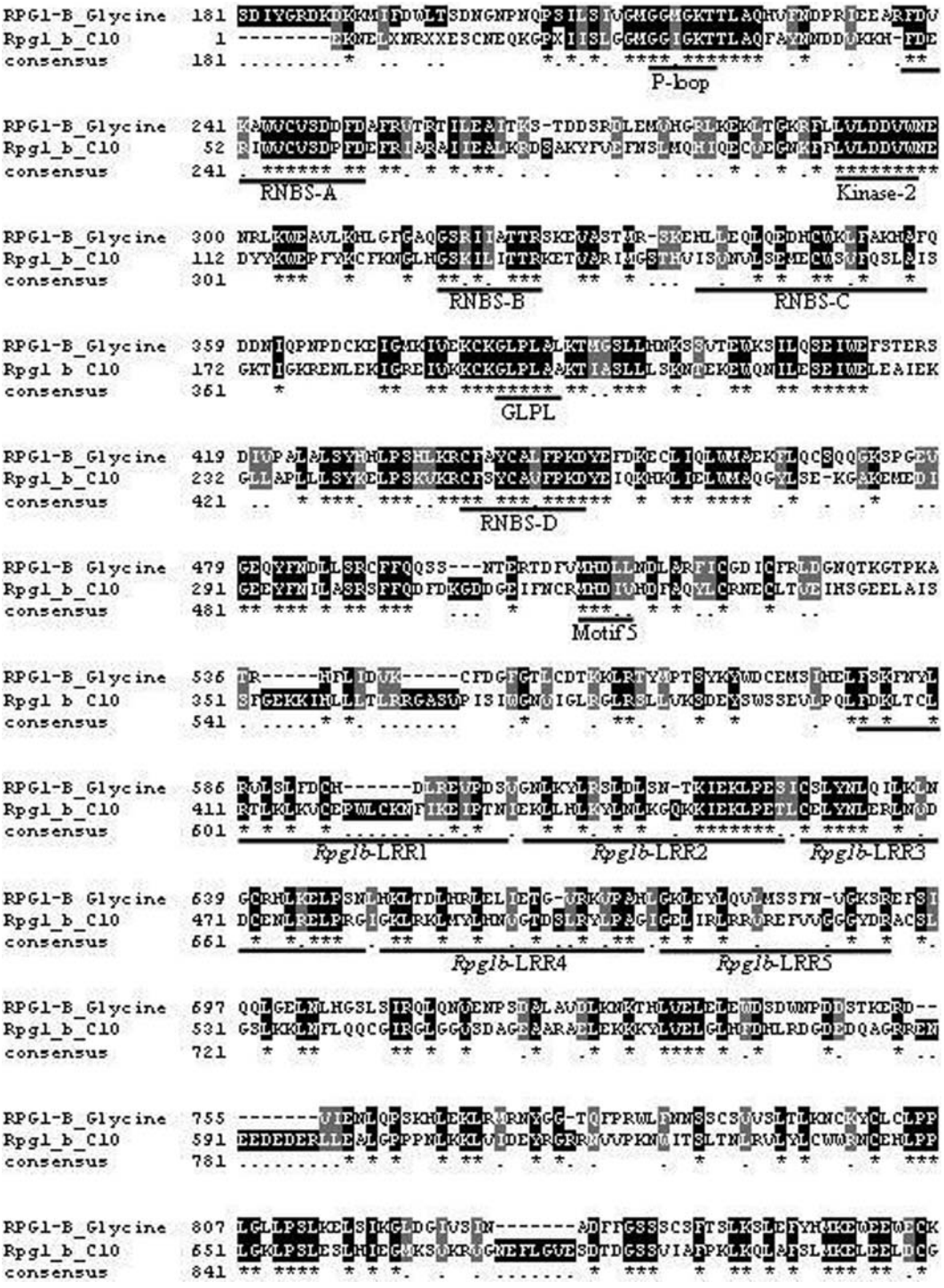

Figure 3 - Alignment of the CitEST contig C10 and RPG1-B from *Glycine max* (GenBank Accession Number AAR19097). Amino acids boxed in black and (*) are invariant, whereas residues shaded in gray and (.) are conserved in > 75% of the sequences. Domains and motifs characteristic are underlined.

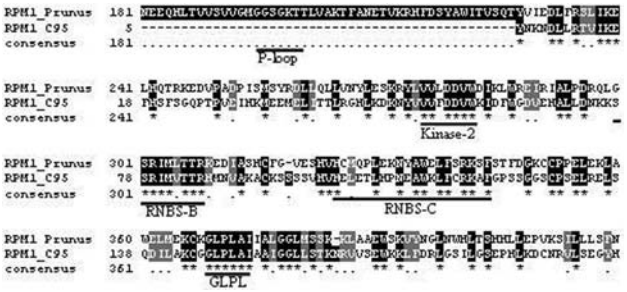

Figure 4 - Alignment (ClustalW) of the CitEST contig C95 and RPM1 from *Prunus persica* (GenBank Accession Number AAT09451). Amino acids boxed in black and (*) are invariant, whereas residues shaded in grey and (.) are conserved in > 75% of the sequences. Domains characteristic are underlined.

2005) were found in the CitEST database (Table S1). Similarities to RPP13 (CC-NBS-LRR) resulted in four contigs (C189, C225, C196 and C139). Contig 189 (1e-49) is formed by two reads from leaf libraries; one of these reads is from *C. reticulata* derived from *X. fastidiosa*-infected leaf. Contig 225 (5e-39) contains six reads; three of them are from *C. reticulata X. fastidiosa*-infected leaf libraries. Contig 196 (2e-54) shows two reads: one is from *C. sinensis* fruit library and the other is from *C. reticulata X. fastidiosa*-infected leaf library (30 days after infection). Contig 139 (1e-31) contains two reads, one is from the *C. sinensis* genome library and the other is from *C. reticulata X. fastidiosa*-infected leaf library (60 days after infection), suggesting its possible expression in *C. reticulata X. fastidiosa*-infected tissue. *The* possible expression of this protein in *C. sinensis* is not clear because, in this species, the read is from the genomic library. Contig 54, similar (9e-38) to RPP5 (TIR-NBS-LRR) consisted of three reads, two of them from *C. sinensis* fruit libraries and one from the *C. reticulata X. fastidiosa*-infected leaf library. Data indicated a possible species-specific expression pattern of this gene.

The *Arabidopsis thaliana* RPS4 (TIR-NBS-LRR) protein confers resistance to *Pseudomonas syringae* carrying avrRps4 (Zhang and Gassmann, 2003), and in the CitEST database there are sequences similarities with this protein (Table S1). Contig 117 (3e-30) is formed by three reads, one from non-infected *Citrus aurantifolia* library and two from *C. sinensis X. fastidiosa*-infected leaf libraries, suggesting a different pattern of species-dependent expression, and a possible role in *C. sinensis* defense.

Ve1 and *Ve2* genes from tomato confer race-specific resistance to the pathogenic fungi *Verticillium albo-atrum* in potato (Kawchuk *et al.*, 2001). One contig (C113), with similarity (3e-75) to the putative Ve2 from rice was identified in the CitEST database. This contig consists of six reads from *C. sinensis* and *C. reticulata Xylella*-infected leaf libraries (Figure 2), which also suggests a putative role in pathogen defense.

