Identification and characterization of NBS–LRR class resistance gene analogs in faba bean (*Vicia faba* L.) and chickpea (*Cicer arietinum* L.)

C. Palomino, Z. Satovic, J.I. Cubero, and A.M. Torres

**Abstract:** A PCR approach with degenerate primers designed from conserved NBS–LRR (nucleotide binding site – leucine-rich repeat) regions of known disease-resistance (R) genes was used to amplify and clone homologous sequences from 5 faba bean (*Vicia faba*) lines and 2 chickpea (*Cicer arietinum*) accessions. Sixty-nine sequenced clones showed homologies to various R genes deposited in the GenBank database. The presence of internal kinase-2 and kinase-3a motifs in all the sequences isolated confirm that these clones correspond to NBS-containing genes. Using an amino-acid sequence identity of 70% as a threshold value, the clones were grouped into 10 classes of resistance-gene analogs (RGA01 to RGA10). The number of clones per class varied from 1 to 30. RGA classes 1, 6, 8, and 9 were comprised solely of clones isolated from faba bean, whereas classes 2, 3, 4, 5, and 7 included only chickpea clones. RGA10, showing a within-class identity of 99%, was the only class consisting of both faba bean and chickpea clones. A phylogenetic tree, based on the deduced amino-acid sequences of 12 representative clones from the 10 RGA classes and the NBS domains of 6 known R genes (*I2* and *Prf* from tomato, *RPP13* from *Arabidopsis*, *Gro1–4* from potato, *N* from tobacco, *L6* from flax), clearly indicated the separation between TIR (Toll/interleukin-1 receptor homology: *Gro1–4*, *L6*, *N*, RGA05 to RGA10)- and non-TIR (*I2*, *Prf*, *RPP13*, RGA01 to RGA04)-type NBS–LRR sequences. The development of suitable polymorphic markers based on cloned RGA sequences to be used in genetic mapping will facilitate the assessment of their potential linkage relationships with disease-resistance genes in faba bean and chickpea. This work is the first to report on faba bean RGAs.

**Key words:** disease resistance, NBS–LRR resistance gene analogs (RGAs), *Vicia faba* L., *Cicer arietinum* L., multiple sequence alignment, genetic mapping.

**Résumé :** Une approche PCR à l’aide d’amorces dégénérées, inspirées des régions NBS–LRR (site de liaison nucléotidique – répétition riche en leucines) de gènes de résistance (R) connus, a été employée pour amplifier et cloner des séquences homologues chez 5 variétés de féverole (*Vicia faba*) et 2 accessions de pois chiche (*Cicer arietinum*). Soixante-neuf clones séquencés montraient de l’homologie à divers gènes R présents au sein de GenBank. La présence de motifs internes kinase-2 et kinase-3a chez toutes les séquences a confirmé que ces clones correspondent à des gènes à motif NBS. En prenant pour seuil une identité de 70 % de la séquence peptidique, les clones ont été groupés en 10 classes d’analogues de gènes de résistance (RGA01 à RGA10). Le nombre de clones par classe variait entre 1 et 30. Les classes RGA 1, 6, 8 et 9 comprenaient uniquement des clones provenant de la féverole tandis que les classes 2, 3, 4, 5 et 7 n’incluaient que des clones provenant du pois chiche. La classe RGA10 affichait une identité intra-classe de 99 % et elle était la seule comprenant des clones de la féverole et du pois chiche. Un arbre phylogénétique fondé sur les séquences peptidiques prédites de 12 clones représentatifs au sein des 10 classes RGA ainsi que sur les domaines NBS de 6 gènes R connus (*I2* et *Prf* de la tomate, *RPP13* d’*Arabidopsis*, *Gro1–4* de la pomme de terre, *N* du tabac et *L6* du lin) a permis d’établir clairement une distinction entre les séquences NBS–LRR de type TIR (*Toll/Interleukin-1 Receptor : Gro1–4*, *L6*, *N*, RGA05 à RGA10) et de type non-TIR (*I2*, *Prf*, *RPP13*, RGA01 à RGA04). Le développement de marqueurs polymorphes à partir des séquences clonées de RGA permettra de les placer sur une carte génétique et facilitera l’examen de leur liaison potentielle avec des gènes de résistance chez la féverole et le pois chiche. Le présent travail est le premier à décrire des RGA chez la féverole.


