Crop Breeding and Applied Biotechnology 7: 179-185, 2007
Brazilian Society of Plant Breeding. Printed in Brazil

Targeting and genotyping RGAs in a mapping population of the AA genome of wild *Arachis*

Soraya Cristina de Macedo Leal-Bertioli[1*], Patrícia Messenberg Guimarães[1], and David John Bertioli[2]

Received 20 October 2005

Accepted 22 November 2006

ABSTRACT - *Peanut is an important legume crop. Although it has high levels of morphological diversity, it lacks genetic variability and sources of disease resistance. The transference of resistance genes from wild species is difficult due to the different ploidy level of the wild and cultivated species. Recently, amphidiploids have been produced that can be used as 'bridges' to introgress wild genes. Molecular markers are useful to pyramidize desirable genes and track them through generations of backcrossings. Molecular markers based on Resistance Gene Analogs have improved chances to be present in or linked to resistance gene loci. This study describes the development and genotyping of molecular markers based on resistance gene motifs. Specific primers were designed based on unique sequences of an Arachis RGA dataset. The identity of the amplified polymorphic bands was confirmed by sequencing. These markers were genotyped on a F_2 population that segregates for resistance to biotic stress types.*

Key words: Peanut, molecular markers, resistance gene analogs, marker-assisted selection, *Arachis*.

INTRODUCTION

Peanut, *Arachis hypogaea* L. is a legume of great importance, especially in tropical and subtropical areas. It is mainly used for human food, oil and animal fodder. Unlike the morphological diversity of peanut, the genetic diversity is low and the crop lacks effective sources of biotic resistance (Kochert et al. 1996). This could be due to its origin: *A. hypogaea* probably arose from a single event of allotetraploidization of a hybrid of two diploid species. The resulting plant was sexually incompatible with its wild parents and remained isolated, resulting in a genetic bottleneck (Kochert et al. 1991). Wild species, on the other hand, are genetically very diverse and a rich source of resistance to the pathogens that attack cultivated species (Halward et al. 1992, Galgaro et al. 1998).

Biotic stresses represent a major challenge for farmers. The transfer of resistance genes from wild species to crop cultivars is particularly difficult in peanut because of the different ploidy levels that cause a fertility barrier. Resistance against the root-knot nematode, *Meloidogyne arenaria,* has been transferred at least once from wild species: diploid species were intricately crossed, chromosomes duplicated and the resulting hybrid plant was backcrossed with *A. hypogaea* (Simpson 2001). This process is however very time-consuming (C. Simpson, pers. comm).

Recently, the species has been resynthesized from the cross of two wild species, *A. duranensis* and *A. ipaënsis* (Fávero et al. 2005). The chromosome number of the hybrid was duplicated and the plant was crossed

[1] Embrapa Recursos Genéticos e Biotecnologia, Parque Estação Biológica, Final W5 Norte, C.P. 02372, 70.770-900, Brasilia, DF, Brasil. *E-mail: soraya@cenargen.embrapa.br
[2] Universidade Católica de Brasília, Campus II, SGAN 916, 70.790-160, Brasília, DF, Brasil

with *A. hypogaea* cv. Runner. The resulting plant was included in pre-breeding programs in order to introgress resistances into peanut. However, when wild accessions are used as donor parents in breeding programs, genes that confer poor agronomic traits are transferred along with the trait of interest. This phenomenon is known as linkage drag and results from the transference of undesirable DNA that flanks the gene of interest. The tagging of resistance genes together with marker-assisted selection should minimize the problems with linkage drag. In a number of reports, RGAs (Resistance Gene Analogs) were found to be clustered on chromosomes near to resistance genes or to be part of them (Zhang et al. 2002). With the aim of understanding the diversity of resistance genes in *Arachis*, we scanned the genome of several wild and cultivated species in search for RGAs (Bertioli et al. 2003), using degenerate primers designed for resistance gene regions described elsewhere for other species (Kanazin et al. 1996, Leister et al. 1996, Yu et al. 1996). The largest RGA dataset for any legume was established for peanut with the 79 completely characterized RGAs. A transferable SSR-based genetic map was also recently constructed, based on a population which represents the AA genome of *Arachis.* It was generated using two wild *Arachis* accessions*, A. stenosperma* V10309 and *A. duranensis* K7988, contrasting in resistance to late leaf spot *Cercosporidium personatum* and root-knot nematode *Meloidogyne arenaria* and *M. javanica* race 4 (Moretszohn et al. 2005).