Reads similar to other *R* genes that confer resistance to fungus were also found in the CitEST database. For example, there are reads with similarities to *I2* locus from tomato, which confer resistance to *Fusarium oxysporum* sp *lycopersici*. In this locus, six homologous were identified, including the *I2C-1*, *I2C-2* (Simons *et al.*, 1998) and *I2C-5* genes (Sela-Buurlage *et al.*, 2001). In the CitEST, two singletons were identified for each of these proteins.

The *Citrus tristeza virus* (CTV) is one of the most important pathogens of citrus (Bernet *et al.*, 2004). A single dominant NBS-LRR class gene, *Ctv*, provides broad-spectrum resistance to CTV in *Poncirus trifoliata* L. Raf. (Gmitter *et al.*, 1996). In the CitEST database, 28 contigs and 35 singletons with similarity to *Ctv* gene (Table S1) were found in *Citrus reticulata*, *Citrus sinensis*, *Citrus aurantium*, and *Poncirus trifoliata* libraries. Contig 181 is highly similar (e = 0.0) to CTV resistance protein, and consists of two reads from the *C. reticulata* leaf library derived from *X. fastidiosa*-infected leaf (30 days after infection). Contig 65 (3e-54) consists of seven reads; three of these are from the *C. sinensis X. fastidiosa*-infected leaf library (30 days after infection). Contig 138 (e = 0.0) is formed by three reads: two from the *C. sinensis* genomic library and one from the leaf *C. sinensis X. fastidiosa*-infected library. Contig 193 (e-160) is composed of four reads: one from the fruit *C. sinensis* library and three from the *C. reticulata* library from *X. fastidiosa*-infected leaf (30 days after infection). Contig 28 (3e-47) consists of two reads: one is from the *C. sinensis* fruit library and the other is from the *C. reticulata X. fastidiosa*-infected leaf library (60 days after infection). Contig 12 (e-102) is formed by three reads. Two of them are from the *C. sinensis X. fastidiosa*-infected leaf library. Contig 250 (8e-42) presents only two reads, which are from the *P. trifoliata Phytophthora*-infected bark library (Figure 2). Finally, contig 201 (8e-49) contains four reads exclusively from *X. fastidiosa*-infected leaf libraries: three of them are from *C. sinensis* and one is from *C. reticulata* (Figure 2). These results suggested the presence *of Ctv* gene in *Citrus* species and a possible role in disease resistance against pathogens because of its presence in infected libraries. Figure 5 shows the alignment of the largest contigs (C2 and C181) and the CTV resistance protein from *P. trifoliata*. The presence of the characteristic motifs is observed in this alignment: P-loop, kinases and GLPLAL (Cannon *et al.*, 2002)**.**

Contigs and singletons in this work also were identified with similarities to *R* genes related to virus resistance, such as the *KR1*, *KR4*, *3gG2* and *Rsv3* mosaic resistance in soybean (Jeong *et al.*, 2002; He *et al.*, 2003; Hayes *et al.*, 2004; Wang *et al.*, 2004), *ry*-1 resistant to *Potato virus Y* (Vidal *et al.*, 2002) and *N* gene for mosaic virus resistance in tobacco and *Arabidopsis* (Hammond-Kosack and Jones, 1997). In the CitEST, contig 68 shows similarity (1e-44) with the ry-1 (TIR-NBS-LRR) protein. This contig comprises four reads: three of them are from the *C. sinensis* and *C. reticulata* fruit libraries and one is from the *C. sinensis* library derived from *X. fastidiosa*-infected leaf. The single-

ton with similarity (6e-35) to KR1 protein is from the *C. sinensis X. fastidiosa*-infected leaf library.

The N protein, or rather, an N protein-containing complex, is hypothesized to specifically recognize a TMV (*Tobacco mosaic virus*) protein (p50) and triggers signal transduction cascades, leading to the induction of HR, restriction of virus spread, and onset of SAR (systemic acquired resistance) (Liu *et al.*, 2004). Nineteen contigs and thirteen singletons were identified in the CitEST database with similarity to the N (TIR-NBS-LRR) protein from *Arabidopsis thaliana* and *Nicotiana tabacum* (Table S1). Among those contigs, contig 268 (5e-51) is formed by eight reads, four of them are from the *C. sinensis* leaf *X. fastidiosa*-infected library, in an early phase of infection. Contig 1 (1e-39) presents three reads: two are from the *C. sinensis X. fastidiosa*-infected leaf library (in the early stage of infection) and one is from the *C. reticulata* fruit library. In contrast, contig 182 (2e-37) consists of two reads from *C. reticulata X. fastidiosa*-infected leaf libraries and one from the *C. sinensis* fruit library, suggesting a species and tissue-dependent expression of these proteins. Contig 3 (3e-42) shows a species-dependent expression and consists of three *C. sinensis* reads, two are from the *X. fastidiosa*-infected leaf library in an early stage of infection, and one is from the fruit library. Finally, contig 122 (3e-30) presents a specific expression pattern, and contains two reads from *C. sinensis X. fastidiosa*-infected leaf libraries (Figure 2), suggesting a possible role of this protein, similar to viral protein defense, in *C. sinensis* defense against bacteria.

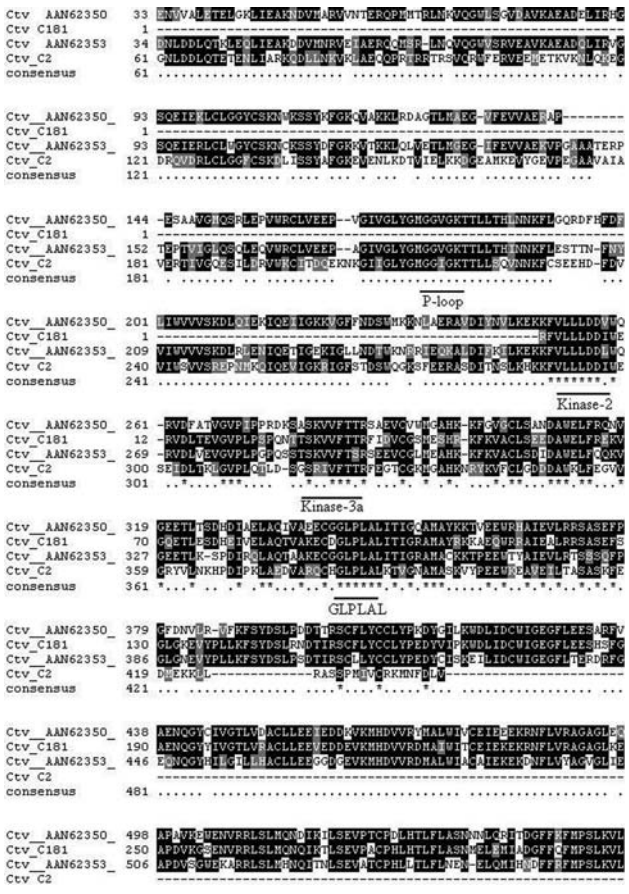

Figure 5 - Alignment (ClustalW) of the contigs C2 and C181 and *Poncirus trifoliata* Ctv proteins (GenBank Accession Numbers AAN62350 and AAN62353). Amino acids boxed in black and (*) are invariant, whereas residues shaded in gray and (.) are conserved in > 75% of the sequences. Characteristic motifs are underlined.