Corresponding Editor: D.J. Somers.

Z. Satovic. Department of Seed Science and Technology, Faculty of Agriculture, University of Zagreb, Svetošimunska 25, 10000 Zagreb, Croatia.

1Corresponding author (e-mail: carmen.palomino.ext@juntadeandalucia.es).
Introduction

Plant diseases caused by fungi, viruses, nematodes, bacteria, and insects present severe threats to crop yields worldwide. The fight against plant pathogens is mostly based on a combination of farming practices, chemical treatments, and the use of disease-resistant cultivars. Breeding for disease resistance is one of the top priorities in plant improvement. Molecular marker technologies have located and mapped resistance genes in many species, facilitating genetic analysis and transfer into desirable lines or cultivars. In recent years, a large number of disease-resistance (R) genes that confer resistance on a diverse spectrum of pathogens have been isolated from a wide range of plant species (Boyes et al. 1998; Cooley et al. 2000; Gassmann et al. 1999; Tai et al. 1999). These genes share striking structural similarities, suggesting that certain signalling events are common to all or most plant defense systems. The most common class of plant R genes cloned so far belong to the NBS–LRR group, which contains nucleotide-binding sites (NBS) and a C-terminal leucine-rich repeat (LRR) of variable length. The NBS domain includes a number of amino-acid motifs, most notably the P-loop, kinase-2, kinase-3a, and GLPL motifs, which are highly conserved in most of the characterized R-gene products, such Arabidopsis RPS2 (Bent et al. 1994; Mindrinos et al. 1994), tobacco N (Whitham et al. 1994), and flax L6 (Lawrence et al. 1995). The R genes encode cell receptors that detect the presence of specific pathogens (Bent 1996; Hammond-Kosack and Jones 1997) and respond by activating signal transduction pathways. Nucleotide triphosphate binding is thought to alter the interaction between R proteins and other proteins acting downstream in the cascade (Bent 1996). LRR domains mediate protein–protein interactions, and are the major determinants of recognition specificity (Fluur 2001).

Plant NBS–LRR proteins can be classified into 2 major groups, on the basis of the presence or absence in the complete protein of an N-terminal region homologous to the Toll and interleukin receptor-like regions (TIR) (Meyers et al. 1999; Pan et al. 2000; Young 2000). The non-TIR proteins often contain a C-terminal coiled-coil (CC) or leucine zipper (LZ) motifs, which might play a role in the interaction with molecules downstream in the signal transduction pathway (Pan et al. 2000). So far, TIR-type genes have not been found in monocot genomes, whereas non-TIR sequences are present in both dicots and monocots (Meyers et al. 1999).

A number of studies have used PCR with degenerate oligonucleotide primers designed from these short conserved motifs of cloned plant R genes to amplify multiple DNA sequences in plant species (Aarts et al. 1998; Deng et al. 2000; Kanazin et al. 1996; Leister et al. 1996; Seah et al. 1998; Shen et al. 1998; Speulman et al. 1998; Wang and Xiao 2002; Yaish et al. 2004; Yu et al. 1996). These sequences have been called resistance-gene analogs (RGAs) (Kanazin et al. 1996) or resistance-gene candidates (RGCs) (Shen et al. 1998). Genetic analyses have associated a number of these sequences to known genes that confer resistance to viruses, bacteria, fungi, and nematodes (Aarts et al. 1998; Deng et al. 2000; Seah et al. 1998; Shen et al. 1998; Speulman et al. 1998; Wang and Xiao 2002; Yaish et al. 2004; Yu et al. 1996). Moreover, mapping studies of RGAs have provided evidence that they cosegregate with fungal disease-resistance markers (Aarts et al. 1998; Meyers et al. 1999). Consequently, this approach provides an attractive strategy to isolate multiple resistance-gene candidate sequences that can be used to develop molecular markers useful in marker-assisted selection or even lead to molecular cloning of previously unknown disease-resistance genes (Deng et al. 2000).