The main objective of this study was to develop a PCR-based strategy to convert these RGAs into molecular markers in order to place them on the genetic map of the AA genome of *Arachis*.

MATERIAL AND METHODS

Genomic DNA preparation

Total genomic DNA was extracted from young leaves using the protocol of Rogers and Bendich (1988) modified by an additional precipitation step of 1.2 M NaCl. DNA concentration was estimated by agarose gel electrophoresis comparing the fluorescence intensities of ethidium bromide-stained samples with those of a Low Mass DNA ladder.

Design of specific primers and RGA amplification

The translated sequences of 79 *Arachis* RGAs (Bertioli et al. 2003) were aligned using Clustalx and primers were designed to be as specific as possible. An example of a region used for primer design is shown in Figure 1, and primer details in Table 1. The Primer Melting Temperature (Tm) was calculated using the formula: Tm= 4(G+C) + 2(A+T).

PCRs were performed as follows: 15 mM $MgCl_2$, 1U Taq DNA polymerase (Invitrogen), 1X enzyme buffer, 200nM dNTPs, 25 ng genomic DNA, 1 pmol μL^{-1} of each primer and autoclaved distilled water to 15 µL. Thermocycling was programmed as follows: 5min at 96 °C, 35 cycles of: 96 °C for 1 min, primer-specific annealing temperature for 1 min (Table 1), 72 °C for 1 min, and a final extension step of 72 °C for 5 minutes. When the annealing temperatures of the forward and reverse primers diverged, the lowest temperature was used for amplifications. Reactions were run on a Mastercycler Gradient Thermocycler (Eppendorff), separated on 1% ethidium bromide-stained agarose gel and visualized under UV light.

The reactions were analyzed by polyacrylamide gel electrophoresis as described by Creste et al. (2001). To confirm that the products were RGAs, bands were excised from the gel, soaked in 100µL autoclaved distilled water overnight and maintained at 95 °C for 5 min. A 5 µL aliquot was used as template for a PCR, under the same conditions as the original amplification. PCR products were visualized by agarose gel electrophoresis, purified using a QIAQuick gel extraction kit Qiagen, and quantified by comparing fluorescence in ethidium bromide-stained agarose gel. The sequencing reactions were run on an ABI 3700 (Applied Biosystems), basically as recommended by the manufacturer.

RESULTS

As expected, the NBS region of different *Arachis* species is conserved (Figure 1). Based on the 79 *Arachis* RGAs, only six regions with insertions or deletions were therefore considered suitable to design specific primers. Of the six primers designed, S48-2F, S534-R and S5-26-F produced bands from both parents and the hybrid plant. All these three primer pairs were derived from sequences isolated from *A. stenosperma.* The primers