RLPs class

Tomato *Cf* genes confer resistance to leaf mold caused by *Cladosporium fulvum* and belong to the RLPs group (Joosten and De Wit, 1999). Contigs and singletons related to a wide variety of *Cf* genes were found in the CitEST (Table S1). Contig 52 is similar (4e-86) to Cf-2.1 from rice and consists of three reads from *C. sinensis*, two of them are from the *X. fastidiosa*-infected leaf library (30 days after infection) and one is from the fruit library, showing a possible species specific expression. The protein Cf-2.2 identified from *L. pimpinellifolium* is represented by seven contigs and one singleton. Contig 136 (e-129) presents eight reads. Five are from the *C. sinensis X. fastidiosa*-infected leaf library (30 days after infection), two are from *C. reticulata* and one is from *C. sinensis* fruit libraries, suggesting a possible role in *C. sinensis* resistance. Contig 204 (2e-54) shows expression only in *X. fastidiosa*-infected leaf libraries, and consists of two reads from *C. sinensis* and one from *C. reticulata* (Figure 6), indicating specific expression in bacteria-infection condition in both species. Contig 232 (1e-71) does not show a pattern of

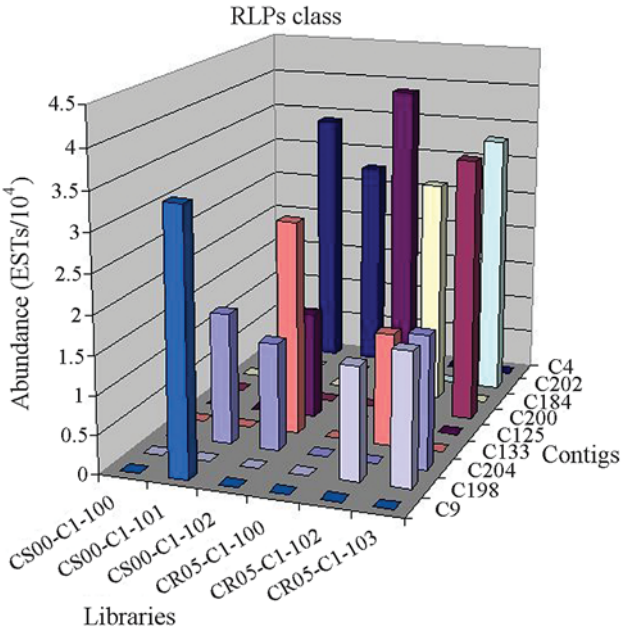

Figure 6 - Transformed data representing the relative abundance of EST by library, expressed in 10^4 reads, in contigs related to RLPs class. CS: *Citrus sinensis*; CR: *Citrus reticulata*; PT: *Poncirus trifoliata*; C1: leaf cDNA; 100: non-infected material; 101: infected with *Xylella fastidiosa*; 102: 30 days after *X. fastidiosa* infection; 103: 60 days after *X. fastidiosa* infection.

expression and consists of two reads, one is from the *C. reticulata* fruit library and the other is from *P. trifoliata Citrus tristeza virus*-infected leaf. Contigs 226 and 200 are examples of the Cf-5 protein from *L. esculentum*, which is also present in the CitEST database (Table S1). Contig 226 (1e-65) shows a tendency of expression in *P. trifoliata Citrus tristeza virus*-infected leaf and *Phytophthora*-infected bark, suggesting that this protein is important in defense against different types of pathogens (virus and fungus). Contig 200 (8e-61) shows species and tissue specific patterns of expression; this contig presents only two reads and they are from the *C. reticulata X. fastidiosa*-infected leaf library (60 days after infection) (Figure 6), indicating a possible role in *C. reticulata* bacterial defense.

The Cf-4 and Cf-9 genes are members of the Hcr9 (homologues of *C. fulvum* resistance gene Cf-9) family (Parniske *et al.*, 1997; Thomas *et al.*, 1997; Takken *et al.*, 1999), whereas the Cf-2 and Cf-5 genes belong to the Hcr2 subgroup (Dixon *et al.*, 1996, 1998). In the citrus EST database, putative proteins are related to Hcr2 and Hcr9 groups from *Lycopersicon* species (Table S1). The contig 144, similar (4e-55) to Hcr2-0B, presents reads from leaf *C. sinensis* libraries, three of them are from juvenile plants and two are from leprose virus-infected plants (with *Brevivalpus* and CiLV), suggesting a role in the early infection of *C. sinensis*. The contigs 184 (9e-76) and C202 (2e-60) are related to the Hcr2-5D protein, and these contigs present two reads from *C. reticulata X. fastidiosa*-infected leaf libraries (60 days after infection) (Figure 6), suggesting a pattern of expression that is species and tissue specific to disease resistance in this plant. With regard to the Hcr9-4C protein, one contig C61 (2e-66) is observed, which contains four reads and a tendency of expression in *C. sinensis X. fastidiosa*-infected leaf (30 days after infection), also indicating a role in disease defense.

HcrVf genes confer resistance to the apple scab pathogen *Venturia inaequalis* (Belfanti *et al.*, 2004) and they are RLPs. In the citrus database, contigs and singletons are related to a wide variety of these genes (Table S1). Contig 264 (7e-92) shows similarity with the *Malus floribunda* HcrVf2 protein, and presents a tendency of expression in *X. fastidiosa*-infected leaf libraries from *C. sinensis* (30 days after infection). Contig 239 (5e-64) is related to *Malus floribunda* HcrVf3 protein and presents three reads; two of them are made up by *P. trifoliata* from *Citrus tristeza virus*-infected leaf library and one is from non-infected leaf from *C. aurantifolia*, suggesting a role in *P. trifoliata* disease resistance.

The RPP27 protein from *A. thaliana* provides resistance to the oomycete pathogen *Peronospora parasitica* and shows extensive homology to tomato Cf proteins (Tör *et al.*, 2004). In the CitEST, 15 contigs and 14 singletons are related to this protein (Table S1). Among these contigs, four (C125, C198, C133 and C9) show specific expression in *X. fastidiosa*-infected leaf libraries. The contigs 125 (e-101) and 133 (1e-24) did not show a species specific expression pattern. Contig 125 presents four reads, three are from *C. reticulata* and one is from *C. sinensis* libraries**;** contig 133 contains two reads from *C. sinensis* and one from *C. reticulata* (Figure 6). The contigs 9 (5e-46) and 198 (4e-23) show species specific expression, since contig 9 presents two reads only from *C. sinensis* and contig 198 is formed by two reads from *C. reticulata* (Figure 6). Among the singletons, 11 are from *X. fastidiosa*-infected leaves and three are from the fruit of *C. sinensis* and *C. reticulata* libraries. These results suggest the involvement of these proteins in *C. sinensis* and *C. reticulata* disease resistance.