The faba bean (Vicia faba L.) is the third most important cool-season food legume in the world, after chickpea and pea. It is grown in many temperate areas and is used as a dry bean for food and livestock feed and as a cooked vegetable for human consumption. The chickpea (Cicer arietinum L.) is mainly cultivated in semi-arid areas of India, East Africa, the Mediterranean countries, North America, Latin America, and Australia. It is a primary human dietary protein source in developing countries where it is consumed in various forms. Recent advances in the development of genetic maps of faba bean and chickpea have allowed the identification of genes and quantitative trait loci (QTLs) for resistance to parasitic plants (Orobanche crenata) (Roman et al. 2002, 2003) and fungal diseases, such as ascochyta blight, rust, and fusarium wilt (Avila et al. 2003; 2004; Cober et al. 2005; Millán et al. 2003; Rubio et al. 2003; Santra et al. 2000; Winter et al. 2000), together with other relevant agronomic traits in both species.

The aim of this study was to isolate and characterize the NBS–LRR class of RGAs sequences from faba bean and chickpea, using degenerate oligonucleotide primers based on conserved motifs of known R genes. The final objective is to design specific primers for these sequences for use in developing markers, such as RGC-based cleaved amplified polymorphic sequences, for resistance-gene tagging and mapping. The combined analysis in both crops should facilitate the saturation of current genetic maps developed by our group and the identification of macrosynthetic relationships between disease-resistance regions in the genomes of the 2 species.

Materials and methods

Plant material and DNA extraction

Plant genomic DNA for PCR amplification was extracted from 5 V. faba lines (Vf6, Vf136, 29H, 2N52, and Vf176) and 2 accessions of C. arietinum (ILC3279 and WR315). All materials are parental lines that have been used in previous mapping studies by our group. Progenies derived from these lines have allowed the detection of genes and
QTLs for important agronomic traits, such as seed weight or resistance to broomrape, Fusarium, Uromyces, and Ascochyta (Avila et al. 2003, 2004; Millán et al. 2003; Roman et al. 2002, 2003; Rubio et al. 2003). Total DNA was extracted from young leaf tissue using the extraction method described by Torres et al. (1993). DNA samples were quantified visually with electrophoresis in 0.8% agarose gels, using the method described by Palomino et al. 2002, 2003; Rubio et al. 2003). Total DNA was extracted from each expected band and parental line were randomly selected. Clone names consisted of a letter (A to E for C. arietinum (Leister et al. 1996). I: ATA, ATC, ATT; R: G or A.

**Table 1.** Degenerate primers used to amplify resistance-gene analogs in Vicia faba and Cicer arietinum (Leister et al. 1996). I: ATA, ATC, ATT; R: G or A.

<table>
<thead>
<tr>
<th>Consensus motif nucleotide binding site</th>
<th>Primer</th>
<th>Oligonucleotide sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGV/IGKT (P-loop)</td>
<td>s1</td>
<td>GGTGGGTTGGGAAGCAACG</td>
</tr>
<tr>
<td></td>
<td>s2</td>
<td>GGIGGIGTGGGIAALCAC</td>
</tr>
<tr>
<td></td>
<td>as1</td>
<td>CAACGCTAGTGGCAATCC</td>
</tr>
<tr>
<td>GLPLAL (hydrophobic domain)</td>
<td>as2</td>
<td>IAGGCIAGGIGIAGGCC</td>
</tr>
<tr>
<td></td>
<td>as3</td>
<td>ARIGCTARI6GGIARI6CC</td>
</tr>
</tbody>
</table>

Degenerate primers and PCR conditions

Primers used in this study (Table 1) were synthesized from the sequences described by Leister et al. (1996). Primers s1 and s2 were designed in the sense direction, corresponding to the amino-acid sequence GVGKTT found in the P-loop of N, L6, and RPS2 genes. Primers as1, as3, and as4 were designed in the antisense direction, based on the sequence GLPAL (hydrophobic region in N, L6, and RPS2). PCRs for all parental lines were performed in a total volume of 25 μL, containing 50 ng of genomic DNA, 1x PCR buffer, 2 mmol/L MgCl₂, 0.1 mmol/L dNTPs, 0.25 μmol/L of each primer, and 1 U Taq polymerase (Biotools, Madrid, Spain). Six combinations of sense and antisense primers were tested: I, s1/as1; II, s2/as3; III, s2/as4; IV, s1/as3; V, s1/as4; and VI, s2/as1. Amplifications were carried out in a gradient thermocycler (T Gradient PCR, Biometra). The optimal annealing temperature was determined for each primer pair analysed.