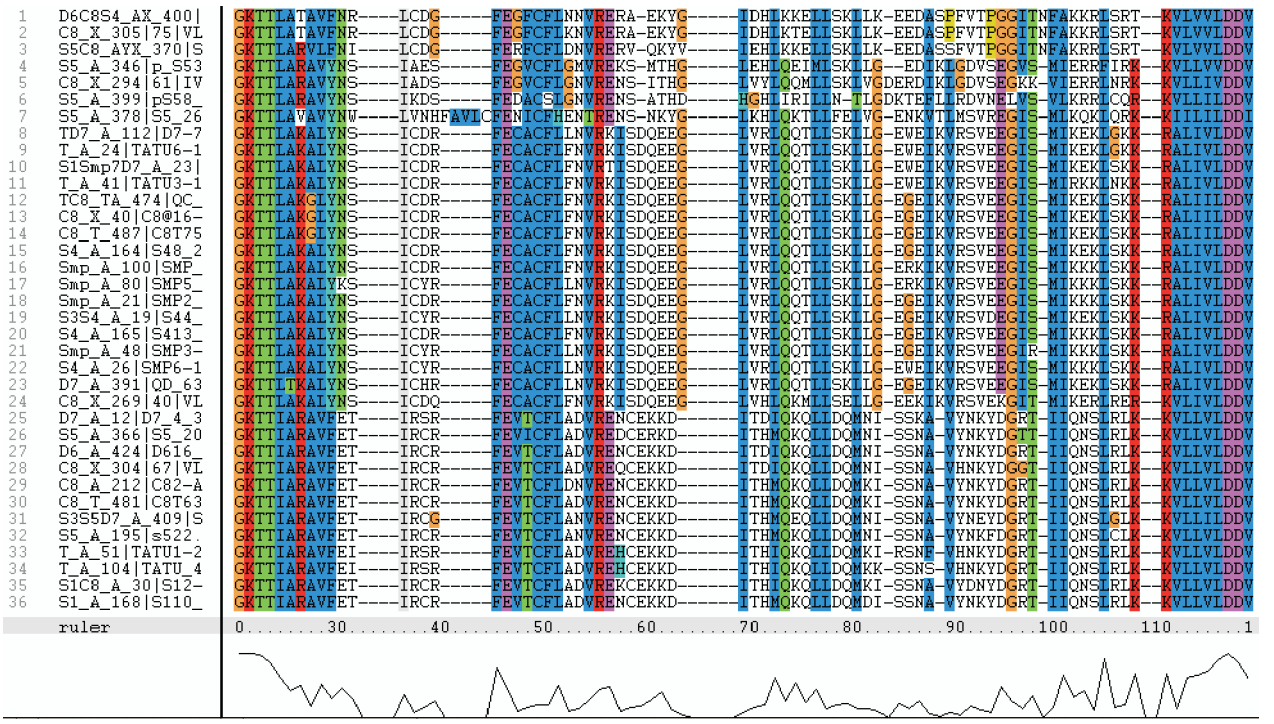

Figure 1. Example of an alignment of aminoacid sequences derived from RGAs of *Arachis* spp. using ClustalX. The region used for the design of primer S5_A_378 – fwd is underlined

Table 1. Primers designed based in unique contigs of *Arachis* spp. used to amplify RGAs by PCR. Size of PCR product was inferred based on the contig sequence from which primers were derived

Primer pair	Contig	Species	Access	Sequence 5'-3'	PCR product size	Tm
S48-2F	S4_164	*A. stenosperma*	V10309	Fwd: CAG.GAG.AAT.GTG.ATT.GGT.TTA	215 bp	58
				Rev: TT.GAG.CAT.AAT.GTA.TTG.CTT		52
C8V11R.abi	incomplete	*A. cardenasii*	GKP10017	Fwd: TCG.CTT.GTT.ACG.TAC.AAA.GCA	260 bp	60
				Rev: TTT.AGG.TGA.CAC.TAT.AGA.ATA		56
S534-R	S5_A_384	*A. stenosperma*	S5	Fwd: CCC.AAT.TTG.AAG.AAC.ACT.GCA	314 bp	60
				Rev: AAA.CAG.TGT.TAC.TCC.ATG.TGA		58
S5-26-F	S5_A_378	*A. stenosperma*	S5	Fwd: GGC.AGT.GTA.TAA.TTG.GCT.TGT	418 bp	60
				Rev: TGA.AAC.TGA.GCA.ATT.CTA.GAG		62
SMP7_3-F	SmpD6_A_56P	*A. simpsonii*	V13710	Fwd: TCC.ATA.TAT.TGT.GTA.ATG.GAT.TT	222 bp	66
				Rev: CAT.CAT.CAA.GAA.CAA.CAA.GAA.CAA		64
D6-6-3-F	incomplete	*A. duranensis*	V14167	Fwd: ATG.AGG.ATC.CTG.ATA.GAT.GGG	240 bp	61
				Rev: GGA.CAA.TTT.CCT.TGC.CGA.TCA		61

D6-6-3-F, C8V11R.abi and SMP7_3-F however produced no bands in the PCRs.

The amplification products from the parental accession and hybrid on agarose gel showed a single band. The same PCR products presented multiple bands when resolved by PAGE, most of which were monomorphic, though all primers presented a few strong polymorphic bands (Figure 2).

These polymorphic bands were isolated, reamplified and sequenced to confirm the identity. Most traces indicated the presence of a single sequence within amplified bands (eg. top and bottom traces, Figure 3), while

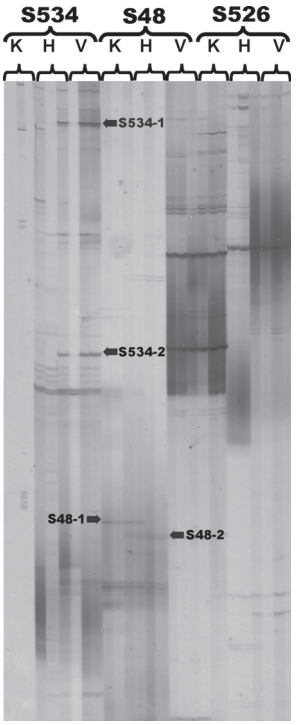

Figure 2. Polyacrylamide gel with products of S534-R and S248-2F primers with parentals and hybrid. Reactions were performed in duplicates as follows: K = *A. duranensis* K7988, H = hybrid, V = *A. stenosperma* V10307. Polymorphic bands are indicated by arrows

other traces suggested that bands are mixtures of sequences (middle trace, Figure 3). The sequence of one polymorphic band (S526-F1, not shown) did not show homology with any sequence of the *Arabidopsis* or *Arachis* databank, whereas sequence analysis of the other six polymorphic bands (S48-2-1, S48-2-2, S48-2-3, S534-1, S534-2, and S534-3) confirmed that they were highly homologous RGAs to the original clone based on which they were designed (not shown). The bands S48-1, S48-2, S534-1 and S534-2 were V10309-derived, whereas S48-3 and S534-3 were K7988-derived (Figure 2). Each set of bands amplified by the same primer pair are variants of the same RGA, with size or mobility difference and a few SNPs (sequence nucleotide polymorphisms). An example of a SNP in the sequences amplified by primer pair S534-R is shown in Figure 3. The strongest bands, S534-1 and S48-2-3, were genotyped on the complete mapping population, which consists of 93 F_2 plants derived from the hybrid of a cross of *A. stenosperma* (accession V10309) and *A. duranensis* (K7988). These species both have the AA genome. An example of genotyping gel is given in Figure 4.

DISCUSSION

This study is part of a project that aims to use resistance genes from wild relatives of peanut in breeding programs of cultivated peanut. Resistance Gene Analogs are often clustered in the genome and are frequently part of resistance genes (Zhang et al. 2002). However, since RGAs belong to a complex gene family, their use as molecular markers is complicated as it requires the use of techniques that are sensitive enough to distinguish similar sequence regions, such as PCR followed by PAGE.

The characterization of a few wild *Arachis* accessions for resistance to several pests showed that all accessions tested were more resistant than cultivated peanut (Nelson et al. 1989, Leal-Bertioli et al. 2000, Leal-Bertioli et al. 2003, Fávero 2005). This consistent high resistance level is very relevant for RGA isolation in these species. A large-scale RGA analysis was performed (Bertioli et al. 2003) but the correlation of RGAs with resistance genes could only be established by converting these into markers and placing them within a genetic map. This map would have to be based on a cross of contrasting parents for disease resistance. A microssatellite-based map for the AA genome was created (Moretszohn et al. 2005) and several attempts were made to use RGAs as molecular markers. The first strategy was Southern Blot, which had been used successfully in other crops (Collins et al. 2001, Madsen et al. 2003). Twelve representative RGAs from different clades were used as probes, with a high polymorphism degree between parents (Guimarães et al. 2005). These markers are now being placed on the *Arachis* genetic map. It is an established technique, although rather time-consuming, and requires large amounts of high quality DNA, which is not always possible for seasonal plants like *Arachis*. Alternative strategies are therefore required, which would make the screening of a larger number of RGAs and their placement on a map possible.

The placement of RGAs on genetic maps has been accomplished for several crops using Southern Blot, PCR using specific or degenerate primers and AFLP using adaptors designed for RGA motifs (Donald et al. 2002, De Giovanni et al. 2004, Xu et al. 2005, Yan et al. 2005). In this study, specific primers were designed for

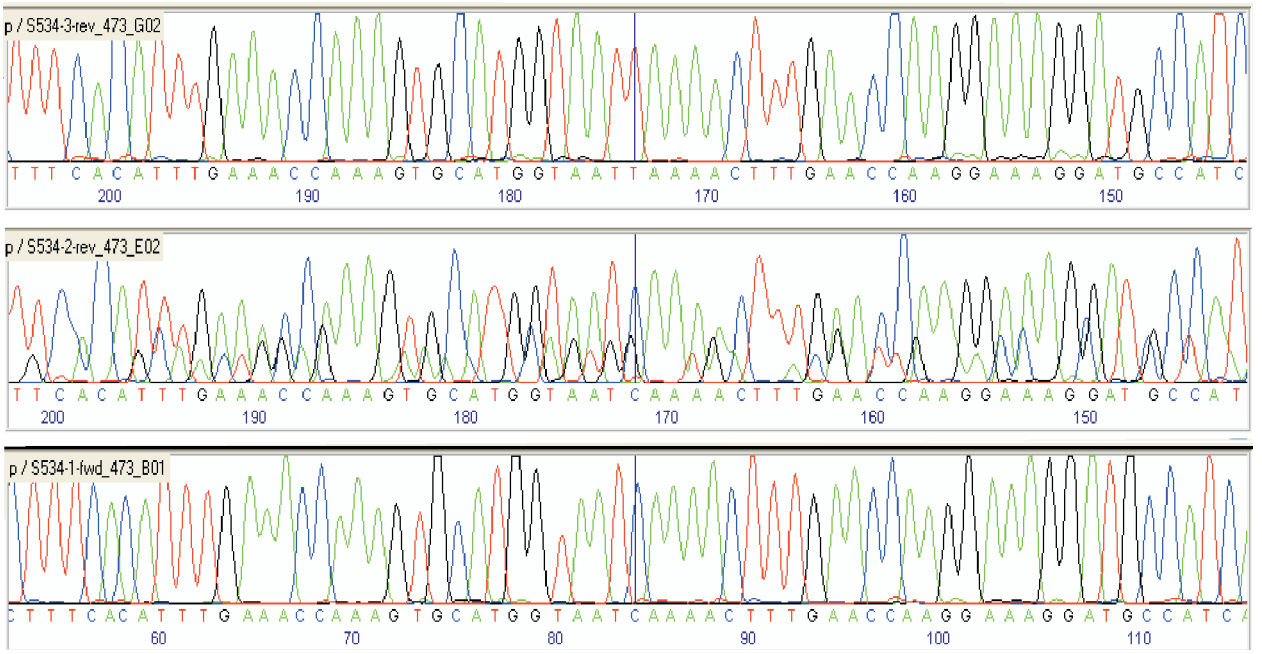

Figure 3. Example of a SNP (sequence nucleotide polymorphism) found in the sequences of polymorphic bands reamplified from polyacrylamide gels of amplified products using primer S534-R (vertical bar)

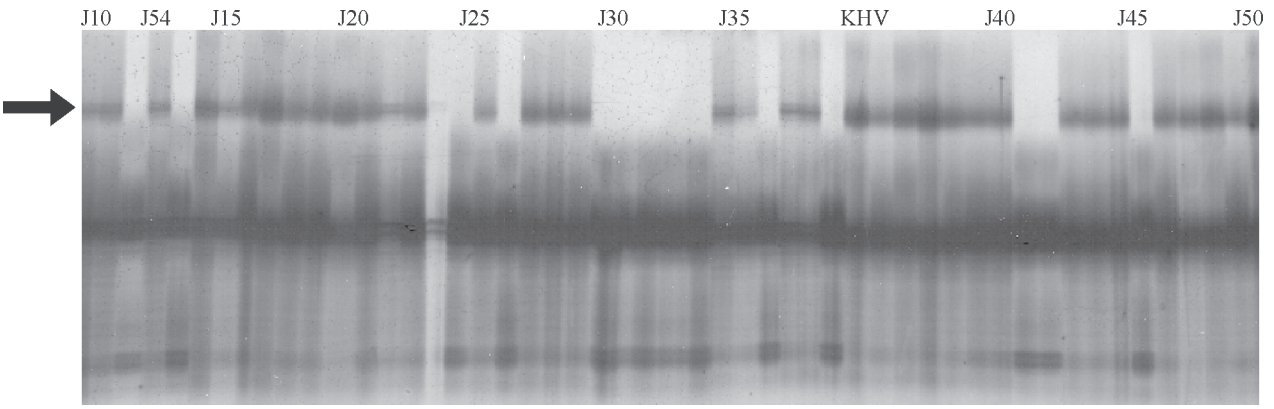

Figure 4. Example of segregation of a polymorphic band S48-2-3 (RGA) produced by PCR with the specific primers S48-2F For and Rev (indicated by black arrow). These primers were designed based on the sequence of the RGA clone S4_A_164 derived from accession V10309 of *A. stenosperma*. Patterns of the individuals J10 to J54 are shown. "K" is the parental *A. duranensis* K7988, "H" is the F_1 hybrid plant and "V" is parental *A. stenosperma* V10309

the previously obtained RGA sequences (Bertioli et al. 2003). For some regions that were unique in the RGA alignment, it was not possible to design primers because (a) the primer pairs were autocomplementary (b) Tm of forward or reverse primers were not compatible (c) the region between primer pairs was too short for PCR. Six pairs were designed in the suitable regions for primer design, so the principle of specific amplification could be tested. All primers amplified the original species from which they were sequenced, but only three amplified DNAs of the accessions that were used for the creation of the segregating population. The other primers did not show any product from these accessions, so their sequence was not homologous to any RGA from V10309 and K7988. Three primers S48-2F, S534-R and S5-26-F showed a single amplification product when resolved on agarose gel. However, when separated on polyacrylamide, several bands of different molecular

weights appeared, some of which were polymorphic. This can be explained by the fact that although the primers were designed for single sequences in the alignment, this did not rule out the possibility of an amplification of other homologous sequences in the genome, as they are part of a multigene family. Of the seven polymorphic bands produced when using these three primers, only one was not significantly homologous to the RGAs (S526-F1). The fact that the primers were designed for an RGA region did not exclude the possibility that they would bind to another, unrelated region of the genome. Therefore, the reamplification and sequencing of the bands generated is mandatory to certify the identity of the genetic marker

The markers developed here were genotyped on a F_2 population that segregates for resistance to biotic stresses. In parallel to this study, bioassays are being performed with this F_2 population in order to associate resistances to molecular markers. Introgression lines are being created by the tetraploidization of wild hybrids and their hybridization with peanut (Favero et al. 2005). These lines are back-crossed with peanut, and their progeny is phenotyped for fungal diseases. The identification of molecular markers tightly linked to genes of interest, in this case, of pathogen resistance, would allow the implementation of a Marker-Assisted Selection program (MAS), which can increase the effectiveness of the back-crosses, overcome linkage drag, and reduce the time for selection of resistant lines. In such cases, the resistance mechanism may be unknown, but the resistance-related loci can be selected through generations of selection by simple PCR (Hinchliffe et al. 2005). This would greatly speed up and optimize the development of improved peanut varieties with incorporation of new wild disease-resistance genes.

Desenvolvimento e genotipagem de marcadores RGAs em população de mapeamento genoma AA de *Arachis* silvestre

RESUMO - *O amendoim é uma leguminosa de grande importância mundial. Embora tenha alta variabilidade morfológica, tem baixa variabilidade genética e poucas fontes de resistências. A transferência de genes a partir de espécies silvestres é difícil devido à diferença de ploidia entre estas e a cultivada. Recentemente, foram produzidos anfidiplóides que podem ser utilizados como 'pontes' para a introgressão de genes silvestres. Marcadores moleculares são úteis para piramidizar estes genes e monitorá-los através das gerações de retrocruzamentos. Marcadores moleculares baseados em regiões análogas a genes de resistência (RGAs) têm maiores chances de estar ligados a locos de resistência. Este trabalho descreve o desenvolvimento e genotipagem de marcadores moleculares baseados em motivos de genes de resistência. Primers específicos foram construídos para seqüências únicas de RGAs de Arachis. A identidade das bandas polimórficas amplificadas foi confirmada por seqüenciamento. Estes marcadores foram genotipados em uma população F_2 segregante para resistência a estresses bióticos.*

Palavras-chave: Amendoim, marcadores moleculares, análogos a genes de resistência, seleção assistida por marcadores, *Arachis*.

REFERENCES

Bertioli DJ, Leal-Bertioli SC, Lion MB, Santos VL, Pappas G Jr, Cannon SB and Guimarães PM (2003) A large scale analysis of resistance gene homologues in *Arachis*. **Molecular Genetics and Genomics 270**: 34-45.

Collins N, Park R, Spielmeyer W, Ellis J and Pryor AJ (2001) Resistance gene analogs in barley and their relationship to rust resistance genes. **Genome 44**: 375-381.

Creste S, Tulmann Neto A and Figueira A (2001) Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Reports 19**: 299-306

Donald TM, Pellerone F, Adam-Blondon AF, Bouquet A, Thomas MR and Dry IB (2002) Identification of resistance gene analogs to a powdery mildew resistance locus in grapevine. **Theoretical and Applied Genetics 101**: 301-308.

De Giovanni C, Dell'Orco P, Bruno A, Ciccarese F, Lotti C and Ricciardi L (2004) Identification of PCR-based markers (RAPD, AFLP) linked to a novel powdery mildew resistance gene (*ol-2*) in tomato. **Plant Science 166**: 41-48.

Fávero AP, Simpson CE, Valls JFM and Vello NA (2005) Study of the evolution of cultivated peanut through crossability studies among *A. ipaënsis*, *A. duranensis* and *A. hypogaea*. **Crop Science 46**: 1546-1552.

Galgaro L, Lopes CR, Gimenes M, Valls JFM and Kochert G (1998) Genetic variation between species of sections *Extranervosae, Caulorrhizae, Heteranthae,* and *Triseminatae* (genus *Arachis*) estimated by DNA polymorphism. **Genome 41**: 445-454.

Guimarães PM, José ACFV, Bertioli DJ and Leal-Bertioli SCM (2005) Desenvolvimento de marcadores moleculares para Análogos a Genes de Resistência em *Arachis* silvestre. **Fitopatologia Brasileira 30**: 663-667.

Halward T, Stalker T, LaRue E and Kochert G (1992) Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). **Plant Molecular Biology 18**: 315-25.