The *EILP* (Takemoto *et al.*, 2000), *EIX1* and *EIX2* (Ron and Avni, 2004) genes also have homology to the *Cf* resistance genes in tomato. An ethylene-inducing xylanase (EIX) is a potent elicitor of plant defense responses in specific cultivars of tobacco (*Nicotiana tabacum*) and tomato (Ron and Avni, 2004). Contig 4 is similar to EIX1 (1e-61) and despite this protein being related to fungal resistance in other species (Ron and Avni, 2004), in citrus EST database this contig is formed by four reads from leaf *X. fastidiosa*-infected libraries of *C. sinensis* (Figure 6), suggesting the possible role also in defense against bacterial pathogens.

RLKs

The genes *Xa21* and *Xa26* from rice provide resistance to a broad spectrum of bacterial blight pathogen *Xanthomonas oryzae* pv. *oryzae* strains (Li *et al.*, 2001; Sun *et al.*, 2004). The resistance activity of *Xa21* is developmentally controlled. Its resistance increases progressively from being susceptible at the juvenile stage to fully resistant at the adult stage (Sun *et al.*, 2004), while *Xa26* confers resistance to *Xanthomonas oryzae* pv. *oryzae* at both seedling and adult stages in rice (Yang *et al.*, 2003) and is constitutively expressed (Sun *et al.*, 2004). In the CitEST database, 22 contigs and 19 singletons (Table S1) similar to *Xa21* that are distributed into different libraries were found. Contig 116 (2e-80) presents three *C. sinensis* reads: two of them are from *X. fastidiosa*-infected leaf libraries and one is from the fruit library. Contig 165 (2e-75) shows tissue specific expression. While in *C. sinensis* the gene is present in the *X. fastidiosa*-infected leaf library (one read), in *C. reticulata*, there is one read in the non-infected leaf library; the opposite occurs in contig 91 (2e-37). Contigs 23 (e-111) and 129 (3e-32) show species specific patterns of expression; these contigs present only *C. sinensis* reads. In contig 23, three of them are from the *X. fastidiosa*-infected leaf library in the early stage of infection, and one is from the fruit library (Figure 7). In contig 129, two reads are from the *X. fastidiosa*-infected leaf library (30 days after infection) and one is from the fruit library, indicating the action in *C. sinensis* defense of these proteins. Contig 21 (5e-94) does not show a pattern of expression. This contig presents two reads: one from the *C. sinensis* fruit library and other from *P. trifoliata Citrus tristeza virus*-infected leaf. Contig 191

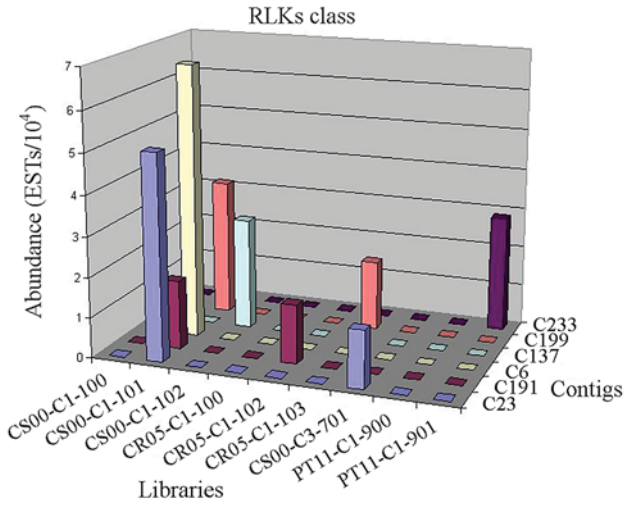

Figure 7 - Transformed data representing the relative abundance of EST by library, expressed in 10^4 reads, in contigs related to RLKs class. CS: *Citrus sinensis*; CR: *Citrus reticulata*; PT: *Poncirus trifoliata*; C1: leaf

(8e-52) shows a pattern of expression in *X. fastidiosa*-infected leaf libraries from *C. sinensis* (one read) and *C. reticulata* (one read) (Figure 7). The contigs 6, 137 and 233 show a pattern of expression that is species and tissue specific. While the contigs 6 (2e-76) and 137 (2e-50) are formed by four and two reads, respectively - they all are from *C. sinensis X. fastidiosa*-infected leaf libraries - contig 233 (6e-56) contains only reads from *P. trifoliata Citrus tristeza virus*-infected leaf library (Figure 7), suggesting a role in plant disease resistance in these species.

Five contigs and one singleton related to *Xa26* were found in the citrus database. The reads are distributed into all different citrus libraries. Contig 110 (e-153) presents 11 reads. These reads are derived from *X. fastidiosa*-infected leaf libraries of *C. sinensis* while the reads from *C. reticulata* and *P. trifoliata* have the tendency of expression in non-infected leaf libraries. Contig 120 (4e-71) comprises reads from *C. sinensis* fruit libraries (two reads) and from *C. sinensis* and *C. reticulata* leaf libraries (four reads), showing a tendency of expression in the *X. fastidiosa*-infected leaf libraries in these species. Contig 199 (4e-75) shows a pattern of expression in *X. fastidiosa*-infected leaf libraries from *C. sinensis* (two reads) and *C. reticulata* (one read) (Figure 7), indicating a possible relation to plant defense.

Cytoplasmic Ser/Thr kinase class

The *Pto* gene confers resistance to strains of *Pseudomonas syringae* pv. *tomato*, expressing *avrPto* and it was introgressed into the cultivated species *L. esculentum* from the related species *L. pimpinellifolium* (Pilowsky and Zutra, 1982; Martin *et al.*, 1993). *Pto* is a small gene. The open reading frame (ORF) consists of 963 nucleotides with no introns and encodes a functional serine-threonine kinase (Martin *et al.*, 1993; Loh and Martin, 1995). In the CitEST database, three contigs were found that were similar to this protein (Table S1). Among them, contig 237 shows similarity (7e-37) with the *Capsicum chinense* Pto protein. This contig presents two reads and they are from the *P. trifoliata Citrus tristeza virus*-infected leaf library, suggesting that in this species there is a possible role in plant defense against pathogen, as in *Capsicum chinense*.

Seven-transmembrane (7-TM) family

The protein of the disease resistance gene of barley, MLO, does not contain the recognizable structural domains of most of the known plant R gene products (Liu and Wang, 2002). This protein is the prototype of a family of seven-transmembrane (7-TM) proteins that is found in flower plants and bryophytes, but not in prokaryotes, yeast, or animals (Büschges *et al.*, 1997; Devoto *et al.*, 1999). In the dicot plant species *Arabidopsis* and the monocot barley, the presence of specific isoforms of the family of MLO proteins is required for successful host-cell invasion by ascomycete powdery mildew fungi species (Panstruga, 2005). Barley genotypes lacking functional MLO, either due to natural genetic variation (Piffanelli *et al.*, 2004) or induced lesions in the *Mlo* gene (Piffanelli *et al.*, 2002), are resistant against all known isolates of the fungal pathogen. In the CitEST database, two contigs and four singletons were found (Table S1) which are similar to the MLO protein from *A. thaliana* and *Oryza sativa*. Contig 236 (2e-59), related to the MLO protein from *A. thaliana*, presents two reads: one of them is from the *P. trifoliata Citrus tristeza virus*-infected leaf library and the other is from *C. sinensis* fruit library. Among the singletons, one of them is derived from the *C. sinensis X. fastidiosa*-infected leaf library (2e-61), and the others are from non-infected tissues.