Cycling conditions consisted of an initial denaturation step at 94 °C for 3 min, followed by 35 amplification cycles at 94 °C for 1 min, 46 °C for 30 s, and 72 °C for 30 s. The size of the amplified fragments was determined using 1% w/v SeaKem agarose, 1% w/v NuSieve agarose gels in 1x TBE buffer. The amplification products were visualized with ethidium bromide staining.

Cloning and sequencing of the PCR products

PCR products of the expected size (510–550 bp, in accordance with the R-gene sequences of RPS2, N, and L6) were excised from the gels. DNA was extracted using the QIAquick Gel Extraction kit (Qiagen, Hilden, Germany). PCR purified products were cloned into the pGEM-T vector system (Promega, Madison, Wis.) and transformed into competent Escherichia coli JM109 cells, in accordance with the supplier’s protocol. Between 20 and 40 clones derived from each expected band and parental line were randomly selected. Clone names consisted of a letter (A to E for V. faba parental lines and F or G for C. arietinum) followed by the primer combination used (I–VI) and the clone number. Positive clones were selected by PCR amplification of the inserts, using universal forward and reverse primers. Clones were then characterized, using restriction analyses with 3 enzymes (RsaI, HaeIII, and HpaII), and grouped by their restriction fragment patterns. After visual inspection on 1% agarose, 1% NuSieve gels, at least 3 clones representative of each group were selected for sequencing. Plasmids were extracted using a lysis-by-boiling miniprep protocol. Double-stranded plasmid DNA was sequenced (forward and reverse) using a BigDye terminator cycle sequencing version 3.1 kit (PE Biosystems, Foster City, Calif.) on an ABI Prism 3100 Genetic Analyzer apparatus (Applied Biosystems, Foster City, Calif.) at the Servicio de Secuenciación Automática de DNA, SCAI (University of Córdoba, Spain). Sequence editing and analysis were conducted with the BioEdit ver. 7.0.1 software program (Hall 1999).

Sequences reported here have been assigned GenBank accession Nos. DQ276887 to DQ276955.

Sequence alignment and phylogenetic analysis

Sequences with conserved NBS domains and homology to resistance genes or resistance-gene analogs were identified using the BLAST web page (http://www.ncbi.nlm.nih.gov/blast) of the National Center of Biotechnology Information (NCBI, Bethesda, Md.).

Multiple alignment of nucleotide and translated protein sequences was carried out using the ClustalW system (Higgins et al. 1994), as implemented in MEGA ver. 3.0 (Kumar et al. 2004). MEGA3 was also used to calculate p-distances between sequences at the amino-acid level and to construct the UPGMA tree. A threshold value for amino-acid identities of 70% was used to group cloned sequences into classes. Phylogenetic analysis was performed to evaluate the relationship between RGAs from V. faba and C. arietinum and plant R genes. Amino-acid sequences of 12 representative clones from 10 RGA classes, as well as the NBS domains of 6 R genes — I2 (Simons et al. 1998) and Prf (Salmeron et al. 1996) from tomato; RPP13 (Bittner-Eddy et al. 2000) from Arabidopsis; Gro1–4 (Paal et al. 2004) from potato; L6 (Lawrence et al. 1995) from flax; and N (Whitham et al. 1994) from tobacco — were aligned, and a neighboring-joining tree was generated using the PAM matrix (Dayhoff 1979). The bootstrap method was applied to evaluate the reliability of the tree branching (Felsenstein 1985). Nucleotide diversity, Pi (Nei 1987), among 12 representative clones was calculated in shifting windows of 30 nucleotides, with a step size of 2, using the DNAsp 4.0 (Rozas et al. 2003). The ratio of nonsynonymous to synonymous substitutions (Ks/Ka) was estimated for each of the RGA classes, using the method described by Nei and Gojobori (1986), with the MEGA3 program.