Hinchliffe DJ, Lu Y, Potenza C, Segupta-Gopalan C, Cantrell RG and Zhang J (2005) Resistance gene analog markers are mapped to homeologous chromosomes in cultivated tetraploid cotton. **Theoretical and Applied Genetics 110**: 1074-1085.

Kanazin V, Marek LF and Shoemaker RC (1996) Resistance gene analogs are conserved and clustered in soybean. **Proceedings of the National Academy of Sciences of the USA 93**: 11746-11750.

Kochert G, Halward T, Branch WD and Simpson CE (1991) RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. **Theoretical and Applied Genetics 81**: 563-570.

Kochert G, Halward T and Stalker HT (1996) Genetic variation in peanut and its implications in plant breeding. In: Pickersgill B and Lock JM (eds.) **Advances in Legume Systematics**. Part 8: Legumes of Economic Importance. The Royal Botanic Gardens Kew, Great Britain.

Leal-Bertioli SCM, Guimarães PM, Bruzzi Lion M, Carneiro RMDG, Valls JFM and Bertioli DJ (2000) Busca de resistência ao nemtaóide das galhas *Meloidogyne* spp. e seqüências análogas a genes de resistência em acessos silvestres de *Arachis*. EMBRAPA, p. 5-17 (Boletim de Pesquisa, 20).

Leal-Bertioli SCM, Guimarães PM, Fávero AP, Moretzsohn MC, Proite K and Bertioli DJ (2003) Amendoim selvagem: uma fonte de resistência a pragas. **Biotecnologia Ciência e Desenvolvimento**. Available at www.biotecnologia.com.br. Assessed in February 2005.

Leister D, Ballvora A, Salamini F and Gebhardt C (1996) A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. **Nature Genetics 14**: 421-429

Madsen LH, Collins NC, Rakwalska M, Backes G, Sandal N, Krusell L, Jensen J, Waterman EH, Jahoor A, Ayliffe M, Pryor AJ, Langridge P, Schulze-Lefert P and Stougaard J (2003) Barley disease resistance gene analogs of the NBS-LRR class: identification and mapping. **Molecular Genetics and Genomics 269**: 150-161.

Nelson SC, Simpson CE and Starr JL (1989) Resistance to *Meloidogyne arenaria* in *Arachis* spp. germplasm. **Journal of Nematology 21**: 654-660 (Supplement).

Moretzsohn MC, Leoi L, Proite K, Guimarães PM, Leal-Bertioli SCM, Gimenes MA, Martins WS, Valls JFM, Grattapaglia D and Bertioli DJ (2005) A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). **Theoretical and Applied Genetics 111**: 1060-1071.

Rogers SO and Bendich AJ (1988) Extraction of DNA from plant tissues. In: Gelvin S and Schilperoot RA (eds.) **Plant Molecular Biology Manual**. Kluwer Academic Press, Boston, p A6: 1-10.

Yan Z, Denneboom C, Hettendorf A, Dolstra O, Debener T, Stam P and Visser PB (2005) Construction of an integrated map of rose with AFLP, SSR, PK, RGA, RFLP, SCAR and morphological markers. **Theoretical and Applied Genetics 110**: 766-777.

Xu Q, Xiaopeng W and Deng X (2005) Isolation of TIR and nonTIR NBS-LRR resistance gene analogues and identification of molecular markers linked to powdery mildew resistance locus in chestnut rose (*Rosa roxburghtii* Tratt). **Theoretical and Applied Genetics 111**: 819-830.

Simpson CE (2001) Use of wild *Arachis* species/introgression of genes into *A. hypogaea* L. **Peanut Science 28**: 114-116.

Yu YG, Buss GR and Maroof MA (1996) Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. **Proceedings of the National Academy of Sciences of the USA 93**: 11751-11756.

Zhang LP, Khan A, Niño-Liu D and Foolad MR (2002) A molecular linkage map of tomato displaying chromosomal locations of resistance gene analogs based on a *Lycopersicon esculentum* x *L. hirsutum* cross. **Genome 45**: 133-146.