Concluding Remarks

Even though plants are constantly threatened by potentially pathogenic microbes, surprisingly, most of them appear generally healthy. Most likely, this healthy state is due to preformed physical barriers as well as to an elaborate surveillance system of plasma membrane-anchored and possibly also cytoplasmic immune receptors. They enable plants to quickly recognize the potential pathogen by detection of conserved molecular structures, as these proteins are produced from *R* genes. This recognition mechanism is the first cellular signal that will trigger a variety of biochemical responses, which will result in the activation of defense mechanisms, lead to a HR development, and ultimately prevent the dispersion of the pathogen through the uninfected tissue. The present work represents, to our knowledge, a first attempt to identify numerous resistance proteins in citrus. In the CitEST database, hundreds of putative proteins were identified that are important to the initial trigger of the defense response in citric plants. The contigs and singletons analyzed show a tendency of pattern expression in infected-libraries similar to the *R* genes, which are con-

sistent with the function of these genes, suggesting a possible role in the plant pathogen defense. Further studies regarding the expression and specific functions of these identified genes would lead to a better understanding of the genetic and biochemical basis of pathogen resistance in citrus and other plant species. Moreover, data described herein represent a useful knowledge base for studies on the manipulation of particular resistance proteins in an attempt to enhance or to induce plant disease resistance. In order to achieve these goals, it is necessary to confirm the protein expressions and functional characterization of these genes, and also to try to overcome the diseases not only in citric species but also in other species of plants.

Acknowledgments

The authors would like to thank CNPq/ Millennium Institute (62.0054/01-8) for financial support of the CitEST project and FAPESP for financial support provided to S. G.-G. (05/58117-4).

References

Ayliffe MA and Lagudah ES (2004) Molecular genetics of disease resistance in cereals. Ann Bot 94:765-773.

Ballvora A, Pierre M, Van den Ackerveken G, Schornack S, Rossier O, Ganal M, Lahaye T and Bonas U (2001) Genetic mapping and functional analysis of the tomato *Bs4* locus, governing recognition of the *Xanthomonas campestris* pv. *vesicatoria* AvrBs4 protein. Mol Plant Microbe Interact 14:629-638.

Belfanti E, Silfverberg-Dilworth E, Tartarini S, Patocchi A, Barbieri M, Zhu J, Vinatzer BA, Gianfranceschi L, Gessler C and Sansavini S (2004) The *HcrVf2* gene from a wild apple confers scab resistance to a transgenic cultivated variety. Proc Natl Acad Sci USA 101:886-890.

Bernet GP, Breto MP and Asins MJ (2004) Expressed sequence enrichment for candidate gene analysis of citrus tristeza virus resistance. Theor Appl Genet 108:592-602.

Bittner-Eddy PD, Crute IR, Holub EB and Beynon JL (2000) *RPP13* is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica.* Plant J 21:177-188.

Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, van Daelen R, van der Lee T, Diergaarde P, Groenendijk J, *et al* (1997) The barley *Mlo* gene: A novel control element of plant pathogen resistance. Cell 88:695-705.

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.

Chini A, Grant JJ, Seki M, Shinozaki K and Loake GJ (2004) Drought tolerance established by enhanced expression of the CC-NBS-LRR gene, ADR1, requires salicylic acid, EDS1 and ABI1. Plant J 38:810-822.

Chini A and Loake GJ (2005) Motifs specific for the ADR1 NBS - LRR protein family in *Arabidopsis* are conserved among NBS - LRR sequences from both dicotyledonous and monocotyledonous plants. Planta 221:597-601.

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

Deng Z, Huang S, Ling P, Chen C, Yu C, Weber CA, Moore GA and Gmitter FG Jr (2000) Cloning and characterization of NBS-LRR class resistance-gene candidate sequences in citrus. Theor Appl Genet 101:814-822.

Devoto A, Piffanelli P, Nilsson I, Wallin E, Panstruga R, Heijne G and Schulze-Lefert P (1999) Topology, subcellular localization, and sequence diversity of the Mlo family in plants. J Biol Chem 274:34993-35004.

Dixon MS, Hatzixanthis K, Jones DA, Harrison K and Jones JDG (1998) The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. Plant Cell 10:1915-1925.

Dixon MS, Jones DA, Keddie JS, Thomas CM, Harrison K and Jones JDG (1996) The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell 84:451-459.

Dodds P, Lawrence G and Ellis J (2001) Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference between the P and P2 rust resistance specificities in flax. Plant Cell 13:163-178.

Flor HH (1971) Current status of the gene-for-gene concept. Annu Rev Phytopathol 9:275-296.

Gmitter FG, Xiao SY, Huang S, Hu XL, Garnsey SM and Deng Z (1996) A localized linkage map of the citrus tristeza virus resistance gene region. Theor Appl Genet 92:688-695.

Grant JJ, Chini A, Basu D and Loake GJ (2003) Targeted activation tagging of the *Arabidopsis* NBS-LRR gene, ADR1, conveys resistance to virulent pathogens. Mol Plant Microbe Interact 16:669-680.

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296:92-100.

Hammond-Kosack KE and Jones JDG (1997) Plant disease resistance genes. Annu Rev Plant Physiol Plant Mol Biol 48:575-607.

Hayes AJ, Jeong SC, Gore MA, Yu YG, Buss GR, Tolin SA and Maroof MA (2004) Recombination within a nucleotide-binding-site/leucine-rich-repeat gene cluster produces new variants conditioning resistance to soybean mosaic virus in soybeans. Genetics 166:493-503.

Heath MC (2000) Hypersensitive response-related death. Plant Mol Biol 44:321-334.

He CY, Tian AG, Zhang JS, Zhang ZY, Gai JY and Chen SY (2003) Isolation and characterization of a full-length resistance gene homolog from soybean. Theor Appl Genet 106:786-793.

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

Innes RW, Bisgrove SR, Smith NM, Bent AF, Staskawicz BJ and Liu YC (1993) Identification of a disease resistance locus in *Arabidopsis* that is functionally homologous to the RPG1 locus of soybean. Plant J 4:813-820.

Jeong SC, Kristipati S, Hayes AJ, Maughan PJ, Noffsinger SL, Gunduz I, Buss GR and Maroof MA (2002) Genetic and sequence analysis of markers tightly linked to the soybean mosaic virus resistance gene, *Rsv3*. Crop Sci 42:265-270.

Joosten MHAJ and De Wit PJGM (1999) The tomato-*Cladosporium fulvum* interaction: A versatile experimental system to study plant-pathogen interactions. Annu Rev Phytopathol 37:335-367.

Kawchuk LM, Hachey J, Lynch DR, Kulcsar F, van Rooijen G, Waterer DR, Robertson A, Kokko E, Byers R, Howard RJ, *et al.* (2001) Tomato *Ve* disease resistance genes encode cell surface-like receptors. Proc Natl Acad Sci USA 98:6511-6515.