© 2006 NRC Canada
Results and discussion

PCR amplification and molecular cloning of RGA sequences

A degenerate oligo-primed PCR approach was chosen as a first step in the identification of putative resistance genes in *V. faba* and *C. arietinum*. The 2 most critical factors in degenerate oligo-primed PCR are the design of the primers and the PCR conditions. The primers should be designed from an amino-acid region with minimal degeneracy in codon usage (Compton 1990), and PCR conditions should be optimized to give a balance between efficiency and specificity. To isolate RGAs in *V. faba* and *C. arietinum*, we assayed 6 combinations of degenerate primers, designed by Leister et al. (1996), on genomic DNA from the parental lines described in Materials and methods (Tables 1 and 2).

It was necessary to optimize the PCR conditions (e.g., annealing temperature and MgCl₂ concentrations) and cycling for each primer combination. The optimum annealing temperature, determined after testing temperatures ranging from 40 to 55 °C, were as follows: 44 °C for primer combinations I and III; 46 °C for II and VI; and 55 °C for IV and V. The products differed considerably among primer combinations (Table 2). Thus, whereas primer combination I yielded 4 amplification products, ranging from 310 to 1016 bp, primer combinations IV and V failed to produce bands of the expected size (389 bp) and were thus excluded from further study. Finally, combinations II, III, and VI generated 2 major fragments each, from which the bands of the expected size (510–550 bp) were selected for further sequence analysis.

In initial experiments, some of the fragments obtained from primer combination I that differed in size from those expected were also cloned. Nevertheless, sequencing data revealed a complete lack of homology with R genes (data not shown); these clones were disregarded. The main products of primer combination III were also cloned and sequenced. However, the BLASTX database search showed that all of the sequence-characterized inserts contained retrotransposon-like sequences (data not shown); these were discarded from the analysis as well. Thus, all the informative amplicons considered in the analysis were obtained by cloning the major band from primer combinations II and VI (Fig. 1).

Restriction analysis of positive clones from primer combinations II and VI showed that the 510–550 bp DNA bands contained heterogeneous fragments, indicating amplification of more than 1 RGA by the degenerate primers used. Clones showing different restriction patterns of the insert and, occasionally, several clones of identical or similar patterns were chosen and further characterized by sequencing and sequence analysis.

Sequence analysis

A total of 149 putative *V. faba* and *C. arietinum* RGA clones (excluding that from primer combination III) were sequenced. Searches of the GenBank database, using the BLASTX algorithm, revealed that 14 clones contained retro-

Table 2. PCR products obtained with the different primer combinations.

<table>
<thead>
<tr>
<th>Primer combination</th>
<th>Size (bp)²</th>
<th>No.³</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (s1/as1)</td>
<td>1016/710/420/310 —</td>
<td>—</td>
</tr>
<tr>
<td>II (s2/as3)</td>
<td>540/409</td>
<td>14/73</td>
</tr>
<tr>
<td>III (s2/as4)</td>
<td>550/409</td>
<td>—</td>
</tr>
<tr>
<td>IV (s1/as3)</td>
<td>389</td>
<td>—</td>
</tr>
<tr>
<td>V (s1/as4)</td>
<td>389</td>
<td>—</td>
</tr>
<tr>
<td>VI (s2/as1)</td>
<td>584/396</td>
<td>55/76</td>
</tr>
</tbody>
</table>

²Size of major fragments from the nucleotide binding site region amplified by each primer combination.

³Number of resistance-gene analog sequences vs. total number of clones sequenced.

Fig. 1. PCR amplification products from genomic DNA of *Vicia faba* and *Cicers arietinum* parental lines, using primer combinations II (s2/as3) and VI (s2/as1). Arrows indicate the PCR products of expected size.

© 2006 NRC Canada
transposon-like sequences, 21 were highly similar to poly-proteins or zinc finger proteins of several species (e.g., *C. arietinum*, *Oryza sativa*, and *Sorghum bicolor*), and 41 clones did not have significant hits on GenBank entries. Four sequences showed similarity to cloned R genes, but 2 of them contained 1 or more stop codons and the other 2 included a sequence of limited size. These clones were excluded from the analysis. The remaining 69 amplicons could be translated into polypeptides containing single open reading frames and read. A search of the NCBI database indicated that their deduced amino-acid sequences showed a high level of identity with several plant R genes, such as *I2* and *Mi1–1* from tomato, *TMV* and *RPP13* from *Arabidopsis thaliana*, *Gro1–4* from potato, and a number of RGA sequences recently isolated using similar PCR-based approaches from genera such as *Lentis*, *Pisum*, *Cicer*, *Phaseolus*, and *Citrus* (Deng et al. 2000; Ferrier-Cana et al. 2003; Huettel et al. 2002; Timmerman-Vaughan et al. 2000; Yaish et al. 2004). Primer combination VI was the most successful in obtaining NBS sequences, with 55 of 149 (36.9%) sequences showing homologies in the database. In contrast, only 14 (9.4%) NBS sequences were derived from primer combination II.

BLAST analysis detected homology to the internal conserved motifs typical of the NBS–LRR type of R genes (kinase-2 and kinase-3a). The presence of an internal kinase-2 (LXXLDDVX) and kinase-3a (GSRI/VIITTR) motif in all the sequences isolated from *V. faba* and *C. arietinum*, independent of the primer sequences used, confirm that these clones correspond to NBS-containing genes. Meyers et al. (1999) reported that more than 1% of the *A. thaliana* genome might correspond to the NBS type of R genes, and genetic mapping studies of RGAs provided evidence that they cosegregate with fungal disease resistance markers (Aarts et al. 1998). Our results indicate that a large number of NBS-like sequences is also present in the *V. faba* and *C. arietinum* genomes. Although some RGAs in *C. arietinum* have been mapped (Huettel et al. 2002), no reports in the literature or GenBank were found on RGAs in *V. faba*. Therefore, this study is the first report on RGAs in this species. All *V. faba* and *C. arietinum* RGAs detected were closely related to sequences of known R genes and RGAs from other species. Thus, some of them might encode resistance-gene products of unknown specificity. Future research will be directed toward the identification of associations between the RGAs obtained from *V. faba* and *C. arietinum* and some diseases evaluated in our segregating populations.

Multiple alignment of the deduced amino-acid sequences of the 69 clones and a number of known R genes confirmed the presence of all the motifs characteristic of the NBS–LRR gene class (Traut 1994). No variants in the GLPL motif were obtained using the degenerate antisense primers as3 and as1 deduced from this conserved motif (Fig. 2). The only hydrophobic motif obtained from our sequences was GLPLAL. The consensus XGXGKT, corresponding to the P-loop motif, is present in almost all NBS–LRR-type plant resistance genes as well, and the corresponding DNA sequences were derived from primers s1 and s2. In our study, primer s2 always generated amplification products of the expected size in combination with the other primer. In con-
Contrast, no amplicons of the expected size were ever obtained with primer s1 (Table 2).

As with the soybean (Yu et al. 1996), the most common P-loop motif in *V. faba* and *C. arietinum* was GGVGKTT. Nevertheless, the variation WGGGKTT was also found in both species (Fig. 2). In other species, such as the common bean or lentil, the most common P-loop motifs reported were GGLGKTT (Ferrier-Cana et al. 2003) and GGMGKTT (Yaish et al. 2004), respectively.

NBS motifs of R-gene products can be categorized into 2 major groups: the TIR and non-TIR linked sequences. The former presents an aspartic-acid residue at the end of the kinase-2 motif (LLVLDDVD), whereas the latter has a characteristic tryptophan residue (LLVLDDWV) (Meyers et al. 1999). TIR and non-TIR types of RGAs have been identified in, among other plants, apple (Lee et al. 2003), alfalfa (Cordero and Skinner 2002), common bean (López et al. 2003), and chickpea (Huettel et al. 2002), using similar pairs of degenerate primers matching the P-loop and GLPL motifs. In this study, primer combination VI (s2/as1) amplified both TIR and non-TIR types of sequences, whereas primer combination II (s2/as3) only amplified TIR-type sequences. Yaish et al. (2004) suggested an association between the primer combination and the type of sequence amplified. If so, subtle changes in primer combinations could be used to amplify specific sets of RGA sequences. In legumes, primers derived from P-loop and GLPL motifs appear to amplify preferentially TIR-NBS rather than non-TIR-NBS encoding regions (Bertioli et al. 2003; Yu et al. 1996). This study produced similar results with *V. faba* and *C. arietinum* (see the Sequence alignment and phylogenetic analysis section, below).