Keen NT and Buzzel RI (1991) New disease resistance genes in soybean against *Pseudomonas syringae* pv. *glycinea*: Evidence that one of them interacts with a bacterial elicitor. Theor Appl Genet 81:133-138.

Liu Y, Schiff M and Dinesh-Kumar SP (2004) Involvement of MEK1 MAPKK, NTF6 MAPK, WRKY/MYB transcription factors, *COI1* and *CTR1* in N-mediated resistance to *Tobacco mosaic virus*. Plant J 38:800-809.

Liu WD and Wang SP (2002) Analysis and mapping of homologous sequences of barley disease resistance gene *Mlo* and maize disease resistance gene *Hm1* in rice. Yi Chuan Xue Bao 29:875-879.

Li ZK, Sanchez A, Angeles E, Singh S, Domingo J, Huang N and Khush GS (2001) Are the dominant and recessive plant disease resistance genes similar? A case study of rice *R* genes and *Xanthomonas oryzae* pv. *oryzae* races. Genetics 159:757-765.

Loh Y-T and Martin GB (1995) The *Pto* bacterial resistance gene and the *Fen* insecticide sensitivity gene encode functional protein kinases with serine/threonine specificity. Plant Physiol 108:1735-1739.

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

McDowell JM, Dhandaydham M, Long TA, Aarts MGM, Goff S, Holub EB and Dangl JL (1998) Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the *RPP8* locus of *Arabidopsis.* Plant Cell 10:1861-1874.

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.

Meyers BC, Kozik A, Griego A, Kuang H and Michelmore RW (2003) Genome-wide analysis of NBS-LRR - encoding genes in *Arabidopsis*. Plant Cell 15:809-834.

Nimchuk Z, Eulgem T, Holt III BF and Dangl JL (2003) Recognition and response in the plant immune system. Annu Rev Genet 37:579-609.

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.

Noutoshi Y, Ito T, Seki M, Nakashita H, Yoshida S, Marco Y, Shirasu K and Shinozaki K (2005) A single amino acid insertion in the WRKY domain of the *Arabidopsis* TIR - NBS - LRR - WRKY-type disease resistance protein SLH1 (sensitive to low humidity 1) causes activation of defense responses and hypersensitive cell death. Plant J 43:873-888.

Panstruga R (2005) Serpentine plant MLO proteins as entry portals for powdery mildew fungi. Biochem Soc Trans 33:389-392.

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.

Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BB and Jones JD (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.

Piffanelli P, Ramsay L, Waugh R, Benabdelmouna A, D'Hont A, Hollricher K, Jorgensen JH, Schulze-Lefert P and Panstruga R (2004) A barley cultivation-associated polymorphism conveys resistance to powdery mildew. Nature 430:887-891.

Piffanelli P, Zhou FS, Casais C, Orme J, Jarosch B, Schaffrath U, Collins NC, Panstruga R and Schulze-Lefert P (2002) The barley MLO modulator of defense and cell death is responsive to biotic and abiotic stress stimuli. Plant Physiol 129:1076-1085.

Pilowsky M and Zutra D (1982) Screening wild tomatoes for resistance to bacterial speck pathogen (*Pseudomonas tomato*). Plant Dis 66:46-47.

Rehmany AP, Gordon A, Rose LE, Allen RL, Armstrong MR, Whisson SC, Kamoun S, Tyler BM, Birch PR and Beynon JL (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by *RPP1* resistance genes from two *Arabidopsis* lines. Plant Cell 17:1839-1850.

Reuber TL and Ausubel FM (1996) Isolation of *Arabidopsis* genes that differentiate between resistance responses mediated by the *RPS2* and *RPM1* disease resistance genes. Plant Cell 8:241-249.

Ron M and Avni A (2004) The receptor for the fungal elicitor ethylene-inducing xylanase is a member of a resistance-like gene family in tomato. Plant Cell 16:1604-1615.

Sela-Buurlage MB, Budai-Hadrian O, Pan Q, Carmel-Goren L, Vunsch R, Zamir D and Fluhr R (2001) Genome-wide dissection of *Fusarium* resistance in tomato reveals multiple complex loci. Mol Genet Genomics 265:1104-1111.

Simons G, Groenendijk J, Wijbrandi J, Reijans M, Groenen J, Diergaarde P, Van der Lee T, Bleeker M, Onstenk J, de Both M, *et al.* (1998) Dissection of the fusarium *I2* gene cluster in tomato reveals six homologs and one active gene copy. Plant Cell 10:1055-1068.

Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai WX, Zhu LH, *et al.* (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270:1804-1805.

Sun X, Cao Y, Yang Z, Xu C, Li X, Wang S and Zhang Q (2004) *Xa26*, a gene conferring resistance to *Xanthomonas oryzae* pv. *oryzae* in rice, encodes an LRR receptor kinase-like protein. Plant J 37:517-527.

Takemoto D, Hayashi M, Doke N, Mishimura M and Kawakita K (2000) Isolation of the gene for EILP, an elicitor-inducible LRR receptor-like protein, from tobacco by differential display. Plant Cell Physiol 41:458-464.

Takken FL, Thomas CM, Joosten MH, Golstein C, Westerink N, Hille J, Nijkamp HJ, De Wit PJ and Jones JD (1999) A second gene at the tomato *Cf-4* locus confers resistance to

Cladosporium fulvum through recognition of a novel avirulence determinant. Plant J 20:279-288.

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796-815.

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.

Tör M, Brown D, Cooper A, Woods-Tör A, Sjölander K, Jones JDG and Holub EB (2004) *Arabidopsis* downy mildew resistance gene *RPP27* encodes a receptor-like protein similar to *CLAVATA2* and tomato *Cf-9*. Plant Physiol 135:1100-1112.

Trotochaud AE, Hao T, Wu G, Yang Z and Clark SE (1999) The *CLAVATA1* receptor-like kinase requires *CLAVATA3* for its assembly into a signaling complex that includes KAPP and a Rho-related protein. Plant Cell 11:393-405.

van der Biezen EA and Jones JDG (1998) The NB-ARC domain: A novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr Biol 8:R226-R227.

van der Vossen E, Sikkema A, Hekkert BL, Gros J, Stevens P, Muskens M, Wouters D, Pereira A, Stiekema W and Allefs S (2003) An ancient *R* gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato. Plant J 36:867-882.

Van Sluys MA, Monteiro-Vitorello CB, Camargo LE, Menck CF, Da Silva AC, Ferro JA, Oliveira MC, Setubal JC, Kitajima JP and Simpson AJ (2002) Comparative genomic analysis of plant-associated bacteria. Annu Rev Phytopathol 40:169-189.

Vidal S, Cabrera H, Andersson RA, Fredriksson A and Valkonen JP (2002) Potato gene *Y*-1 is an *N* gene homolog that confers cell death upon infection with potato virus Y. Mol Plant Microbe Interact 15:717-727.

Wang B, Wang Y, Wang Q, Luo G, Zhang Z, He C, He SJ, Zhang J, Gai J and Chen S (2004) Characterization of an NBS-LRR resistance gene homologue from soybean. J Plant Physiol 161:815-822.