We observed differences in the kinase-2 and kinase-3a motifs with respect to the TIR and non-TIR group. The kinase-2 motif of the TIR group was X(L/I)(I/V)(U/V)(L/I)D(D/N)V, whereas that of the non-TIR group was FL(F/V)VLDD(L/N)W. All the non-TIR sequences had the characteristic tryptophan residue (W), whereas the kinase-2 motifs found in the TIR group of sequences contained either aspartic acid (D), asparagine (N), or serine (S) residues at the last position. In both species, more variations in the TIR-group were observed. In the kinase-3a motif of the non-TIR group, a single sequence was found in all the faba bean clones (GSSVIITTR), in contrast to the variations observed in the chickpea (GS(R/K)(V/I)(I/L)(I/V)ITTR). In the TIR group, the same variations in this motif were observed in faba bean and chickpea GS(R/K)(I/V)(I/V)(T/I)(T/S)R.

In accordance with Pan et al. (2000), short conserved amino-acid residues in the amino-acid motifs NBS-II and NBS-V were also detected (Fig. 2). In NBS-II, a phenylalanine was invariably present in both TIR and non-TIR groups.
and, 5 residues downstream, a phenylalanine or a tryptophan residue was highly conserved in the TIR group and non-TIR group, respectively. Conserved motifs in the NBS domain might play an important role in the function of resistance genes, because several loss-of-function alleles in R genes are due to point mutations within conserved blocks inside the NBS domain (Pan et al. 2000). These small regions of conservation, although poorly understood, are likely to represent functionally relevant sites and useful landmarks in the isolation of R-gene homologues (Bent 1996). The alignment of the deduced amino-acid sequences of representatives of all RGA classes and the sequences of N, L6, Gro1–4, I2, Prf, and RPP13 in the homologous regions is shown in Fig. 2.

### Sequence alignment and phylogenetic analysis

A UPGMA tree, based on alignment of the 69 nucleotide and deduced amino-acid sequences, was constructed using MEGA3 (Kumar et al. 2004). DNA sequences corresponding to the P-loop (GGVGKTT) and hydrophobic motifs (GLPAL) could be contributed by primers and, thus, might not correspond to the exact genomic sequences. However, we still considered it appropriate to include them in the calculation of the percentages of identity, given the stringent PCR conditions and the degeneracy of the oligonucleotide primers used.

Ten classes (RGA1–RGA10) could be defined using a threshold level of amino-acid sequence identity of 70%, corresponding to a nucleotide sequence identity of approximately 80% (Fig. 3). The number of sequences representing each class varied from 1 for RGA03 and RGA07 to 30 for RGA01. RGAs were distributed in 2 major clades: 46 RGAs showed amino-acid sequence similarity to the non-TIR-type NBS and 23 to the TIR type. The non-TIR-type sequences were grouped into 4 classes and the TIR-type sequences into 6 classes. Classes 1, 6, 8, and 9 comprised only of clones isolated from the faba bean, whereas classes 2, 3, 4, 5, and 7 included only chickpea clones. RGA10, showing a within-class identity of 99%, was the only class that contained both faba bean and chickpea clones. The number of classes of RGAs amplified from a given plant species seems to depend both on the type of oligonucleotide primers used and on the variety/cultivar of a particular species. In legume species, such as the common bean, 8 classes of RGAs were reported (Rivkin et al. 1999), whereas 9 to 11 soybean classes have been reported by Kanazin et al. (1996) and Yu et al. (1996), respectively, using different sets of degenerate oligonucleotides and experimental conditions.

Pairwise comparisons of the deduced amino-acid sequences indicated that the percentage of identity within each of the classes ranged from 88% to 100%, whereas the identities between classes ranged from 22% to 68% (Table 3). Two of the 10 classes consisted of a single RGA sequence (RGA03 and RGA07) from C. arietinum, whereas the remaining classes contained multiple members. Classes 1 to 4, grouped as non-TIR sequences, were more similar to the I2 and Prf genes of tomato and the RPP13 gene of Arabidopsis (31% to 51% identity) (Table 3) and, as expected, showed lower levels of identity to sequences of the TIR-type R genes (Gro1–4, N, and L6). Classes RGA05 and RGA06 were sim-
ilar to the Gro1–4 gene of potato (46% average identity) and the N gene of tobacco (42.6% average identity), whereas classes 7 to 10 revealed similar percentages of identity with the 3 TIR-type resistance genes (Gro1–4, N, and L6), ranging from 29.7% to 39.6%.