Wang ZY, Seto H, Fujioka S, Yoshida S and Chory J (2001) BRI1 is a critical component of a plasma-membrane receptor for plant steroids. Nature 410:380-383.

Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C and Baker B (1994) The product of the *Tobacco mosaic virus* resistance gene *N*: Similarity to toll and the interleukin-1 receptor. Cell 78:1101-1115.

Wojtaszek P (1997) Oxidative burst: An early plant response to pathogen infection. Biochem J 322:681-692.

Yang Z, Sun X, Wang S and Zhang Q (2003) Genetic and physical mapping of a new gene for bacterial blight resistance in rice. Theor Appl Genet 106:1467-1472.

Zhang XC and Gassmann W (2003) RPS4-mediated disease resistance requires the combined presence of RPS4 transcripts with full-length and truncated open reading frames. Plant Cell 15:2333-2342.

Internet Resources

Boxshade, http://www.ch.embnet.org/software/ BOX_form.html (May 25, 2006).

CitEST (Citrus ESTs database), http://biotecnologia.centrodecitricultura.br (April 25, 2006).

ClustalW, http://www2.ebi.ac.uk/clustalw/ (April 29, 2006).

Supplementary Material

The following online material is available for this article:

Table S1

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

Associate Editor: Alessandra Alves de Souza

Table S1 - Number of contigs and singletons of the CitEST database, similar to know *R* genes.

Ortholog name	Genbank Accession number	Best e-value	Number	
			contigs*	singletons
NBS-LRR/NBS/LRR				
AIG1	AAM65167 *Arabidopsis thaliana*	e-120	2 (C150; C240)	3
B8	AAK61315 *Phaseolus vulgaris*	3e-29	0	1
B11	AAK61320 *Phaseolus vulgaris*	3e-18	1	1
cD7	AAD13036 *Phaseolus vulgaris*	1e-17	0	3
cD8	AAD13037 *Phaseolus vulgaris*	3e-17	1	2
clone 82-4	AAO89149 *Gossypium barbadense*	6e-30	0	1
clone 409	AAL00995 *Theobroma cacao*	2e-33	0	1
Ctv	AAN62335; AAN62348; AAN62350; AAN62351; AAN62352; AAN62353 *Poncirus trifoliata*	0.0	28 (C2; C12; C28; C65; C138; C181; C193; C201; C250)	35
I2	AAD27815 *Lycopersicon esculentum*	3e-30	0	3
I2C-1	AAB63274 *Lycopersicon esculentum*	2e-32	0	2
I2C-2	AAB63275 *Lycopersicon esculentum*	3e-17	0	2
I2C-5	AAL01986 *Lycopersicon pimpinellifolium*	1e-28	0	2
I2GA-SH194-2	AAW48302 *Solanum tuberosum*	2e-30	1	0

J71	AAK61321 *Phaseolus vulgaris*	9e-41	1	0
J78	AAK61318 *Phaseolus vulgaris*	2e-10	0	1
KR1	AF327903 *Glycine max*	6e-35	1	1
KR4	AAO15846 *Glycine max*	1e-14	1	0
MHD30	AF369833 *Vitis vinifera*	2e-51	0	2
MHD106	AAM21288 *Vitis vinifera*	2e-28	0	3
MsR1	AAN62760 *Medicago sativa*	2e-43	2	0
PM3b	BAD53385 *Oryza sativa*	2e-19	1	0
Pt3	AAN08169 *Citrus grandis* x *Poncirus trifoliata*	2e-66	1	0
Pt6	AAN08166 *Citrus grandis* x *Poncirus trifoliata*	2e-57	1	0
Pt14	AAN08168 *Citrus grandis* x *Poncirus trifoliata*	3e-28	0	1
Pt19	AAN08179 *Citrus grandis* x *Poncirus trifoliata*	7e-34	0	1
RCa2	AAO38214 *Manihot esculenta*	1e-27	0	3
RGA	CAD56833 *Lens culinaris*	7e-39	2	1
RGA-II24	AF516646 *Malus prunifolia*	2e-12	0	1
RGA s-reg19	CAD45035 *Hordeum vulgare*	2e-16	1	2
rga S-120	AJ507100 *Hordeum vulgare*	1e-17	1	0
rga S-226	CAD45027 *Hordeum vulgare*	6e-13	0	2
rga S-9201	CAD45029 *Hordeum vulgare*	9e-46	3 (C271)	5
RGA1	AAP45163 *Solanum bulbocastanum*	e-106	1	0
RGA2	CAA72178 *Arabidopsis thaliana*	e-114	1 (C149)	0

RGA2	AAP86601 *Solanum bulbocastanum*	6e-14	0	3
RGA3	AAP45165 *Solanum bulbocastanum*	1e-52	1	1
RGA4	AAP45166 *Solanum bulbocastanum*	2e-12	0	3
RGA 9	AAU89643 *Poncirus trifoliata*	5e-31	0	1
RGA 22	AY746418 *Poncirus trifoliata*	8e-19	1	0
RGC1b	AF017751 *Lactuca sativa*	4e-34	4 (C197; C231)	5
RGC2	AAQ72576 *Lactuca sativa*	6e-54	5	6
RGC2a	AF017752 *Lactuca sativa*	8e-11	0	2
RGC2B	AAD03156 *Lactuca sativa*	7e-11	0	1
RGC2J	AAD03671 *Lactuca sativa*	9e-11	0	1
RGC2K	AAD03672 *Lactuca sativa*	4e-27	0	4
RGC2K	AAP44460 *Lactuca serriola*	4e-22	2	0
RGC20	AAD03673 *Lactuca sativa*	9e-28	0	4
RGH1	AAO37645 *Manihot esculenta*	1e-31	3	4
RGH1	BAD08985 *Oryza sativa*	5e-25	0	3
RGH2	AAO37646 *Manihot esculenta*	7e-33	0	1
RPc	AY526717 *Quercus suber*	5e-81	21 (C194; C249)	4
RPH8A	BAC15497 *Oryza sativa*	2e-11	0	1
Rsv3	AAL76166 *Glycine max*	4e-58	1	2
SlVe1	AAP20229 *Solanum lycopersicoides*	2e-29	0	1
Ve1	AAK58682 *Lycopersicon esculentum*	9e-28	0	2
Ve2	XP_550023 *Oryza sativa*	3e-75	1 (C113)	0
YR1	AAN03738 *Oryza sativa*	3e-27	0	1
YR5	AAN03740 *Oryza sativa*	5e-12	0	1

YR9	AAK93796 *Oryza sativa*	3e-12	0	1
3gG2	AY518517 *Glycine Max*	2e-33	1	1
5gG3	AY518518 Glycine Max	2e-81	3	0
11P31	AAN08158 *Citrus grandis* x *Poncirus trifoliata*	8e-41	1 (C235)	1
	BAB02054 *Arabidopsis thaliana* leucine-rich repeat disease resistance protein-like	3e-43	0	1
	AAO00824; BAB10347; T46170 *Arabidopsis thaliana* (disease resistance protein)	9e-16	0	3
	AAD50010; AAG51872; AAG51873; AAF01520; AAO50553;BAC42094; CAB40943 *Arabidopsis thaliana* (putative disease resistance protein)	4e-91	2	8
	CAA06201 *Glycine Max* (resistance protein)	6e-12	0	1
	AAG21897; NP_915900 *Oryza sativa* (putative resistance protein)	5e-15	0	2
	NP_908793 *Oryza sativa* (putative stripe rust resistance protein)	6e-10	0	1
	AAF78445 *Arabidopsis thaliana*	e-122	1	9
CC-NBS-LRR				
ADR1	NP_195056 *Arabidopsis thaliana*	0.0	2 (C216; C219)	2
B149	AAR29073 *Solanum bulbocastanum*	3e-36	0	1

clone Ha-NTIR3B	AY490797 *Helianthus annuus*	6e-95	1	0
MLA6	NP_919661 *Oryza sativa*	5e-10	0	1
RPG1-B	AAR19097 *Glycine max*	e-108	1 (C10)	1
RPM1	AAD41050 *Arabidopsis lyrata*	3e-12	0	1
RPM1	AAT09451 *Prunus persica*	e-102	2 (C95; C210)	0
RPP8	AAP82810 *Arabidopsis thaliana*	0.0	5	2
RPP13	NP_190237 *Arabidopsis thaliana*	e-158	8 (C189; C225; C196; C139)	10
RPS2	AAK96709 *Arabidopsis thaliana*; AAM90858 *Arabidopsis lyrata*; AAF19803 *Brassica oleracea*	1e-13	0	3
RPS2	BAD53266 *Oryza sativa*	4e-37	2	0
StEIG-A51	AB124833 *Solanum tuberosum*	3e-44	2	1
TIR-NBS-LRR				
Bs4	AAR21295 *Lycopersicon esculentum*	1e-63	1 (C115)	0
LM6	AAG09951 *Glycine max*	3e-23	0	1
L20a	AAG48132 *Glycine max*	6e-12	0	1
MRGH63	AAO45749 *Cucumis melo*	1e-29	1	0
N	CAA16928; CAB40942; BAB08447; BAB11635 *Arabidopsis thaliana*	0.0	20 (C1; C3; C122; C182; C268)	12
N	AAT37497; BAD12594 *Nicotiana tabacum*	e-104	3	1
NBS7	AAL07542 *Helianthus annuus*	2e-22	1	0

NBS9	AAL07544 *Helianthus annuus*	7e-14	0	1
NL25	CAA08797 *Solanum tuberosum*	6e-58	0	1
PU3	AAL07535 *Helianthus annuus*	9e-30	0	1
RPP1	CAB96660 *Arabidopsis thaliana*	2e-43	1	1
RPP1-WsA	AAC72977 *Arabidopsis thaliana*	2e-12	0	1
RPP1-WsB	NP_197270 *Arabidopsis thaliana*	3e-50	0	3
RPP5	CAB40943 *Arabidopsis thaliana*	4e-72	3 (C54)	1
RPS4	BAB11393 *Arabidopsis thaliana*	2e-42	6 (C117)	5
RRS1	Q9FH83 *Arabidopsis thaliana*	7e-17	0	1
ry-1	CAC82811 *Solanum tuberosum*	2e-61	3 (C68)	2
R4	AAQ93076 *Malus baccata*	4e-45	0	1
R11	AAQ93077 *Malus* x *domestica*	9e-70	1	0
	NP_176305 *Arabidopsis thaliana* (disease resistance protein (TIR class)	5e-82	3	2
RLP class				
Cf-2.1	BAC22244 *Oryza sativa*	4e-86	3 (C52)	8
Cf-2.1	T10504 *Lycopersicon pimpinellifolium*	4e-54	0	14
Cf-2.2	AAC15780 *Lycopersicon pimpinellifolium*	e-129	7 (C136; C204; C232)	1
Cf-2	BAB64604 *Oryza sativa*	6e-30	0	3
Cf2/Cf5	CAD42634 *Hordeum vulgare*	9e-10	0	1
Cf2/Cf5	NP_917533 *Oryza sativa*	2e-11	0	1
Cf-4	CAA05268 *Lycopersicon hirsutum*	2e-26	1	0
Cf-4A	CAA73187 *Lycopersicon esculentum*	2e-37	1	0
Cf-5	AAN15323 *Arabidopsis thaliana*	e-166	1	1

Cf-5	AAC78591 *Lycopersicon esculentum*	1e-65	7 (C200; C226)	5
Cf-9	CAA05274 *Lycopersicon pimpinellifolium*	2e-54	1	3
Cf-9	AAP03881 *Nicotiana tabacum*	9e-31	0	3
Cf-9	BAD54033 *Oryza sativa*	1e-38	2	0
EILP	BAA88636 *Nicotiana tabacum*	4e-38	2	0
Eix1	AY359965 *Lycopersicon esculentum*	1e-61	1 (C4)	0
Eix2	AY359966 *Lycopersicon esculentum*	4e-72	1	0
HcrVf1	CAC40825 *Malus floribunda*	e-104	2	4
HcrVf2	CAC40826 *Malus floribunda*	7e-92	5 (C264)	4
HcrVf3	CAC40827 *Malus floribunda*	e-111	3 (C239)	3
Hcr2-0A	AAC78592 *Lycopersicon esculentum*	8e-46	1	4
Hcr2-0B	AAC78593 *Lycopersicon esculentum*	4e-55	2 (C144)	6
Hcr2-2A	AAC78594 *Lycopersicon pimpinellifolium*	6e-25	1	0
Hcr2-5B	AAC78595 *Lycopersicon esculentum*	5e-34	1	0
Hcr2-5D	AAC78596 *Lycopersicon esculentum*	9e-76	4 (C184; C202)	8
Hcr9-4C	CAA05267 *Lycopersicon hirsutum*	2e-66	1 (C61)	3
Hcr9-9D	CAA05275 *Lycopersicon pimpinellifolium*	9e-30	0	1
Hcr9-NL0D	AAD13301 *Lycopersicon esculentum*	6e-95	1	2
Hcr9-SC0A	AAD13305 Lycopersicon esculentum	1e-13	0	1
Peru 2	AAV41396 *Lycopersicon peruvianum*	8e-34	1	0
RPP27	CAE51864 *Arabidopsis thaliana*	e-101	15 (C9; C125; C133; C198)	14

RLK class				
Xa21	AAU44210; BAD69455; NP_919177; XP_464646; XP_464648; XP_464649; XP_468209; XP_481680 *Oryza sativa*	e-125	22 (C6; C21; C23; C91; C116; C129; C137; C165; C191; C233)	19
Xa26	AY364476 *Oryza sativa*	e-153	5 (C110; C120; C199)	1
Cytoplasmic Ser/Thr kinase class				
Pto	AAQ82657 *Capsicum chinense*	7e-37	1 (C237)	0
Pto	AAF76306 *Lycopersicon pimpinellifolium*	6e-66	2	0
Seven-transmembrane (7-TM) family				
Mlo	AAM45040; NP_201398; NP_565902; O49621 *Arabidopsis thaliana*	2e-59	1 (C236)	3
Mlo	AAG46114; AAN17391 *Oryza sativa*	2e-61	1	4

*Between parentheses are represented the contigs (C) commented in the Results and Discussion.