The amino-acid sequences of 12 representative clones from the 10 RGA classes were chosen to generate a neighbor-joining tree, based on the multiple alignment of amino-acid sequences of these representative RGAs and the NBS domains of 6 known R genes (I2 and Prf from tomato, RPP13 from Arabidopsis, Gro1–4 from potato, L6 from flax, N from tobacco) (Figs. 2 and 4). The 2 major branches resolved in the phylogenetic tree (Fig. 4) include groups of R genes that can be differentiated on the basis of the presence or absence of the TIR region, supporting previous results (Meyers et al. 1999; Pan et al. 2000). Most of the faba bean and chickpea RGA sequences (RGA05, RGA06, RGA07, RGA08, RGA09, and RGA10) formed a major cluster with N, L6, and Gro1–4. Most of these sequences shared an aspartic-acid residue (D) at the final position of the kinase-2 motif. The only exception was RGA06, which presented an asparagine (N) residue at this position (Fig. 2). The other major cluster comprised classes RGA01, RGA02, RGA03, and RGA04, which clustered together with I2, Prf, and RPP13. In these classes, the end of the kinase-2
motif contained a tryptophan residue (W), as expected for the non-TIR group.

The analysis of the nucleotide polymorphism and the diversity of the 12 representative clones showed lower polymorphism rates at the kinase-2 and kinase-3a motifs compared with the rest of the sequence (Fig. 5). The identification of divergent sequence regions will facilitate the future design of specific primers for the development of molecular markers that might be useful in marker-assisted selection.

We calculated the ratio of nonsynonymous \((K_a)\) to synonymous \((K_s)\) nucleotide substitutions \((K_a/K_s)\) among triplets encoding the amino acids of faba bean and chickpea RGAs for each class to determine either diversifying \((K_a/K_s > 1)\) or purifying \((K_a/K_s < 1)\) selection. The \(K_a/K_s\) ratio for each group ranged from 0.0 to 0.84. These ratios are similar to those of the NBS domains from cloned R genes (Michelmore and Meyers 1998). However, the faba bean NBS domain recently obtained with the first RGAs isolated from strawberries is subject to purifying selection. Similar results have been recently obtained with the first RGAs isolated from strawberries (Martínez Zamora et al. 2004). However, the faba bean NBS domain within group RGA06 showed a ratio value higher than 1, which could indicate that diversifying selection was predominant between the sequences included in that group. Ferrier-Cana et al. (2003) did not find any evidence of diversifying selection in the NBS region of 7 common bean RGA sequences, but they did find diversifying selection in the LRR and spacer regions of these sequences.

In this paper, we report the amplification, cloning, and characterization of NBS–LRR class resistance-gene candidate sequences from V. faba and C. arietinum, using degenerate primers based on conserved motifs of known R genes. This is the first report on the isolation of RGAs in V. faba using a PCR-based approach.

Currently, we are carrying out a phylogenetic analysis, which includes these faba bean and chickpea RGA sequences and the RGAs of other legume species, as well as cloned R genes, to determine the phylogenetic relationships and to study the diversity of the RGAs obtained. Furthermore, specific primers for each RGA class, in combination with a number of restriction enzymes, are being assayed to detect polymorphic cleaved amplified polymorphic markers that will be genotyped in our faba bean and chickpea recombinant inbred line populations segregating for resistance to parasitic plants and different fungal diseases (Avila et al. 2003, 2004; Cobos et al. 2005; Millán et al. 2003; Roman et al. 2002, 2003; Rubio et al. 2003). Because RGAs tend to cluster in close proximity to disease-resistance loci (Leister et al. 1996), these PCR-based markers might provide a method for selecting resistant germplasm or progeny in faba bean and chickpea breeding programs. The combined approach might further identify macromolecular relationships between disease resistance regions in the genomes of the 2 legume species.

Acknowledgements

This study was supported by the European Community project FP6–2002-FOOD-1-506223 GRAIN LEGUMES Integrated Project. The authors are grateful to Dr. Antonio Di Pietro for critical revision of the manuscript.

References


Felsenstein, J. 1985. Confidence limits on phylogenies: An ap-


This article has been cited by:


