

MARTA FONSECA MARTINS

CARACTERIZAÇÃO DE SONDAS HIPERVARIÁVEIS E DE UMA SEQÜÊNCIA  
GENÔMICA DE SOJA HOMÓLOGA A GENES DE RESISTÊNCIA A  
DOENÇAS

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

VIÇOSA  
MINAS GERAIS - BRASIL

MARÇO - 2000

Ficha catalográfica preparada pela Seção de Catalogação e  
Classificação da Biblioteca Central da UFV

T

M386c  
2000

Martins, Marta Fonseca, 1966-

Caracterização de sondas hipervariáveis e de uma sequência genômica de soja homóloga a genes de resistência a doenças / Marta Fonseca Martins. – Viçosa : UFV, 2000.

71p. : il.

Orientador: Everaldo Gonçalves de Barros

Tese (doutorado) – Universidade Federal de Viçosa, 1999

1. Soja - Variabilidade genética. 2. Soja - Genes de resistência - Isolamento. 3. Genes de resistência - Detecção por RT-PCR. 4. RFLP. 5. Genética molecular. I. Universidade Federal de Viçosa. II. Título.

CDD 19.ed. 633.343

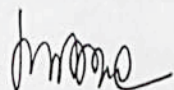
CDD 20.ed. 633.343

MARTA FONSECA MARTINS

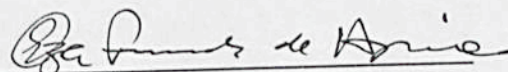
CARACTERIZAÇÃO DE SONDAS HIPERVARIÁVEIS E DE UMA SEQUÊNCIA  
GENÔMICA DE SOJA HOMÓLOGA A GENES DE RESISTÊNCIA A  
DOENÇAS

Tese apresentada à Universidade  
Federal de Viçosa, como parte das  
exigências do Programa de Pós-  
Graduação em Genética e  
Melhoramento, para obtenção do  
título de *Doctor Scientiae*.

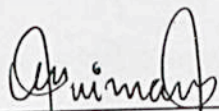
APROVADA: 25 de outubro de 1999.



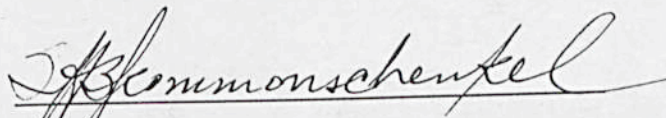
Prof. Maurilio Alves Moreira  
(Conselheiro)



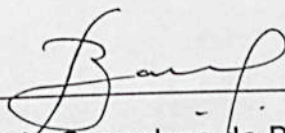
Profa. Elza Fernandes de Araújo  
(Conselheira)



Pesq. Claudia Teixeira Guimarães



Prof. Sérgio Hermínio Brommonschenkel



Prof. Everaldo Gonçalves de Barros  
(Orientador)

## AGRADECIMENTO

Ao Professor Everardo Gonçalves de Barros, pela orientação, pela amizade e pela confiança demonstradas durante o desenvolvimento deste trabalho.

À minha mãe.

Ao Professor Maurício, pela paciência e pelo apoio.

Ao Cristiano.

À Professora Elza Fe de Azevedo, pelos conselhos e pela indicação de metas a seguir.

À Bruna.

Ao grande amigo Marco Antonio, pela dedicação, pela amizade, pela convivência e pelo apoio constante.

A Nathalia, pela colaboração, pela paciência e pela amizade durante a condução dos experimentos.

Aos colegas e amigos dos Laboratórios de Genética Molecular da Planta I e II e Seq-DNA: Keul, Wagner, Francineer, Rorson, Ivan, Ana Lilia, Geraldo, Márcia, Rita, Fábio, Pedro, Cynthia, Abelmon, Valéria, Tatiana, Jay, Henrique, Vilmer, Geraldo, Maria Luiza, Marley, Luigi, Sebastião, Marcinho, Inês, Sílvia, Cláudia Fátima, Marcelinho, Newton e Marcilza, pela amizade.

... Laboratório de Biotecnologia Aplicada à Agricultura  
... Estabelecimento de Favião, Aloisio, Márcio, Alino, Graça,  
... - São Paulo, 1990.

Aos funcionários e colegas do Curso de Genética e Melhoramento,  
pela colaboração.

À Universidade Federal de Viçosa e ao BIOAGRO, pela minha  
formação acadêmica.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico  
(CNPq), pela concessão de bolsa de estudo durante o mestrado e o  
doutorado.

### **AGRADECIMENTO**

A todos que contribuíram para a conclusão de mais esta etapa da  
minha formação profissional.

Ao Professor Everaldo Gonçalves de Barros, pela orientação, pela  
amizade e pela confiança demonstradas durante o desenvolvimento deste  
trabalho.

Ao Professor Maurilio Alves Moreira, pela confiança e pelo apoio.

À Professora Elza Fernandes de Araújo, pelos conselhos e pela  
indicação de metas a seguir.

Ao grande amigo Marco Antonio, pela dedicação, pela amizade, pela  
convivência e pelo apoio constante.

À Nathalia, pela colaboração, pela persistência e pela amizade durante  
a condução dos experimentos.

Aos colegas e amigos dos Laboratórios de Genética Molecular de  
Plantas I e II e Seq-DNA: Kcal, Vagner, Francismar, Ronan, Ivan, Ana Lilia,  
Geraldo, Márcia, Rita, Fábio, Pedro, Cynthia, Abelmon, Valéria, Tatiana, Jay,  
Henrique, Vilmar, Gerardo, Maria Luiza, Marley, Luigy, Sebastião, Marcinho,  
Inês, Sílvia, Cláudia Fortes, Marcelinho, Newton e Maricília, pela amizade.

Aos funcionários do Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO), especialmente a Fausto, Aloísio, Márcio, Aline, Graça, Alessandra e Márcia, pela convivência agradável.

Aos funcionários e colegas do Curso de Genética e Melhoramento, pela colaboração.

À Universidade Federal de Viçosa e ao BIOAGRO, pela minha formação acadêmica.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão de bolsa de estudo durante o mestrado e o doutorado.

A todos que contribuíram para a conclusão de mais esta etapa da minha formação profissional.

Eu, filha de Sebastião Martins e Geralda Fonseca Martins, nasci em 12 de outubro de 1966, em Sete Lagoas, Minas Gerais.

Curso Ciências Biológicas na Universidade Federal de Minas Gerais, tendo concluído o Bacharelado em Genética no ano de 1991.

Em agosto de 1993, iniciei o Curso de Mestrado em Genética e Melhoramento na Universidade Federal de Viçosa (UFV), em Viçosa, MG, submetendo-se à defesa da tese em julho de 1995.

Em agosto de 1995, iniciei o Curso de Doutorado em Genética e Melhoramento na UFV, defendendo a tese em 25 de outubro de 1999.

## BIOGRAFIA

### CONTEUDO

Marta Fonseca Martins, filha de Sebastião Martins e Geralda Fonseca Martins, nasceu em 12 de outubro de 1966, em Sete Lagoas, Minas Gerais.

ABST. cursou Ciências Biológicas na Universidade Federal de Minas Gerais, tendo concluído o Bacharelado em Genética no ano de 1991.

2. RE Em agosto de 1993, iniciou o Curso de Mestrado em Genética e Melhoramento na Universidade Federal de Viçosa (UFV), em Viçosa, MG, submetendo-se à defesa de tese em julho de 1995.

2.4 Em agosto de 1995, iniciou o Curso de Doutorado em Genética e Melhoramento na UFV, defendendo tese em 25 de outubro de 1999.

3.1. Material Vegetal	16
3.2. Isolamento de DNA plasmidial	18
3.3. Isolamento de RNA total	20
3.4. Isolamento de DNA genômico	21
3.5. Isolamento de fragmentos de DNA de géis de agarose	21
3.6. Construção de biblioteca genômica de variedade de soja FT-Cristalina	21
3.7. "Screening" da biblioteca genômica	22
3.8. Mibração e obtenção das sondas para "screening"	23

3.12. Sequenciamento de DNA	28
3.13. RT-PCR	29
3.14. Clonagem Molecular	30
<b>4. RESULTADOS E DISCUSSÃO</b>	<b>33</b>
4.1. Caracterização das sondas de RFLP	33
4.2. Caracterização dos polimorfismos por PCR	41
4.3. Amplificação do DNA genômico com "primers" específicos para os clones de RFLP	44
4.4. Isolamento de clones genômicos homologos às sondas de RFLP	50
4.5. Caracterização do molecular do clone CR44	
	<b>Página</b>
EXTRATO	viii
ABSTRACT	x
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	3
2.1. Genoma e diversidade genética da soja	3
2.2. Resistência a doenças em plantas	5
2.3. Genes de resistência ( <i>R</i> )	10
2.4. Evolução de genes <i>R</i>	15
3. MATERIAL E MÉTODOS	18
3.1. Material Vegetal	18
3.2. Isolamento de DNA plasmidial	18
3.3. Isolamento de RNA total	20
3.4. Isolamento de DNA genômico	21
3.5. Isolamento de fragmentos de DNA de géis de agarose	21
3.6. Construção da biblioteca genômica da variedade de soja FT-Cristalina	21
3.7. "Screening" da biblioteca genômica	22
3.8. Marcação e obtenção das sondas para "screening"	23

3.9. Isolamento de DNA de fagos recombinantes.....	25
3.10. Reações de amplificação por PCR.....	26
3.11. Seqüenciamento de DNA.....	26
3.12. "Southern Blot".....	28
3.13. RT-PCR.....	29
3.14. Clonagem Molecular.....	30
4. RESULTADOS E DISCUSSÃO.....	33
4.1. Caracterização das sondas de RFLP.....	33
4.2. Caracterização dos polimorfismos por PCR.....	41
4.3. Amplificação do DNA genômico de FT-Cristalina com "primers" específicos para os clones de RFLP.....	44
4.4. Isolamento de clones genômicos homólogos às sondas de RFLP.....	50
4.5. Caracterização do molecular do clone CR44.....	50
4.6. Análise da seqüência dos subclones A53-09 E2 e A53-09 P4.....	56
4.7. Análise da expressão órgão-específica de A45-10, A53-09 e A75- 10.....	58
5. RESUMO E CONCLUSÕES.....	63
REFERÊNCIAS BIBLIOGRÁFICAS .....	66

Durante a construção de um mapa de RFLP de soja, na DuPont (EUA) as sondas analisadas foram, em sua maioria, monoméricas, porém algumas se mostraram altamente poliméricas (A1-10, A2-08, A45-10, A53-09 e A75-10). A fim de melhor caracterizar esse padrão hipervariável, essas sondas foram seqüenciadas. Duas delas (A1-10 e A2-08) não apresentaram similitude relevante com nenhuma seqüência do "GenBank", entretanto, uma característica marcante dessas sondas foi a sua riqueza em seqüências A-T. As outras três sondas continham seqüências homólogas a genes que codificam proteínas de resistência a doenças. Devido ao padrão de hibridização extremamente polimórfico revelado pela sonda A45-10, foram desenvolvidos "primers" flanqueando essa sonda para amplificar regiões homólogas em diversos genótipos de soja para que esse perfil de amplificação pudesse ser utilizado como "marcador". Os produtos de amplificação obtidos, na faixa de 1,5 a 2,0 kb, foram capazes de identificar os

... para de uma biblioteca genômica de soja. A partir de uma biblioteca genômica de soja, foram isolados quatro clones, usando a sonda A53-09. O clone CR44, contendo um inserto de aproximadamente 70 kb. Dois fragmentos de EcoRI e PstI do clone CR44, que hibridizaram com a sonda A53-09, foram subclonados e seqüenciados. Esses dois fragmentos formaram um contíguo, denominado GR, de 4.198 pb. A comparação da seqüência de aminoácidos predita com outras existentes no "GenBank" indicou uma homologia entre GR e a proteína "Cl-2.1-βke" de *Arabidopsis thaliana*, proteína da resistência a doenças homóloga à Cl-2/Cl-5 de *Hordeum vulgare* e outras proteínas de resistência de *Lycopersicon*. Além disso, foi identificada uma ORF de 783 nucleotídeos, que codifica uma proteína de 262 aminoácidos com domínios LRRs, uma das características estruturais de maioria das proteínas de resistência. Um RT-PCR foi feito para detectar

## EXTRATO

MARTINS, Marta Fonseca, D. S., Universidade Federal de Viçosa, março de 2000. **Caracterização de sondas hipervariáveis e de uma seqüência genômica de soja homóloga a genes de resistência a doenças.** Orientador: Everaldo Gonçalves de Barros. Conselheiros: Maurilio Alves Moreira e Elza Fernandes de Araújo.

Durante a construção de um mapa de RFLP de soja, na DuPont (EUA), as sondas analisadas foram, em sua maioria, monomórficas, porém algumas se mostraram altamente polimórficas (A1-10, A2-08, A45-10, A53-09 e A75-10). A fim de melhor caracterizar esse padrão hipervariável, essas sondas foram seqüenciadas. Duas delas (A1-10 e A2-08) não apresentaram similaridade relevante com nenhuma seqüência do "GenBank"; entretanto, uma característica marcante dessas sondas foi a sua riqueza em seqüências A:T. As outras três sondas continham seqüências homólogas a genes que codificam proteínas de resistência a doenças. Devido ao padrão de hibridização extremamente polimórfico revelado pela sonda A45-10, foram desenhados "primers" flanqueando essa sonda para amplificar regiões homólogas em diversos genótipos de soja, para que esse perfil de amplificação pudesse ser utilizado como "fingerprint". Os produtos de amplificação obtidos, na faixa de 1,5 a 2,0 kb, foram capazes de identificar os

diferentes genótipos analisados. A partir de uma biblioteca genômica da variedade de soja FT-Cristalina, foram isolados quatro clones, usando-se a sonda A53-09. Um deles, o clone CR44, continha um inserto de aproximadamente 16 kb. Dois fragmentos de *EcoRI* e *PstI* do clone CR44, que hibridizaram com a sonda A53-09, foram subclonados e seqüenciados. Esses dois fragmentos formaram um contíguo, denominado GR, de 4.198 pb. A comparação da seqüência de aminoácidos predita com outras existentes no "GenBank" indicou uma homologia entre GR e a proteína "Cf-2.1-like" de *Arabidopsis thaliana*, proteína de resistência a doenças homóloga à Cf-2/Cf-5 de *Hordeum vulgare* e outras proteínas de resistência de *Lycopersicon*. Além disso, foi identificada uma ORF de 789 nucleotídios, que codifica uma proteína de 262 aminoácidos com domínios LRRs, uma das características estruturais da maioria das proteínas de resistência. Um RT-PCR foi feito para detectar transcritos homólogos a A53-09, A75-10 e A45-10, usando-se "primers" que flanqueavam a região de maior similaridade com genes de resistência. Em amostras de RNA extraídas de folhas, raízes e haste foi detectada a presença de transcritos, usando-se os "primers" derivados de A53-09, o que indicou que A53-09 não é expresso de modo órgão-específico. Os produtos de amplificação presentes nos três órgãos foram seqüenciados e alinhados perfeitamente com a ORF de 789 nucleotídios. Para A75-10, foi detectada a presença de várias bandas, nos três órgãos examinados, indicando que, possivelmente, os "primers" estariam pareando com mais de um transcrito. Para A45-10, não foi detectada a presença de nenhum transcrito nos três órgãos analisados.

## ABSTRACT

MARTINS, Marta Fonseca, D. S., Universidade Federal de Viçosa, March, 2000. **Characterization of hypervariable probes and of a soybean genomic sequence homologous to disease resistance genes.** Adviser: Everaldo Gonçalves de Barros. Committee Members: Maurilio Alves Moreira and Elza Fernandes de Araújo.

During the construction of one soybean RFLP map at DuPont (USA), the majority of the probes analyzed were monomorphic, however, a few of them were highly polymorphic (A1-10, A2-08, A45-10, A53-09, and A75-10). To better understand the hypervariable pattern revealed by these probes, they were sequenced. Two of them (A1-10 and A2-08) did not present any similarities with sequences deposited in the GenBank. Their main feature was that they were highly A:T rich. The other three probes contained sequences homologous to known disease resistance genes in plants. Due to the hypervariable pattern revealed by probe A45-10, primers flanking this probe were designed and used to amplify the corresponding region in several soybean genotypes. The amplification products, in the range of 1.5 to 2 kb, allowed the identification of each of the different genotypes analyzed. Probe A53-09 was used to screen a genomic library prepared with DNA from cultivar FT-Cristalina. Four clones were isolated. One of them, clone CR44 of approximately 16 kb, was further analyzed. Restriction of this clone with *EcoRI* and *PstI* revealed two bands hybridizing with probe A53-09. These were

subcloned and sequenced. The two fragments partially overlapped and formed a contig of 4,198 bp designated GR. Blast analyses of GR with sequences deposited in the GenBank showed that it potentially encode a protein homologous to "Cf-2.1-like" protein of *Arabidopsis thaliana*, to disease resistance protein homologous to Cf-2/Cf-5 of *Hordeum vulgare* and to disease resistance proteins of *Lycopersicon*. An open-reading frame of 789 nucleotides was identified in GR. It potentially encode a sequence of 262 amino acid residues with LRR motifs, a structural characteristic of most disease resistance proteins described so far. RT-PCR was performed with primers flanking the regions homologous to disease resistance genes present in probes A53-09, A75-10, and A45-10. RNA samples extracted from leaves, roots, and stems contained transcripts which were amplified with primers derived from probe A53-09, indicating that expression of A53-09 is not organ-specific. The amplification products from the three organs were sequenced and presented perfect homology with part of the 789 bp ORF present in A53-09. The primers derived from A75-10 amplified several bands with different sizes in the three organs analyzed indicating that more than one transcript homologous to A75-10 was present in the different organs. The primers derived from A45-10 did not detect any transcripts in the organs analyzed.

...no presente trabalho, foram utilizadas as sondas A45-10, A53-09 e A75-10, que apresentavam perfil de hibridização extremamente semelhante com as sondas de resistência a doenças em plantas.

A análise de seqüência dessas sondas hipervariáveis revelou que existe similaridade de seqüência com genes que codificam proteínas de resistência a doenças em plantas. Essa caracterização levou ao segundo objetivo deste trabalho, que foi isolar regiões homólogas a essas sondas no cultivar brasileiro FT-Cristalina, resistente à mancha olho-de-rã causada por *Cercospora sojina* e ao tumor causado por *Diaporthe phaseolorum*.

## 1. INTRODUÇÃO

A clonagem de genes de resistência tem fornecido muitas oportunidades para estratégias de obtenção de plantas resistentes e seleção baseada em marcadores moleculares. Os melhoristas sempre quiseram

A soja cultivada, *Glycine max* (L.) Merrill, cultura autógama com  $2n = 2x = 40$ , apresenta base genética bastante estreita, que é explicada não só pelo seu modo de propagação, mas também pela enorme seleção artificial à qual ela tem sido submetida ao longo da sua domesticação e de seu melhoramento. Embora como espécie apresente extensa base genética, vários trabalhos têm demonstrado que a soja cultivada possui base genética muito estreita (DELANNAY et al., 1983; HIROMOTO e VELLO, 1986; ABDELNOOR et al., 1995; MACHADO, 1999).

O avanço das técnicas moleculares, nos últimos anos, tem permitido análises mais precisas do genoma da soja, bem como fazer inferências sobre a evolução da espécie e melhor escolha de progenitores para os programas de melhoramento. A análise da variabilidade do genoma de variedades de soja por hibridização com sondas RFLP (*Restriction Fragment Length Polymorphism*) revelou grande número de bandas monomórficas e algumas polimórficas de pequeno número de cópias, sendo estas últimas utilizadas normalmente na construção de mapas de ligação. Além disso, número bastante reduzido de sondas mapeia em regiões hipervariáveis do genoma, revelando grande número de bandas de DNA polimórficas entre diferentes cultivares.

A partir de uma biblioteca genômica da variedade Bônus (RAFALSKI e TINGEY, 1993), foram selecionadas algumas sondas de RFLP (A45-10, A53-09 e A75-10), que apresentavam perfil de hibridização extremamente

polimórfico, para caracterização molecular no presente trabalho. Portanto, o primeiro objetivo deste trabalho foi o de compreender a razão e a origem desse polimorfismo, revelado por sondas, em uma cultura com tão pouca variabilidade, como é o caso da soja.

A análise de seqüência dessas sondas hipervariáveis revelou que existia similaridade de seqüência com genes que codificam proteínas de resistência a doenças em plantas. Essa caracterização levou ao segundo objetivo deste trabalho, que foi isolar regiões homólogas a essas sondas no cultivar brasileiro FT-Cristalina, resistente à mancha olho-de-rã causada por *Cercospora sojina* e ao cancro-da-haste causado por *Diaporthe phaseolorum*.

A clonagem de genes de resistência tem fornecido muitas oportunidades para estratégias de obtenção de plantas resistentes e seleção baseada em marcadores moleculares. Os melhoristas sempre quiseram piramidar genes de resistência que fossem efetivos contra distintas raças de um único patógeno ou diferentes patógenos. A transferência de genes por métodos moleculares pode ser uma opção, já que apresenta algumas vantagens quando comparada com os métodos tradicionais de melhoramento. O tempo requerido para introdução é menor, independentemente do número de genes da construção a ser transferida, e permite combinar diferentes tipos de resistência.

O estudo da organização dos genes de resistência no genoma da planta e da sua dinâmica de evolução é importante na geração de novos conhecimentos e tecnologias que possam ser usados no desenvolvimento de novos cultivares.

Nos Estados Unidos, DELANNAY et al. (1982) verificaram que 84% do conjunto gênico do genoplasmato do norte do país derivava de 10 ancestrais, e 74% do conjunto gênico do genoplasmato do sul provinha de somente sete progenitores. Em outro estudo com 14 ancestrais de soja, verificou-se que 25% da base genética foi erodida dos conjuntos gênicos do norte e do sul e que, ao longo do tempo, houve uma perda de diversidade genética nos programas de melhoramento de soja, o que pode ser atribuído ao fato de que a diversidade genética de soja cultivada está sendo perdida em 50 anos no conjunto gênico de

HIROMOTO e VELLO (1986) relataram que, para a soja agrícola do ano 1983/1984, 100% do conjunto gênico tinha origem em apenas 26 ancestrais, nove desses ancestrais haviam contribuído com 80% do conjunto gênico e apenas quatro progenitores contribuíram com cerca de 50%.

Trabalhos com marcadores moleculares vêm sendo muito utilizados no estudo da variabilidade genética em plantas, por fornecerem melhor entendimento da estrutura genética e, assim, auxiliarem na

## 2. REVISÃO DE LITERATURA

escolha de progenitores para os programas de melhoramento. ABDELNOOR et al. (1995), utilizando marcadores RAPD (Random Amplified Polymorphic), agruparam 38 variedades brasileiras de soja em cinco subgrupos distintos (A),

### 2.1. Genoma e diversidade genética da soja

que contém pelo menos 50% do conjunto gênico do grupo A. Já no subgrupo B1, que era composto de 14 variedades, os ancestrais Renshou, Tokyo e PI 54610 contribuíram com 37%

do conjunto gênico. A soja (*Glycine max* L. Merrill) é uma espécie autógama, com  $2n = 40$ . O seu genoma haplóide é constituído de  $1,81 \times 10^9$  pares de nucleotídeos, sendo estes distribuídos em seqüências não-repetitivas (40%) e seqüências repetitivas (60%) (GOLDBERG, 1978).

Apesar da grande variabilidade genética do germoplasma mundial da espécie, a soja cultivada possui base genética bastante estreita. Isso se deve, principalmente, ao pequeno número de progenitores utilizados nos programas de melhoramento. Vários autores têm relatado esse fato (DELANNAY et al., 1983; HIROMOTO e VELLO, 1986; KEIM et al., 1989; ABDELNOOR et al., 1995; POWELL et al., 1996; MACHADO, 1999).

Nos Estados Unidos, DELANNAY et al. (1983) verificaram que 88% do conjunto gênico do germoplasma do norte do país derivava de 10 ancestrais, e 70% do conjunto gênico do germoplasma do sul provinha de somente sete progenitores. Em outro estudo com 14 ancestrais de soja, verificou-se que 25% da base genética foi erodida dos conjuntos gênicos do norte e do sul e que, se a taxa de endogamia entre as linhagens utilizadas nos programas de melhoramento daquele país continuar constante, metade da diversidade genética da soja cultivada será perdida em 58 anos no conjunto gênico do

norte e em 89 anos no conjunto gênico do sul dos Estados Unidos (GIZLICE et al., 1993).

No Brasil, situação semelhante também tem sido encontrada. HIROMOTO e VELLO (1986) relataram que, para a safra agrícola do ano 1983/1984, 100% do conjunto gênico tinha origem em apenas 26 ancestrais, nove desses ancestrais haviam contribuído com 80% do conjunto gênico e apenas quatro progenitores contribuíram com cerca de 50%.

Trabalhos com marcadores moleculares vêm sendo muito utilizados no estudo da variabilidade genética em plantas, por fornecerem melhor entendimento da estrutura genética do germoplasma e, assim, auxiliarem na escolha de progenitores para os programas de melhoramento. ABDELNOOR et al. (1995), utilizando marcadores RAPD (*Random Amplified Polymorphic*), agruparam 38 variedades brasileiras de soja em cinco subgrupos distintos (A1, A2, B1, B2 e B3). PI 60406 e Tanloxi contribuíram com pelo menos 50% do conjunto gênico do grupo A. Já no subgrupo B1, que era composto de 14 variedades, os ancestrais Roanoke, Tokyo e PI 54610 contribuíram com 37% do conjunto gênico. O subgrupo B2 continha quatro variedades que tinham PI 240664 como ancestral comum, o mesmo acontecendo com o subgrupo B3, com seis variedades, que tinham PI 37335 como progenitor comum.

Em outro estudo, utilizando marcadores SSRs - "simple sequence repeats" - nucleares e do cloroplasto, POWELL et al. (1996) investigaram os níveis e padrões de variabilidade em genótipos de *G. soja* e *G. max*. De um total de 700 genótipos analisados com sondas de RFLP, foi detectado que 12 genótipos eram responsáveis por 92% da variabilidade desse conjunto gênico. Quando se analisaram 12 genótipos, juntamente com 37 acessos de *G. soja* e *G. max*, com 11 SSRs nucleares, foram encontrados 129 alelos. Os 12 genótipos foram responsáveis por 40% da variabilidade alélica.

MAUGHAN et al. (1996), estudando 23 acessos de soja cultivada e silvestre, verificaram que 36% dos fragmentos de AFLP (*Amplified Fragment Length Polymorphisms*) gerados eram polimórficos e, desses, 17% eram polimórficos em *G. max* e 31% em *G. soja*. De acordo com o que era esperado, a diversidade genética nos acessos silvestres foi maior que nos cultivados. A análise estatística desses dados agrupou os acessos de *G. soja* separadamente dos de *G. max*; dentro deste último, as variedades de soja

cultivadas se agruparam de forma mais densa, evidenciando-se, mais uma vez, que a base genética da soja cultivada é estreita.

MACHADO (1999) ao analisar 263 acessos de soja, sendo 31 de *G. soja*, 10 de *G. max* adaptados e 222 de *G. max* não-adaptados, utilizando também a técnica de AFLP, verificou que os acessos de *G. soja* apresentaram valor médio de dissimilaridade de 0,58; os de *G. max* não-adaptados, valor de 0,43; e os de *G. max* adaptados, no valor de 0,29. Mais uma vez, demonstrou-se que a soja cultivada apresenta base genética bastante estreita e que os acessos silvestres possuem grande diversidade, que pode ser mais bem explorada pelos programas de melhoramento.

## 2.2. Resistência a doenças em plantas

As plantas, assim como os animais, estão constantemente sendo desafiadas por ataques de patógenos (vírus, bactérias, fungos e nematóides), mas, mesmo assim, eventos de doença são raros. Isso se deve ao fato de que, como outros organismos, as plantas utilizam mecanismos para se protegerem. Há três motivos principais que levam os patógenos ao fracasso. O primeiro é que nem sempre determinada planta é capaz de fornecer todos os requerimentos para o crescimento do patógeno, tornando-se, assim, uma planta não-hospedeira. O segundo é que a planta possui estruturas pré-formadas, que podem ser barreiras físicas (parede celular e cutícula) ou compostos tóxicos (substâncias fenólicas e alcalóides em alta concentração ou proteínas antimicrobianas), que impedem que a infecção tenha sucesso. O último motivo é o reconhecimento, pela planta, do patógeno, o que desencadeia uma série de mecanismos de defesa que vão levar, quase sempre, à morte celular e do tecido no local da infecção, impedindo, assim, que a doença se dissemine pelo resto do organismo. Todos esses três motivos levam a uma interação planta-patógeno incompatível, mas só o último depende da indução de uma resposta, que é controlada por inúmeros genes que estão envolvidos desde a detecção e discriminação do patógeno até a resposta final (JOHAL et al., 1995; KIRÁLY e HORNOK, 1997).

A existência de genes que controlam a resistência a patógenos tem sido usada pelos melhoristas há muito tempo, mas só em 1905, com a redescoberta dos trabalhos de Mendel, Biffen demonstrou que apenas um único *locus* condicionava o caráter de resistência à ferrugem-amarela (*Puccinia striiformis*) em trigo (CRUTE e PINKI, 1996). Só mais tarde foi que Flor elucidou os mecanismos genéticos que controlam as interações entre hospedeiro e patógeno. Com base em seus trabalhos, ele elaborou a teoria gene-a-gene, na qual demonstrou que a resistência do linho (*Linum usitatissimum*) à ferrugem (*Melampsora lini*) era consequência da interação de genes complementares no patógeno e hospedeiro. Os cultivares de linho que têm o gene de resistência dominante (*R*) são resistentes a raças de ferrugem que carregam o gene de avirulência dominante (*Avr*) correspondente, entretanto cultivares que têm o gene que codifica suscetibilidade (*r*), normalmente recessivo, desenvolvem doença. Do ponto de vista molecular, essa hipótese pode ser interpretada como a interação do produto do gene *Avr* com o produto do gene *R*, em uma interação do tipo receptor-ligante, em analogia com o sistema imunológico de vertebrados (FLOR, 1956 e 1971; LAMB, 1994; JACKSON e TAYLOR, 1996; STASKAWICZ et al., 1995).

Nessas interações do tipo gene-a-gene, a resposta de defesa do hospedeiro começa com o reconhecimento ou com a percepção, pela planta, de algumas estruturas ou características químicas do patógeno ou, ainda, com estresse ou dano celular causado pela invasão. Esse reconhecimento ativa a produção ou a liberação de moléculas sinalizadoras que são responsáveis pela indução da resposta final de resistência. Tal resistência ao patógeno é manifestada em várias vias e, freqüentemente, está correlacionada com a resposta de hipersensibilidade (HR). Classicamente, HR é definida como a morte das células do hospedeiro dentro de poucas horas após o contato com o patógeno, formando lesões necróticas que impedem o crescimento do patógeno para tecidos saudáveis (HAMMOND-KOSACK e JONES, 1996; JOHAL et al., 1995).

A indução de HR, ocorrida após a interação de elicitores codificados pelos genes de avirulência (*Avr*) do patógeno com os receptores codificados pelos genes *R* do hospedeiro, é consequência de uma cascata de eventos de transdução de sinal. Essas respostas incluem a combustão oxidativa (liberação

de radicais livres), o reforço da parede celular por lignificação ou deposição de calose, a produção de compostos antimicrobianos (fitoalexina e enzimas hidrolíticas), a alteração do potencial de membrana e a indução de proteínas relacionadas à patogênese (PR) (DIXON et al., 1994; STASKAWICZ et al., 1995; KNOGGE, 1996; DANGL et al., 1996; OUSBOURN, 1996; KAMOUN et al., 1999).

Um dos aspectos da resposta HR foi recentemente esclarecido. Uma das dúvidas era conhecer como os radicais livres ( $O_2^-$ ) e  $H_2O_2$ , liberados após o contato com o patógeno, poderiam atuar como sinalizadores da morte celular. O superóxido ( $O_2^-$ ), primeiro produto da combustão oxidativa, era o candidato menos cotado, já que ele tem meia-vida muito curta e é rapidamente convertido a peróxido de hidrogênio e, além disso, não se difunde rapidamente. O peróxido de hidrogênio ( $H_2O_2$ ), o outro candidato, é tóxico a altas concentrações, mas nem sempre é produzido, durante a resposta de resistência, em quantidades suficientes que possam levar à morte celular.

Dado esse impasse, foi lançada uma hipótese de que haveria outra molécula envolvida que fosse gerada durante a combustão oxidativa. Em macrófagos, células do sistema de defesa de mamíferos, os radicais livres agem juntamente com o óxido nítrico (NO) para matar patógenos bacterianos. DELLEDONNE et al. (1998) e DURNER et al. (1998) indicaram que NO é a molécula que, juntamente com os radicais livres, ativa a transcrição de genes de defesa de plantas e a resposta de hipersensibilidade. DURNER et al. (1998) observaram que, quando folhas de tabaco resistentes a TMV eram infectadas com esse vírus, a atividade da enzima NO sintase aumentava, o que não ocorria em plantas suscetíveis. Além disso, a administração de doadores de NO ativava a expressão de genes relacionados à defesa, os quais codificavam as proteínas PR-1 e fenilalanina amônia liase (PAL). DELLEDONNE et al. (1998), trabalhando com cultura de células de soja em suspensão, verificaram que NO potencializa a indução da morte das células, juntamente com os radicais livres, e atua, independentemente desses intermediários, na indução de genes para defesa. Além disso, inibidores da síntese de NO comprometiam a resposta HR de folhas de *Arabidopsis thaliana* resistentes a *Pseudomonas syringae*, causando o aparecimento da doença e o crescimento bacteriano. A partir desses dois trabalhos e de outros já publicados, DANGL (1998) propôs

um modelo de resposta de defesa contra a invasão de patógenos (Figura 1). Nesse modelo, o NO é responsável, juntamente com os radicais livres, pela ativação do programa de morte celular programada e pela síntese de ácido salicílico (SA).

A resposta HR, no local da infecção, deixa o patógeno isolado do resto da planta. A formação de lesões necróticas, como parte da HR ou de um sintoma da doença causada por um patógeno, é responsável pela indução coordenada de um conjunto de genes. Essa ativação de genes não é feita só localmente, estendendo-se pelo resto da planta e assegurando diminuição da severidade de doenças causadas por todas as classes de patógenos. Esse tipo de resposta é denominado resistência sistêmica adquirida (SAR), cuja indução SAR coincide com aumento do nível endógeno de SA no local da infecção e, sistemicamente, em todos os tecidos não-afetados. Acredita-se que SA seja o principal componente sinalizador da via de resposta SAR (PIETERSE e LOON, 1999; HAMMOND-KOSACK e JONES, 1996).

Vários modelos têm sido propostos para interação entre os produtos dos genes *Avr* e os dos genes *R*. BENT (1996) propôs um modelo para explicar a interação entre *P. syringae* e tomate, envolvendo o gene de avirulência *avrPto* e os genes de resistência *Pto* e *Prf*, que seguem o modelo da interação gene-a-gene. Esse modelo foi proposto com base nos dados experimentais de MARTIN et al. (1994), ROMMENS et al. (1995), SALMERON et al. (1994), ZHOU et al. (1995) e de dados não publicados de outros laboratórios. Nesse modelo, os produtos dos genes *Pto* e *Prf* estão associados à face interna da membrana plasmática. Quando o produto do gene *avrPto* (elicitor) se liga ao complexo *Pto:Prf*, ocorrem autofosforilação de *Pto* e fosforilação de *Pti1* por *Pto*, sendo *Pti* uma segunda cinase na cascata *Pto* cinase. *Pti1* ativa direta ou indiretamente os componentes intermediários da via de transdução de sinal, como o complexo NADPH oxidase, que gera uma liberação de radicais livres, e a liberação de SA. Tais fatores estimulam as respostas de defesa, como a lignificação da parede celular e a expressão de genes que sintetizam proteínas de defesa, como  $\beta$ -1, 3-glucanase e PAL. Outras respostas incluem a abertura de canais iônicos e indução de HR. Uma segunda possível via de sinalização inclui a ativação de *Pti5* por *Pto*, promovendo a expressão de genes de defesa.

Embora esse modelo seja confirmado por muitos dados experimentais, há algumas perguntas a serem respondidas e, também, a necessidade da identificação de mais componentes na via de transdução de sinal. Qual é o tipo de interação química entre Pto e Prf? É Prf quem dá a especificidade de reconhecimento de avrPto? Pto atua sozinho no passo de reconhecimento? Como Pti5 ativa a transcrição dos genes de defesa? Esse modelo pode ser aplicado a outras espécies e tipos de interação?

que existiria uma sequência de eventos de defesa comuns. De acordo com as suas características estruturais, os genes R podem ser agrupados em seis diferentes classes, como mostrado na Tabela 1 (STASKAVICZ et al., 1995; BAKER et al., 1997; HAMMOND-KOSACK & JONES, 1997).

A primeira classe é representada pelo gene *Hm1* do milho, que foi o primeiro gene R a ser clonado (JOHAL & BRIGGS, 1992). Esse gene, que confere resistência à rotação do fungo *Colletotrichum graminicola*, foi isolado por

"transposon tagging" usando-se o vírus de milho  $\mu$  gene *Hm1* codifica uma proteína de 1000 aminoácidos que ativa a toxina HC produzida pelo fungo *C. graminicola*. Essa interação entre o gene-*av* gene, já

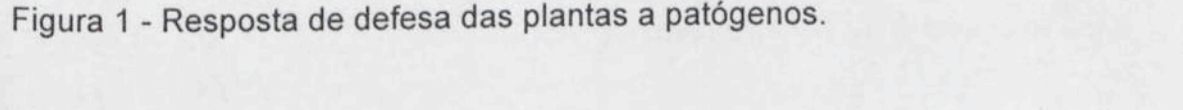
HC, sendo o único gene de resistência de milho que confere resistência à *C. graminicola* (JOHAL & BRIGGS, 1992). A

estratégia para a identificação de genes R em plantas de milho foi fortemente ligado ao gene *Pto* (DODD & JOHAL, 1993). Uma biblioteca

genética construída a partir de milho com o gene *Pto* e o gene *Hm1* foi usada para identificar genes R. A análise da sequência de clones revelou que o gene *Hm1* codifica uma proteína de 1000 aminoácidos que ativa a toxina HC produzida pelo fungo *C. graminicola*. Essa interação entre o gene-*av* gene, já

indica que a proteína Pto pertence a um complexo proteico, junto com Prf, ancorado na membrana que reconhece o elicitor e emite um sinal para iniciar a resposta de defesa (BENT, 1993).

Figura 1 - Resposta de defesa das plantas a patógenos.



Fonte: DANGL (1998)

### 2.3. Genes de resistência (R)

Nos últimos anos, grande número de genes *R* foi clonado em várias espécies de plantas. Embora esses genes sejam responsáveis pela resistência a diversos tipos de patógenos, como bactérias, fungos, vírus e nematóides, os seus produtos gênicos possuem similaridades estruturais, o que leva a pensar que existiria uma seqüência de eventos de defesa comum. De acordo com as suas características estruturais, os genes *R* podem ser agrupados em seis diferentes classes, como mostrado na Tabela 1 (STASKAWICZ et al., 1995; BAKER et al., 1997; HAMMOND-KOSACK e JONES, 1997).

A primeira classe é representada pelo gene *Hm1* de milho, que foi o primeiro gene *R* a ser clonado (JOHAL e BRIGGS, 1992). Esse gene, que confere resistência à raça 1 do fungo *Cochliobolus carbonum*, foi isolado por "transposon tagging", usando-se o transposon *Mu* de milho. O gene *Hm1* codifica uma redutase dependente de NADPH (HTRC), que inativa a toxina HC produzida pelo fungo *C. carbonum*. Essa interação foge do modelo gene-a-gene, já que nenhum componente *Avr* está envolvido na degradação da toxina HC, sendo o único gene de resistência clonado até hoje que difere desse modelo.

A segunda classe de genes inclui *Pto*, que confere resistência à bactéria *P. syringae* pv. *tomato*, que expressa o gene *avrPto*, sendo o primeiro gene a ser clonado que segue o modelo gene-a-gene (MARTIN et al., 1993). A estratégia usada foi a clonagem baseada em mapa. Um marcador de RFLP fortemente ligado ao *locus Pto* foi usado em um "screening" de uma biblioteca genômica construída em YAC (cromossomo artificial de levedura). O clone de YAC identificado foi usado para isolar o cDNA correspondente ao gene *Pto*. A análise da seqüência do clone indicou que esse gene codifica uma proteína cinase do tipo serina/treonina. Os dados experimentais gerados até hoje têm indicado que a proteína *Pto* participa de um complexo protéico, junto com *Prf*, ancorado na membrana que reconhece o elicitor e emite um sinal para outros mensageiros, *Pti1* e *Pti5*, que desencadeiam uma cascata de eventos que culmina na resposta de resistência (BENT, 1996).

A terceira classe inclui os genes que codificam proteínas receptoras citoplasmáticas que contêm domínios repetidos ricos em leucina (LRR) e um sítio de ligação a nucleotídios (NBS). Pertencem a essa classe os genes *RPS2* (BENT et al., 1994; MINDRINOS et al., 1994), *RPM1* (GRANT et al., 1995), *RPP5* (PARKER et al., 1997), *RPP1* (BOTELLA et al., 1998), *RPP8* (McDOWELL et al., 1998), *Prf* (SALMERON et al., 1996), *I2* (ORI et al., 1997), *Mi-1* (VOS et al., 1998), *N* (WHITHAM et al., 1994), *L<sup>6</sup>* (LAWRENCE et al., 1995), *M* (ANDERSON et al., 1997), *locus L* (ELLIS et al., 1999), *Rx* (BENDAHMANE et al., 1999) e a família de genes *RGC2* (MEYERS et al., 1998b). Apesar de todos os membros dessa classe apresentarem os domínios LRR e NBS em comum, a região N-terminal possui algumas diferenças. Alguns membros possuem domínio, denominado TIR, que apresenta alta homologia com o domínio citoplasmático do gene de desenvolvimento *Toll* de *Drosophila* e do gene interleucina 1 (IL 1R), que codifica uma proteína de resposta do sistema imune de mamíferos. Outros membros possuem, na sua região N-terminal, um motivo "leucine zipper" (LZ), ao passo que outros, ainda, só possuem os domínios LRR e NBS (Tabela 1).

A quarta classe é representada por genes que codificam proteínas que estão ancoradas na membrana, possuindo domínios LRRs extracitoplasmáticos e transmembrana (TM). Fazem parte dessa classe os genes *Cf-9* (JONES et al., 1994), *Cf-2* (DIXON et al., 1996), *Cf-4* (THOMAS et al., 1997), *Hs1<sup>pro-1</sup>* (CAI et al., 1997) e *Cf-5* (DIXON et al., 1998).

Na quinta classe são agrupados os genes que codificam uma proteína transmembrana com domínios LRRs extracitoplasmáticos e um domínio celular que possui atividade cinase serina/treonina. O gene *Xa21* (SONG et al., 1995) é o único gene dessa classe clonado até o momento.

Recentemente, outros sete membros da família gênica *Xa21* foram clonados e agrupados em duas classes, de acordo com a similaridade de seqüência (SONG et al., 1997). A classe *Xa21* contém os membros *Xa21*, D e F. A segunda classe, denominada A2, é representada pelos membros A1, A2, C e E, os quais foram testados quanto aos padrões de herança e ao fenótipo de resistência a oito raças de *Xantomonas oryzae* pv. *oryzae*, usando-se plantas transgênicas de arroz (WANG et al., 1998). Observou-se que os membros A1, A2, C, E e F não conferiam nenhum fenótipo de resistência e que

Tabela 1 – Genes R isolados de plantas

Classe	Gene R	Planta	Patógeno	Estrutura*
1	<i>Hm1</i>	Milho	<i>Cochliobolus. Carbonum</i> raça 1	Toxina
2	<i>Pto</i>	Tomate	<i>Pseudomonas. Syringae</i> pv. <i>tomato</i>	Cinase
3	<i>RPS2</i>	<i>Arabidopsis</i>	<i>P. syringae</i> pv. <i>tomato</i>	LZ-NBS-LRR
	<i>RPM1</i>	<i>Arabidopsis</i>	<i>P. syringae</i> pv. <i>maculicola</i>	LZ-NBS-LRR
	<i>RPP8</i>	<i>Arabidopsis</i>	<i>Peronospora parasitica</i>	LZ-NBS-LRR
	<i>Prf</i>	Tomate	<i>P. syringae</i> pv. <i>tomato</i>	LZ-NBS-LRR
	<i>Mi-1</i>	Tomate	<i>Meloidogyne incognita</i> e <i>Macrosiphum euphorbiae</i>	LZ-NBS-LRR
	<i>Rx</i>	Batata	PVX	LZ-NBS-LRR
	<i>RPP5</i>	<i>Arabidopsis</i>	<i>P. parasitica</i>	TIR-NBS-LRR
	<i>RPP1</i>	<i>Arabidopsis</i>	<i>P. parasitica</i>	TIR-NBS-LRR
	<i>N</i>	Tabaco	TMV	TIR-NBS-LRR
	<i>L<sup>6</sup></i>	Linho	<i>Melampsora lini</i>	TIR-NBS-LRR
	<i>M</i>	Linho	<i>M. lini</i>	TIR-NBS-LRR
	locus <i>L</i>	Linho	<i>M. lini</i>	TIR-NBS-LRR
	<i>l2</i>	Tomate	<i>Fusarium oxysporum</i> f. sp. <i>Lycopersici</i> , raça 2	NBS-LRR
	<i>RGC2</i>	Alface	<i>Bremia lactucae</i>	NBS-LRR
4	<i>Cf-9</i>	Tomate	<i>Cladosporium fulvum</i>	LRR-TM
	<i>Cf-2</i>	Tomate	<i>C. fulvum</i>	LRR-TM
	<i>Cf-4</i>	Tomate	<i>C. fulvum</i>	LRR-TM
	<i>Cf-5</i>	Tomate	<i>C. fulvum</i>	LRR-TM
	<i>Hs1<sup>pro-1</sup></i>	Beterraba	<i>Heterodera schachtii</i>	LRR-TM
5	<i>Xa21</i>	Arroz	<i>Xantomonas oryzae</i> pv. <i>oryzae</i>	LRR-TM- Cinase
6	<i>Xa21D</i>	Arroz	<i>Xantomonas oryzae</i> pv. <i>oryzae</i>	LRR

\*Repetições ricas em leucina (LRR), "leucine zipper" (LZ), sítio de ligação a nucleotídeos (NBS), região com homologia a Toll-IL 1R (TIR) e domínio transmembrana (TM).

somente o membro D conferia resistência intermediária àquela conferida por Xa21. Análises de seqüência indicaram que Xa21D codificava uma proteína que tinha perdido os domínios cinase e transmembrana, mas que, mesmo assim, era transcrita e exercia sua função. A perda desses domínios originou-se da integração do retrotransposon *Retrofit*, que introduziu um códon de parada entre o domínio LRR e os domínios transmembrana e cinase, criando, assim, uma nova proteína, com uma provável localização celular diferente, que pode representar uma possível nova classe de genes de resistência.

Esses estudos de estrutura e função de genes *R* evidenciaram que existe grande similaridade de seqüência entre os seus produtos protéicos, o que indica um alto grau de conservação das vias de respostas de defesa. A clonagem dos genes *R* permitiu verificar que alguns domínios estruturais são conservados em diferentes plantas, em genes *R*, para a mesma classe de patógenos, bem como para diferentes classes (vírus, bactérias e fungos). Os produtos protéicos dos genes *R* apresentam os seguintes domínios estruturais conservados: cinase serina-treonina, LRR, NBS, TIR e LZ. A partir dessa informação, podem-se fazer inferências, ligando a estrutura dos domínios e as suas prováveis funções àquelas a que estão relacionados.

A partir da clonagem do gene *Pto*, tornou-se possível visualizar a importância dos eventos de transdução de sinal mediado por cinases na resistência de plantas a doenças. O estado de fosforilação é um mecanismo muito comum nos seres vivos, que o usam para controlar a atividade de proteínas e expressão gênica, mas ainda falta estabelecer relação mais direta entre proteínas de resistência contendo domínios com atividade cinase e reconhecimento do patógeno. Normalmente, proteínas cinases não estão envolvidas em interações do tipo receptor-ligante, que geralmente são exercidas por domínios LRRs, entretanto foi demonstrada, no sistema *Pto-avrPto*, a interação física e altamente específica entre uma proteína cinase e o correspondente produto do gene de avirulência (SCOFIELD et al., 1996; TANG et al., 1996). Uma proteína de resistência que parece encaixar-se no modelo de receptor e sinalizador é a proteína Xa21, que contém os dois domínios que parecem ser as mais importantes para exercerem a função de proteínas de resistência; Xa21 possui o domínio que seria responsável pelo reconhecimento

do patógeno, o domínio LRR, e a porção com atividade cinase, que ativaria a via de transdução de sinal.

O domínio LRR contém leucina ou outros resíduos hidrofóbicos em intervalos regulares, podendo conter prolina e asparagina espaçadas regularmente. Em proteínas de resistência a doenças em plantas, esses domínios estão repetidos ao longo da molécula. Em termos de função, os domínios LRRs de proteínas de levedura, *Drosophila*, humanos e outras espécies têm sido envolvidos na mediação de interações proteína:proteína. Alguns exemplos incluem interações entre componentes intracelulares de uma cascata de transdução de sinal e ligação de hormônios peptídicos e receptores transmembranas.

Devido ao fato do motivo LRR ser recorrente na maior parte dos genes de resistência clonados, ele tem sido apontado como a região da molécula envolvida na especificidade das interações do tipo gene-a-gene e, como consequência, na determinação da especificidade de diferentes alelos. A importância desses domínios LRRs foi demonstrada em experimentos com alelos mutantes de *RPS2* e *RPM1*, que não eram funcionais devido à troca de um único aminoácido dentro da região LRR (BENT et al., 1994; MINDRINOS et al., 1994; GRANT et al., 1995). A deleção de repetições LRRs em alelos mutantes do gene *M* era essencial para a função de resistência e, ou, especificidade (ANDERSON et al., 1997). Outra evidência para o envolvimento dos domínios LRRs no reconhecimento e na ativação das vias de defesa vem do trabalho de WANG et al. (1998), que demonstrou que, mesmo perdendo os domínios transmembrana e cinase, a proteína Xa21D foi capaz de transduzir resposta de resistência parcial de modo raça-específico. Além disso, ELLIS et al. (1999) evidenciaram que a região LRR das proteínas L era a principal responsável pela especificidade gene-a-gene.

Os domínios NBS ocorrem em várias proteínas de mamíferos com atividade de ligação a nucleotídeos e são essenciais para o seu funcionamento. Isso é confirmado, com relação aos genes de resistência, em experimentos com mutantes de *RPS2* (BENT, 1996). Um papel hipotético para os domínios NBS seria a alteração das interações entre os produtos dos genes *R* e outros membros da via de transdução de sinal. No entanto, experimentos de

mutagênese sítio-dirigida dessas regiões são necessários para melhor compreensão do papel dos domínios NBS no funcionamento das proteínas R.

Em *Drosophila*, Toll é uma proteína receptora transmembrana ativada proteoliticamente pelo seu ligante, Spatzle. Quando ativada, Toll interage com a proteína citoplasmática Tube, que então ativa uma cinase serina/treonina, Pelle. A cinase fosforila a proteína inibitória Cactus, que se desliga do fator de transcrição Dorsal, que vai então estimular a formação do eixo dorso ventral. Essa cascata de eventos também ocorre quando a proteína IL 1R, de mamíferos, interage com o seu ligante, IL 1, e induz respostas imunes. Por analogia com essas duas proteínas, é postulado que as proteínas N, L6 e Rpp5, que possuem domínios citoplasmáticos similares ao das proteínas Toll e IL 1R, participariam de uma cascata de eventos de transdução de sinal quando se ligassem a proteínas Avr correspondentes (BAKER et al., 1997; JONES e JONES, 1997).

Os genes *Prf*, *RPS2* e *RPM1*, que têm os motivos LRR e NBS, codificam proteínas com domínios LZ. Esse tipo de motivo tem sido correlacionado com interações proteína-proteína. Alguns fatores de transcrição de eucariotos possuem seqüências LZ, que são importantes na homo e heterodimerização dos fatores de transcrição (BENT, 1996). Entretanto, ainda não se conseguiu nenhum dado experimental que pudesse esclarecer a função de domínios LZ em proteínas R de plantas.

#### 2.4. Evolução de genes R

Normalmente, os genes *Avr* não exercem funções essenciais para a sobrevivência do patógeno, e assim eles podem ser facilmente perdidos, o que permite que haja razoável diversidade dentro da população. Por isso, quando um cultivar que contém um gene *R* que reconhece uma raça específica do patógeno com o correspondente gene *Avr* é plantado em uma grande área, ocorre seleção na população do patógeno de indivíduos que tenham sofrido mutações em *Avr*, aumentando o número daqueles capazes de estabelecer interações de compatibilidade. Por sua vez, essa mudança no patógeno exerce

forte pressão de seleção na população de plantas, em que ocorrerá a seleção de indivíduos capazes de reconhecer essa nova forma do gene *Avr*. Nesse ponto, a situação novamente se invertará, e o ciclo de seleção de novos genes *R* e *Avr* será repetido. É esse o tipo de jogo, entre o patógeno e a planta, responsável pela diversidade de genes *R* encontrada nas populações das espécies hospedeiras e genes de *Avr* que ocorrem nas diferentes raças dos patógenos (JOHAL et al., 1995; STASKAWICZ et al., 1995).

Os mecanismos geradores dessa diversidade nos genes *R* têm sido esclarecidos nos últimos anos por meio de estudos genéticos e moleculares, desenvolvidos em razão do aumento do número de genes clonados. Análises genéticas têm evidenciado que muitos genes *R* pertencem a famílias multigênicas, estando arrançados em grandes blocos formando *loci* complexos. Esses blocos seriam alvo de eventos de recombinação que levariam à criação de novas especificidades. Várias evidências têm levado ao pensamento de que eventos de "crossing-over" desigual são o principal mecanismo de geração de diversidade nas seqüências dos genes *R*.

Muitos genes de resistência clonados até hoje possuem seqüências que codificam domínios LRR repetidos. Os alelos que possuem essas seqüências, devido ao seu caráter repetitivo, podem sofrer "crossing-over" desigual. Tal fato foi observado em mutantes do *locus M* (ANDERSON et al., 1997) em *RPP5* (PARKER et al., 1997) e no tamanho de regiões LRRs dos alelos *L* (ELLIS et al., 1997) e dos genes *RGC2* (MEYERS et al., 1998a e b). Outros estudos evidenciaram que ocorre "crossing-over" desigual também fora da região LRR. Foram detectados eventos de recombinação nas regiões que flanqueavam o *locus Cf-4/9* (PARNISKE et al., 1997). Desse modo, os eventos de recombinação estariam envolvidos na geração de novas especificidades de resistência e no número de cópias dos genes de resistência.

Além dos eventos de recombinação, as mutações de ponto também são importantes na evolução de novas especificidades. Alguns trabalhos têm evidenciado a presença de mutações de ponto em regiões do domínio LRR que estariam expostas na proteína (PARNISKE et al., 1997; WANG et al., 1998; MEYERS et al., 1998b). O acúmulo dessas substituições poderia afetar a ligação da proteína *R* à proteína *Avr* e, assim, contribuir para novas especificidades.

Em resumo, vários são os mecanismos que podem causar diversidade nos genes de resistência. Um dos mecanismos principais é o "crossing over" desigual, que pode gerar diferentes especificidades, seja na criação de novos genes que podem evoluir independentemente, seja, então, no aumento do número de LRRs, sobre as quais mutações podem agir.

## MATERIAL E MÉTODOS

### 3.1. Material vegetal

Neste trabalho, foram utilizados DNA de oito genótipos de *Glycine max*: Noir, Hardin, Bonus, PI 410936, PI 410937, PI 51293, PI 153293 e N85-2176; e dois genótipos de *Glycine soja*: PI 440913 e PI 41782. O DNA desses genótipos foi obtido na DuPont (EUA). Além disso, foi utilizada a variedade FT-Cristalina, obtida do Genótipo de Grupo de Melhoramento de Soja do BIOAGRO/UFV. Todo o material vegetal, como folhas, haste e raízes coletados no estágio de desenvolvimento V3 da variedade FT-Cristalina, foi congelado em nitrogênio líquido logo após a sua coleta e armazenado a  $-80^{\circ}\text{C}$ , até que fosse utilizado para extração de ácidos nucleicos.

### 3.2. Isolamento de DNA plasmidial

O isolamento do DNA plasmidial em pequena escala foi feito por meio de dois métodos diferentes, de acordo com a finalidade: por ruptura com formaldeído-clorofórmio (SAMBROOK et al., 1989; AUSUBEL et al., 1989) para diagnóstico de clonagem e construção de mapa de restrição e por purificação

### 3. MATERIAL E MÉTODOS

#### 3.1. Material vegetal

Neste trabalho, foram utilizados DNA de oito genótipos de *Glycine max*: Noir, Hardin, Bonus, PI 230970, PI 416937, PI 51293, PI 153293 e N85-2176; e dois genótipos de *Glycine soja*: PI 440913 e PI 81762. O DNA desses genótipos foi obtido na DuPont (EUA). Além disso, foi utilizada a variedade FT-Cristalina, obtida do Germoplasma do Grupo de Melhoramento de Soja do BIOAGRO/UFV. Todo o material vegetal, como folhas, haste e raízes coletados no estágio de desenvolvimento V3 da variedade FT-Cristalina, foi congelado em nitrogênio líquido logo após a sua coleta e armazenado a  $-80^{\circ}\text{C}$ , até que fosse utilizado para extração de ácidos nucleicos.

#### 3.2. Isolamento de DNA plasmidial

O isolamento do DNA plasmidial, em pequena escala, foi feito por meio de dois métodos diferentes, de acordo com a finalidade: por purificação com fenol: clorofórmio (SAMBROOK et al., 1989; AUSUBEL et al., 1998), para diagnóstico de clonagem e construção de mapa de restrição; e por purificação

em colunas contidas no "QIAprep Spin Miniprep Kit" (Qiagen Inc.), para seqüenciamento e subclonagem. O isolamento do DNA plasmidial, em média escala, foi feito por meio de purificação em colunas tipo 100 contidas no "QIAGEN Plasmid Midi Kit" (Qiagen Inc.). Em todos os três procedimentos, inicialmente, as células foram rompidas pelo método de lise alcalina, sendo o volume das soluções ajustado de acordo com o volume inicial da cultura. Basicamente, as bactérias transformadas cresceram em meio LB (Sigma) a 37°C, contendo 100 µg/mL de ampicilina, por 12 a 16 horas. Após centrifugação a 5.000 g por dois minutos, a 4°C, as células foram ressuspensas na solução I [glicose 50 mmoles/L, Tris-HCl 25 mmoles/L pH 8,0 e EDTA (ácido etilenodiaminotetracético) 20 mmoles/L]. Em seguida, a solução de lise [NaOH 0,2 mol/L e SDS (dodecil sulfato de sódio) 1%] foi adicionada e a mistura incubada por cinco minutos, à temperatura ambiente. A essa mistura foi adicionado acetato de potássio, à concentração final de 1 mol/L, prolongando-se a incubação por mais 15 minutos. O DNA foi separado dos resíduos insolúveis, como parede celular, e do DNA cromossômico, por meio de centrifugação a 11.000 g, por 10 minutos. O sobrenadante recebeu diferentes tratamentos, de acordo com o método de isolamento de DNA plasmidial usado.

Para isolamento de DNA plasmidial, usando-se as colunas da Qiagen Inc., a lise alcalina foi conduzida conforme descrito anteriormente, à exceção da RNase A, que foi adicionada à solução I na concentração final de 100 µg/mL. Após a lise alcalina, o DNA plasmidial foi isolado do sobrenadante, usando-se a resina pré-empacotada QIAprep Spin e QIAGEN tipo 100 (Qiagen Inc.), conforme as recomendações do fabricante.

No método de preparo de DNA plasmidial em pequena escala, para diagnóstico e mapa de restrição após a lise alcalina, as amostras foram desproteinizadas por tratamento com fenol: clorofórmio: álcool isoamílico (24:24:1), seguido de clorofórmio: álcool isoamílico (24:1), sendo os ácidos nucléicos precipitados com etanol a uma concentração final de 70% e ressuspensos em TE contendo 50 µg/mL de RNase A.

### 3.3. Isolamento de RNA total

Todas as etapas para isolamento do RNA foram conduzidas a 4°C, sendo todas as soluções e água usadas tratadas com dietilpirocarbonato (DEPC) 0,1%, um inibidor de RNase, a 37°C, por três horas, e, em seguida, autoclavadas.

Cerca de 200 mg de raízes, hastes ou folhas da variedade de soja FT-Cristalina, no estágio de desenvolvimento V3, foram triturados em nitrogênio líquido. A cada amostra foram adicionados 500 µL de tampão de extração (citrato de sódio 25 mmoles/L, pH 7,0; isotiocianato de guanidina 4 moles/L; lauril sarcosine de sódio 1,5%; e β-mercaptoetanol 100 mmoles/L) e 500 µL de fenol: clorofórmio: álcool isoamílico (24:24:1) (WADSWORTH et al., 1988). A amostra foi vigorosamente agitada em vórtex, por 15 segundos. As fases orgânica e aquosa foram separadas por centrifugação a 16.000 g, por cinco minutos. A fase aquosa foi reextraída com fenol: clorofórmio: álcool isoamílico (24:24:1) por mais duas vezes. A fase aquosa foi adicionada de 400 µL de LiCl 6 mmoles/L e incubada no gelo por, no mínimo, 60 minutos. O RNA foi coletado por centrifugação a 16.000 g por 10 minutos e o pelete, ressuspensão em 1 mL de LiCl 3 moles/L, sendo, então, centrifugado por 10 minutos a 4°C. Essa fase de lavagem foi repetida mais duas vezes, nas mesmas condições. O precipitado foi ressuspensão em 400 µL de acetato de potássio 2% e o RNA, reprecipitado pela adição de 1 mL de etanol e incubação por 12 horas, a -20°C. O RNA foi separado por centrifugação a 16.000 g por 15 minutos, lavado em etanol 70% e ressuspensão em 50 µL de água. O RNA foi quantificado por espectrofotometria a 260 nm e armazenado a -80°C até o seu uso. A integridade do RNA isolado foi avaliada por eletroforese em géis de agarose 1,8% contendo 0,1 mg/mL de brometo de etídio.

### 3.4. Isolamento de DNA genômico

A extração do DNA foi feita conforme descrito por DOYLE e DOYLE (1990). O DNA foi ressuspensão em H<sub>2</sub>O e armazenado a -20°C até a sua utilização. Tanto a quantidade de DNA nas amostras quanto a sua integridade foram avaliadas por eletroforese em géis de agarose 0,8% contendo 0,1 mg/mL de brometo de etídio.

### 3.5. Isolamento de fragmentos de DNA de géis de agarose

O isolamento e a purificação de fragmentos de DNA de géis de agarose foram feitos pela técnica de adsorção de DNA em membrana de sílica-gel, utilizando-se o "QIAquick Gel Extraction Kit" da Qiagen Inc., conforme as recomendações do fabricante. Após a purificação, os fragmentos de DNA foram quantificados em géis de agarose 1% contendo 0,1 mg/mL de brometo de etídio.

### 3.6. Construção da biblioteca genômica da variedade de soja FT-Cristalina

A biblioteca genômica foi construída, utilizando-se o vetor λEMBL3 fornecido no "Lambda EMBL3 Vector Kit" (Stratagene). O DNA genômico usado foi extraído da variedade de soja FT-Cristalina. Todos os procedimentos para construção da biblioteca foram feitos segundo as recomendações do fabricante (Stratagene). O DNA genômico foi digerido parcialmente com *Sau3A* I. Os fragmentos de 9 a 23 kb foram separados em um gradiente de sacarose de 10 a 40% e, posteriormente, ligados ao vetor λEMBL3 pré-digerido com *BamH* I, fornecidos no "kit". O DNA ligado foi empacotado, usando-se o extrato de empacotamento "Gigapack<sup>®</sup> II Gold Packaging Extracts" (Stratagene), e

infectado na bactéria hospedeira *Escherichia coli* XL1-Blue MRA (P2) (Stratagene) para seleção dos fagos recombinantes. O título da biblioteca foi  $4 \times 10^6$  pfu/mL.

### 3.7. "Screening" da biblioteca genômica

Para o "screening" da biblioteca, as células hospedeiras *E. coli* XL1 Blue MRA foram crescidas a 37°C em meio LB (Sigma) contendo 0,2% de maltose e  $\text{MgSO}_4$  10 mmoles/L até atingirem  $\text{OD}_{600}$  entre 0,2 e 0,5 (SAMBROOK et al., 1989; AUSUBEL et al., 1998). Em seguida, as células foram sedimentadas por centrifugação a 3.000 g a 4°C, por 10 minutos, e ressuspensas em volume de  $\text{MgSO}_4$  10 mmoles/L, correspondente a 1/10 do volume inicial da cultura. Em seguida, 6,9  $\mu\text{L}$  ( $3 \times 10^4$  pfu) da biblioteca genômica de FT-Cristalina foram incubados com 600  $\mu\text{L}$  de células hospedeiras (XL1 Blue MRA) a 37°C, por 15 minutos. Às células, foram adicionados 6 mL de meio semi-sólido (meio LB e 0,7% de agarose), após o que foram vertidas em placas de Petri esterilizadas de 150 mm contendo meio LB. As placas foram incubadas a 37°C por, no mínimo, oito horas. Em seguida, as placas foram esfriadas a 4°C por duas horas. O DNA das placas de lise foi transferido para membranas de náilon por dois minutos. Em seguida, a membrana foi incubada em uma solução de desnaturação ( $\text{NaCl}$  1,5 mol/L e  $\text{NaOH}$  0,5 mol/L) por dois minutos e, depois, na solução de neutralização ( $\text{NaCl}$  1,5 moles/L e  $\text{Tris-HCl}$  0,5 mol/L, pH 8,0) por cinco minutos e por 30 segundos em  $\text{Tris-HCl}$  0,2 mol/L, pH 7,5 e SSC 2X ( $\text{NaCl}$  300 mmoles/L e citrato de sódio 30 mmoles/L, pH 7,0). Logo após, o DNA foi imobilizado nas membranas, usando-se uma irradiação de 120.000 microjoules/ $\text{cm}^2$  por 30 segundos, no "Stratalinker UV Crosslinker" (Stratagene). As membranas foram secadas e armazenadas à temperatura ambiente até o seu uso. As placas foram guardadas a 4°C para coleta das pfus positivas.

Para hibridização, as membranas, depois de secadas, foram novamente hidratadas em  $\text{H}_2\text{O}$  e, depois, incubadas em 10 mL de "QuikHyb

Hybridization Solution" (Stratagene) por 15 minutos a 50 ou 55°C. Em seguida, adicionaram-se a elas 100 ng de sonda marcada e 1 mg de DNA de esperma de salmão, fazendo uma incubação a 50 ou 55°C, por duas horas. Após o período de incubação, a membrana foi lavada em SSC 0,15X e SDS 0,1%, por uma vez à temperatura ambiente e duas vezes por 15 minutos, a 50 ou 55°C. Todos os passos de detecção da sonda foram conduzidos à temperatura ambiente, usando-se o "Kit Illuminator Chemiluminescent Detection System" (Stratagene). O DNA hibridizado foi visualizado por auto-radiografia.

As regiões circulares em torno das pfus positivas foram cortadas do ágar e incubadas em 1 mL de SM (NaCl 0,58%; MgSO<sub>4</sub> 0,2%; Tris-HCl 50 mmoles/L, pH 7,5; e gelatina 0,1%) contendo 0,2% de clorofórmio. As pfus resultantes do primeiro "screening" foram usadas em um segundo e terceiro "screening", com o objetivo de isolar o clone desejado.

### 3.8. Marcação e obtenção das sondas para "screening"

As sondas de RFLP usadas para o "screening" foram cedidas pelo Dr. Rafalski (DuPont, EUA) e pertencem a uma biblioteca genômica da variedade de soja Bônus, clonada no sítio *Pst* I do vetor pBluescript (STRATAGENE) (RAFALSKI e TINGEY, 1993). A sonda A45-10 foi obtida pela digestão do clone de RFLP A45-10 com a enzima *Pst* I, sendo o inserto de aproximadamente 2,1 kb isolado de géis de agarose, conforme descrito anteriormente. A sonda A53-09 foi obtida pela amplificação do inserto do clone de RFLP A53-09 com os "primers" T3 (5'-AATTAACCCTCACTAAAGGG-3') e 5309A (Tabela 2); o fragmento de aproximadamente 400 pb foi purificado, usando-se o "QIAquick PCR Purification Kit" (Qiagen Inc.), conforme recomendações do fabricante. A sonda A75-10 foi obtida pela amplificação do clone de RFLP A75-10 com os "primers" T7 (5'-GTAATACGACTCACTATAGGGC-3') e 7510A (Tabela 2); o fragmento de aproximadamente 500 pb também foi purificado, usando-se o "QIAquick PCR Purification Kit" (Qiagen Inc.). As sondas foram marcadas com flúor-12-dUTP,

Tabela 2 - "Primers" usados nas reações de PCR, seqüenciamento e RT-PCR

Clone	Nome	Posição	Seqüência (5' - 3')
A53-09	5309A	357 a 376	CAGCGAGAAGGAAAGGAATG
	5309B	357 a 376	CATTCCTTTCTTCTCGCTG
	5309C	723 a 742	GGTCCTAGCCCTCCAACT
	5309D	82 a 101	CTCCAATGAGCCTAGGAAGC
	5309E	604 a 624	AATGGTGGTGGTTGCAAATG
A45-10	4510A	474 a 493	CATTTGCAGACACACTTGGGC
	4510B	474 a 493	GCCAAGTGTGTCTGCAAATG
	4510T7B1	1.011 a 1.028	GGCGGGTGAAAGATGACTT
	4510T3C1	1.889 a 1.912	TCAAGTCTCAAGCTCGTCTAG TTC
	4510T3F	2.089 a 2.109	CAGAATCCTATGCAAGCTCCA
	4510T7F	31 a 50	TTGACTGGTGTGTGCCTGT
A2-08	208A	2.003 a 2.028	TTCCTGAATAAAGAACTTACTT CTTG
	208B	254 a 274	CCATAGAGCATGTGTGATGAG
	208T3A2	1.616 a 1.643	TAAAAGGAAGGAAAAGAAAAT GATAAAA
	208T7B2	529 a 551	AACAAACTAACGTCCGTGAAA AA
	208T3A3	1.349 a 1.374	GAATTTTTATTGTCAAACAATT AGCAA
	208T7B3	886 a 906	CCCATTTTAAAACCTCAATGGC
	208T3F	2.240 a 2.259	TCACCATCCTCAGTCCCTTC
	208T7F	25 a 44	ACCCTCTTGTGTACGTTCC
A75-10	7510A	2.174 a 2.193	ACAAGCAGCCCATCCAAATTC
	7510B	2.171 a 2.193	GAATTTGGATGGCTGCTTGT GG
	7510C	461 a 484	CATGGATATGGTTGGAACTT TTT
	7510T3F	3 a 25	GCAGGTAAATGTTTCAGCTGT TT
	7510T7F	2.634 a 2.653	GCAGAGCCATCTTCCAGTTC
	7510T7B2	2.422 a 2.441	TCAGAGAGCACAGTGATGGC
	7510T7B3	1.670 a 1.691	TTTGCTGGGTAATTATGTTCA
	7510T3C2	175 a 195	TGGTTAGAGGAGGCATGACTT
7510T3C3	832 a 911	CACTGGAGGAAGAAGCATGA	
7510T3C4	971 a 992	AGTCAAGCTCAGACATGCTGT C	
Contíguo GR	2R	1.894 a 1912	AAGCTGGTTGTTCTGCAGGT
	4A	4.079 a 4.098	CACCGAGTGTGTATTCCAG
	4B	2.592 a 2.610	GAAAGTGCTGAATGTTGCCA

utilizando-se "Primer-It<sup>®</sup> Fluor Fluorescence Labeling Kit", da Stratagene, conforme recomendações do fabricante.

### 3.9. Isolamento de DNA de fagos recombinantes

Para isolamento de DNA de fagos recombinantes contendo o inserto dos clones selecionados durante o "screening", as células hospedeiras *E. coli* XL1 Blue MRA cresceram a 37°C em meio LB (Sigma), acrescido de 0,2% de maltose e 10 mmoles/L de MgSO<sub>4</sub>, até atingir OD<sub>600</sub> de 0,2 a 0,5 (SAMBROOK et al., 1989; AUSUBEL et al., 1998). Em seguida, as células foram sedimentadas por centrifugação a 3.000 g, a 4°C, por 10 minutos, e ressuspensas em volume de MgSO<sub>4</sub> 10 mmoles/L, correspondente a 1/10 do volume inicial da cultura. Os fagos recombinantes, aproximadamente 5.000 pfus, ressuspensos em SM, foram incubados com 200 µL de células hospedeiras a 37°C, por 15 minutos. As células foram diluídas em 50 mL de meio NZY, pH 7,5 (caseína hidrolisada 1%, NaCl 0,5%, MgSO<sub>4</sub> 0,2% e bactoextrato de levedura 0,5%) e crescidas a 37°C, sob forte agitação, até que a lise das células fosse observada. A fim de assegurar a lise completa, 100 µL de clorofórmio foram adicionados, prolongando-se a incubação por mais 10 minutos, a 37°C. Após a adição de 50 µg de DNase I (Stratagene) e 50 µg de RNase A (Sigma), a incubação prosseguiu por mais 30 minutos, a 37°C. Em seguida, foi-lhe adicionado NaCl, a uma concentração final de 1 mol/L, incubando-se a solução a 0°C, por uma hora. O resíduo insolúvel foi removido por centrifugação a 11.000 g, por 10 minutos, a 4°C. Ao sobrenadante foi adicionado PEG (polietileno glicol) 8.000, a uma concentração final de 10%, sendo a mistura incubada a 4°C, por 12 horas. Os fagos foram coletados por centrifugação a 10.000 g, por 30 minutos, a 4°C, e ressuspensos em 500 µL de SM. Em seguida, foi adicionado à suspensão um volume igual de clorofórmio, e após a centrifugação, por cinco minutos, adicionaram-se à fase aquosa 20 µL de EDTA 0,5 mol/L, pH 8,0; 25 µL de SDS 10%; e 10 µL de Pronase (50 mg/mL), após o que foi feita a incubação a 37°C, por 60 minutos. O DNA

recombinante foi isolado por meio de extração com fenol: clorofórmio: álcool isoamílico (24:24:1), reextração com clorofórmio: álcool isoamílico (24:1) e precipitação pela adição de acetato de amônio e etanol, nas concentrações finais de 80 mmoles/L e 70%, respectivamente. O precipitado foi sedimentado por centrifugação a 16.000 g por 15 minutos, lavado em etanol 70%, secado à temperatura ambiente e ressuspenso em 200  $\mu$ L de H<sub>2</sub>O.

### 3.10. Reações de amplificação por PCR

As reações de amplificação foram conduzidas no termociclador da PE Applied Biosystems, modelo 9.600, com período inicial de desnaturação a 94°C, por um minuto, seguido de 35 ciclos (94°C por 30 segundos, 55°C por um minuto e 72°C por dois minutos) e período adicional de polimerização a 72°C, por sete minutos. As reações foram realizadas em um volume final de 100  $\mu$ L contendo 2,5 mmoles/L de cada dNTP, 1 U de *Taq* DNA Polimerase, 0,4 pmol de cada "primer", 50 ng DNA-molde, KCl 50 mmoles/L, Tris-HCl 10 mmoles/L, H<sub>2</sub>SO<sub>4</sub> 8,3 e MgCl<sub>2</sub> 1,5 mmol/L.

Os produtos das reações foram separados por eletroforese em géis de agarose contendo brometo de etídio e visualizados por meio de luz ultravioleta.

### 3.11. Seqüenciamento do DNA

O DNA foi seqüenciado, usando-se o "ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit" comercializado pela PE Applied Biosystems. Esse "kit" é adequado para o seqüenciamento de DNA, usando-se o "ABI PRISM 310 Genetic Analyzer", da PE Applied Biosystems. As reações de seqüenciamento se baseiam na técnica de seqüenciamento por terminação de cadeia por dideoxynucleotídios (ddNTPs), descrita por SANGER et al. (1977).

As reações de amplificação foram conduzidas pela amplificação linear por PCR, usando-se o termociclador da PE Applied Biosystems, modelo 9.600, programado para um período inicial de desnaturação a 96°C, por dois minutos, seguido de 30 ciclos (96°C por 30 segundos, 50°C por 20 segundos e 60°C por quatro minutos). As reações de seqüenciamento foram realizadas em um volume final de 20 µL contendo 8 µL de "mix" - fornecido no "kit" de seqüenciamento -, 10 pmoles de um "primer" e 200 a 500 ng do DNA a ser seqüenciado. Os produtos de extensão, produzidos na reação de seqüenciamento, foram purificados pela adição de 80 µL de isopropanol 75% e pela incubação, à temperatura ambiente, por 15 minutos. Os fragmentos foram coletados por centrifugação a 16.000 g por 20 minutos. O pelete foi lavado com 250 µL de isopropanol 75%, centrifugado por cinco minutos e secado à temperatura ambiente. O pelete foi ressuspenso em 15 µL de "Template Suppression Reagent" (PE Applied Biosystems), desnaturados a 95°C por cinco minutos e mantidos no gelo até a sua aplicação. Os produtos da extensão foram separados por eletroforese capilar, utilizando-se como matriz o polímero POP6 (PE Applied Biosystems), e a fluorescência emitida foi coletada por uma câmara CCD, sendo a informação processada pelo "ABI PRISM 310 Genetic Analyzer" (PE Applied Biosystems). Os dados foram automaticamente convertidos em seqüências de deoxinucleotídeos.

As seqüências geradas foram montadas em contíguos, alinhadas e analisadas, usando-se o programa DNASTAR (DNASTAR Inc.). A busca por similaridades com outras seqüências depositadas no "GenBank" foi feita, usando-se o programa BLAST (ALTSCHUL et al., 1997) e a predição de possíveis ORFs (seqüência de leitura aberta) com o programa ORF Finder, ambos NCBI ("National Center for Biotechnology Information") (<http://www.ncbi.nlm.nih.gov>). Os possíveis sinais de poliadenilação foram localizados, usando-se o programa "Gene-Finder" (<http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html>). A predição dos possíveis domínios protéicos das ORFs foi feita, usando-se o programa Pfam, do "Sanger Center" (<http://www.sanger.ac.uk/Software/Pfam/search.shtml>).

O programa usado para desenhar os “primers” que seriam usados nas reações de PCR, seqüenciamento e RT-PCR foi o “Primer3 Input Program” (<http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi>).

### 3.12. “Southern Blot”

O DNA dos clones genômicos (2 µg) ou DNA total (10 µg) de soja foram digeridos com enzimas de restrição adequadas, sendo os fragmentos separados em gel de agarose 0,7%. Após a eletroforese, o gel foi incubado por 15 minutos na solução de desnaturação, por dois períodos de 30 minutos cada. O gel foi, então, transferido para a solução de neutralização, e o DNA foi transferido por capilaridade, para uma membrana de náilon em SSC 10X, por 18 horas (SAMBROOK et al., 1989; AUSUBEL et al., 1998). Após a transferência, a membrana foi lavada em SSC 2X por cinco minutos e secada ao ar, sendo o DNA imobilizado no “Stratalinker UV Crosslinker” (Stratagene). A membrana foi pré-hibridizada na solução de hibridização (NaCl 0,5 mol/L, reagente de bloqueio 4% - Amersham-Pharmacia e “Hybridization Buffer” - Amersham-Pharmacia), a 55°C, por 15 minutos. A reação de hibridização foi conduzida a 55°C por, no mínimo, 18 horas, em solução de hibridização contendo aproximadamente 10 ng/mL de sonda. Esta foi marcada, seguindo-se as recomendações do fabricante contidas no manual do “Gene Images Alkphos Direct Labelling”, da Amersham-Pharmacia. Após o período de incubação, a membrana foi lavada na solução de lavagem primária (uréia 2 moles/L; SDS 0,1%; fosfato de sódio 50 mmoles/L pH 7,0; NaCl 150 mmoles/L; MgCl<sub>2</sub> 10 mmoles/L; e reagente de bloqueio 0,2%) por duas vezes, durante 10 minutos a 55°C.

Todos os passos de detecção da sonda foram conduzidos à temperatura ambiente. A membrana foi incubada na solução de lavagem secundária (Tris-HCl 50 mmoles/L, pH 10,0; NaCl 100 mmoles/L e MgCl<sub>2</sub> 2 mmoles/L) por dois períodos de cinco minutos cada. Em seguida, a membrana foi imersa e incubada por cinco minutos no reagente de detecção

CDP-Star™, fornecido pelo “Gene Images Alkphos Detection System” (Amersham-Pharmacia). Após esse período, um filme de raios-X foi exposto à membrana e o DNA hibridizado, visualizado por auto-radiografia.

### 3.13. RT-PCR

As amostras de RNA total de folha, raiz e haste da variedade FT-Cristalina, no estágio V3 de desenvolvimento, foram usadas nas reações de RT-PCR. Todas as amostras, antes de serem amplificadas, foram tratadas com “RQ1 RNase-free DNase” (Promega), conforme as recomendações do fabricante. As amostras foram incubadas em tampão da “RQ1 RNase-free DNase” 1X (Tris-HCl 40 mmoles/L, pH 8,0; MgSO<sub>4</sub> 10 mmoles/L e CaCl<sub>2</sub> 1 mmol/L) por 30 minutos a 37°C, extraídas com igual volume de fenol: clorofórmio: álcool isoamílico (24:24:1), reextraída com clorofórmio: álcool isoamílico (24:1) e precipitadas pela adição de acetato de sódio e etanol. O RNA foi incubado a -20°C por, no mínimo, 12 horas e, depois, coletado por centrifugação a 16.000 g por 30 minutos. O pelete foi lavado em etanol 70%, secado à temperatura ambiente e ressuspenso em H<sub>2</sub>O.

A primeira fita de cDNA foi sintetizada, usando-se o “SuperScript Kit Preamplification System for First Strand cDNA Synthesis” (Life Technologies), de acordo com as recomendações do fabricante. As amostras de RNA total (2,5 µg) foram incubadas com 0,5 µg de oligo (dT)<sub>12-18</sub> a 70°C por 10 minutos e, em seguida, no gelo. Logo após, foram adicionados às amostras tampão PCR 1X (Tris-HCl 20 mmoles/L, pH 8,4 e KCl 50 mmoles/L), MgCl<sub>2</sub> 5 mmoles/L, os quatro dNTPs 0,5 mmol/L cada e DTT (ditiotreitól) 5 mmoles/L, as quais foram incubadas a 42°C por cinco minutos. Em seguida, foram adicionados 200 U da enzima “SuperScript II RNase H<sup>-</sup> Reverse Transcriptase” (Life Technologies), sendo as amostras incubadas por mais 50 minutos a 42°C. Após a síntese da primeira fita de cDNA, o RNA foi degradado pela adição de 2 U de RNase H (Life Technologies) a 37°C, por 20 minutos. Para cada amostra, foi feito um

controle negativo, que continha todos os reagentes, menos a Transcriptase Reversa.

As reações de síntese da segunda fita e de amplificação foram conduzidas no termociclador da PE Applied Biosystems, modelo 9.600, com período inicial de desnaturação a 94°C por um minuto, seguidas de 35 ciclos (94°C por 30 segundos, 55°C ou 50°C por um minuto e 72°C por dois minutos) e de um período adicional de polimerização a 72°C por sete minutos. As reações de amplificação foram realizadas em um volume final de 25 µL contendo 2,5 mmoles/L de cada dNTP, 1 U de *Taq* DNA Polimerase, 0,4 pmol de cada "primer", 2,5 µL da reação de síntese da primeira fita, KCl 50 mmoles/L, Tris-HCl 10 mmoles/L, pH 8,3 e MgCl<sub>2</sub> 1,5 mmol/L. Nessa etapa foi feita uma reação-controle, usando-se os "primers" de actina 3 de soja (SHAH et al., 1982), F - 5'-CCCCTCAACCCAAAGGTCAACAG-3' e R - 5'-GGAATCTCTCTGCCCAATTGTG-3', com as finalidades de verificar a contaminação com DNA genômico e normalizar as concentrações de cDNA-molde. Na presença de DNA genômico de soja, a reação de amplificação, com o "primer" de actina 3, produz um fragmento de 520 pb, ao passo que na presença de cDNA se observa a presença de um fragmento de 440 pb, devido à presença de um intron de 80 pb.

### 3.14. Clonagem molecular

O vetor usado nas subclonagens de fragmentos com extremidades coesivas foi pUC18, previamente clivado com a enzima de restrição apropriada e desfosforilado com a enzima fosfatase alcalina de intestino de bezerro (Life Technologies), ao passo que nos fragmentos com extremidades abruptas foi o pGEM-T, usando-se o "pGEM-T Easy Kit" (Promega).

As reações de ligação dos fragmentos de DNA ao vetor foram feitas, usando-se a enzima T4 DNA Ligase (Life Technologies) em volume final de 10-15 µL, de acordo com as técnicas de clonagem descritas por SAMBROOK et al. (1989) e AUSUBEL et al. (1998). As reações de ligação foram conduzidas

em Tris-HCl 50 mmoles/L, pH 7,6; MgCl<sub>2</sub> 10 mmoles/L; ATP 1 mmol/L; DTT 1 mmol/L; PEG 8.000 5%; o fragmento de DNA a ser clonado e o vetor, na relação 3:1; e T4 DNA Ligase. A reação foi incubada a 4°C, por 24 horas.

As células competentes, *E. coli* DH5 $\alpha$ , foram preparadas de acordo com INOUE et al. (1990). Essas células foram crescidas em meio SOB (bactotripton 2%; extrato de levedura 0,5%; NaCl 10 mmoles/L; KCl 2,5 mmoles/L; MgCl<sub>2</sub> 10 mmoles/L; e MgSO<sub>4</sub> 10 mmoles/L, pH 7,0) até atingirem uma OD<sub>600</sub> de 0,6, sendo posteriormente incubadas a 0°C, por 10 minutos, e concentradas por centrifugação a 2.500 g por 10 minutos, a 4°C. As células foram ressuspensas em meio TB {PIPES (piperazina-N,N'-bis [2-ácido etanossulfônico] 10 mmoles/L; MnCl<sub>2</sub> 55 mmoles/L; CaCl<sub>2</sub> 15 mmoles/L; e KCl 250 mmoles/L, pH 6,7} e incubadas por mais 10 minutos a 0°C. As células foram coletadas por centrifugação a 2.500 g por 10 minutos, a 4°C, e ressuspensas, novamente, em TB e DMSO (dimetilsulfóxido de sódio), a uma concentração final de 7%. As células foram novamente incubadas a 0°C por 10 minutos e, em seguida, aliqüotadas e armazenadas a -80°C até o seu uso.

Para as transformações, a 200  $\mu$ L de células competentes foram adicionados 10 a 15  $\mu$ L da reação de ligase, sendo a suspensão mantida a 0°C por 30 minutos. Após o choque térmico de dois minutos a 42°C, 800  $\mu$ L de meio SOB foram adicionados, prolongando-se a incubação por mais uma hora a 37°C. As células foram concentradas por centrifugação, ressuspensas em 100  $\mu$ L de meio SOB e espalhadas em placas contendo meio LB com ampicilina a 100  $\mu$ g/mL, X-Gal (5-bromo-4-cloro-3-indolil- $\beta$ -D-galactosídeo) e IPTG (isopropiltio- $\beta$ -galactosídeo), para seleção dos clones transformantes e recombinantes. As colônias brancas foram tocadas levemente com a ponta de uma ponteira P10, inoculadas em uma placa de referência e incubada em uma solução contendo todos os componentes necessários para amplificar o inserto. Cada reação continha tampão de PCR 1X, cada um dos dNTPs a 2,5 mmoles/L, "M13 Forward Primer" (5'-CCCAGTCACGACGTTGTTAAACG-3') e "M13 Reverse Primer" (5'-AGCGGATAACAATTTACACAGG-3'), cada um a 10 pmoles e 1 U da enzima *AmpliTaq* DNA Polimerase (PE Applied Biosystems), em um volume final de 25  $\mu$ L. As reações de amplificação foram conduzidas no termociclador da PE Applied Biosystems, modelo 9.600, com

período inicial de lise a 94°C, por 10 minutos, seguidas de 30 ciclos (94°C por 30 segundos, 55°C por um minuto e 72°C por um minuto) e um período adicional de polimerização de 72°C por sete minutos. A placa de referência foi incubada a 37°C por 16 horas, e os fragmentos resultantes da reação de PCR foram separados em um gel de agarose 1%. As colônias que continham o inserto de interesse foram crescidas em 2 mL de meio LB contendo ampicilina. O DNA plasmidial, isolado das bactérias transformadas, foi digerido com as enzimas de restrição apropriadas e separado por eletroforese em géis de agarose 1%, para novo diagnóstico dos clones. As colônias recombinantes foram armazenadas em 40% de glicerol a -80°C, para caracterizações adicionais.

#### 4.1. Caracterização das sondas de RFLP

RAFALSKI e TIMNEY (1993) construíram um mapa genético de ligação, usando sondas de RFLP com baixo número de cópias obtidas de uma biblioteca genômica de *Glycine max* cv. Bónus. Quando os progenitores usados no mapa (*G. max* X *G. soja*) e a população derivada desse cruzamento foram analisados com essas sondas, observou-se que 70% das bandas eram monomórficas. Tal percentagem é bastante elevada, principalmente se for considerado que a população de mapeamento era derivada de um cruzamento interespecífico. Dentre as sondas que revelavam polimorfismos, algumas mapeavam em regiões altamente polimórficas ou hiper-mórficas. Das 2.000 sondas analisadas, somente 0,5% foram extremamente polimórficas, sendo cinco delas (A1-10, A2-08, A45-10, A53-09 e A75-10) selecionadas para serem caracterizadas molecularmente neste trabalho.

Com a finalidade de verificar se essas sondas eram também polimórficas em outras variedades e acessos de soja, elas foram hibridizadas com o DNA genômico de seis genótipos de soja (Figura 1). O padrão de hibridização das cinco sondas foi muito semelhante ao observado anteriormente (RAFALSKI e TIMNEY, 1993), ou seja, apenas duas variedades e 6 acessos, as sondas foram altamente polimórficas ou hiper-mórficas.

sondas de RFLP foram selecionadas. Uma vez que essas sondas estavam clonadas no sítio de *Pst*I no vetor pBS (Stratagene), as amostras de DNA de cada clone foram digeridas com a enzima *Pst*I, e, após a análise eletroforética dos fragmentos, se determinou o tamanho de cada inserto, usando-se o programa *DNA* (Stratagene). A sonda A45-10 tinha um tamanho de aproximadamente 1,0 kb; A53-09 e 1,0 kb; A1-10 e 0,8 kb; A75-10 e 2,6 kb, e A2-08, 2,4 kb (dados não mostrados).

## 4. RESULTADOS E DISCUSSÃO

### 4.1. Caracterização das sondas de RFLP

RAFALSKI e TINGEY (1993) construíram um mapa genético de ligação, usando sondas de RFLP com baixo número de cópias obtidas de uma biblioteca genômica do cultivar de soja Bônus. Quando os progenitores usados no mapa (*G. max* X *G. soja*) e a população derivada desse cruzamento foram analisados com essas sondas, observou-se que 75% das bandas eram monomórficas. Tal percentagem é bastante elevada, principalmente se for considerado que a população de mapeamento era derivada de um cruzamento interespecífico. Dentre as sondas que revelavam polimorfismos, algumas mapeavam em regiões altamente polimórficas ou hipervariáveis. Das 2.000 sondas analisadas, somente 0,5% foram extremamente polimórficas, sendo cinco delas (A1-10, A2-08, A45-10, A53-09 e A75-10) selecionadas para ser caracterizadas molecularmente neste trabalho.

Com a finalidade de verificar se essas sondas eram também polimórficas em outras variedades e acessos de soja, elas foram hibridizadas com o DNA genômico de seis genótipos de soja (Figura 2). O padrão de hibridização das cinco sondas foi muito semelhante ao observado anteriormente (RAFALSKI e TINGEY, 1993), ou seja, nessas outras variedades e acessos, as sondas foram altamente polimórficas ou hipervariáveis.

A fim de determinar a base molecular desse padrão hipervariável, as sondas de RFLP foram seqüenciadas. Uma vez que essas sondas estavam clonadas no sítio de *Pst* I no plasmídeo pBS (Stratagene), as amostras de DNA de cada clone foram digeridas com a enzima *Pst* I, e, após a análise eletroforética dos fragmentos, foi determinado o tamanho de cada inserto, usando-se o programa Onedscan (Stratagene). A sonda A45-10 tinha um tamanho de aproximadamente 2,0 kb; A53-09 e 1,0 kb; A1-10 e 0,6 kb; A75-10 e 2,6 kb; e A2-08, 2,4 kb (dados não mostrados).

Os clones foram inicialmente seqüenciados em ambos os sentidos, utilizando-se os "primers" T3 e T7, que estão ancorados no vetor pBS. Uma vez que a maioria dos clones excedia a 1 Kb, o seqüenciamento foi feito em etapas. Com base nas seqüências obtidas em cada etapa do seqüenciamento, novos "primers" eram desenhados, sintetizados e utilizados na etapa seguinte (Tabela 2). Dessa forma, os cinco clones foram seqüenciados completamente (Figuras 3, 4, 5, 6 e 7).

Os dois clones de RFLP, A1-10 e A2-08 não apresentaram similaridade com seqüências depositadas no "GenBank". Entretanto, esses dois clones são ricos em A:T: A2-8 com 74% de A:T e A1-10 com 64% de A:T. O fato de seqüências ricas em A:T serem freqüentemente encontradas em regiões de DNA repetitivo pode explicar o padrão de hibridização altamente polimórfico encontrado nessas sondas. Nessas regiões ocorrem, freqüentemente, "crossing-over" desigual e erros de replicação pela DNA polimerase, fenômeno conhecido como "slippage" da polimerase, provocando alta taxa de mutação (KORNBERG e BAKER, 1991).

A comparação das seqüências dos outros três clones de RFLP, A45-10, A53-09 e A75-10 (Figuras 4, 5 e 7), com outras seqüências presentes no "GenBank", revelou que esses clones apresentam alta similaridade com outras proteínas codificadas por genes de resistência a doenças, como pode ser visto na Tabela 3. Os genes de resistência freqüentemente pertencem a famílias multigênicas e estão arranjados em blocos gênicos. Além disso, são encontradas, na maioria dos genes de resistência, regiões com motivos ricos em leucina (LRRs – "leucine rich repeats"), que são bastante conservados entre diferentes genes. Tais motivos repetidos nas proteínas de resistência variam, em número, em diferentes proteínas, havendo, também,

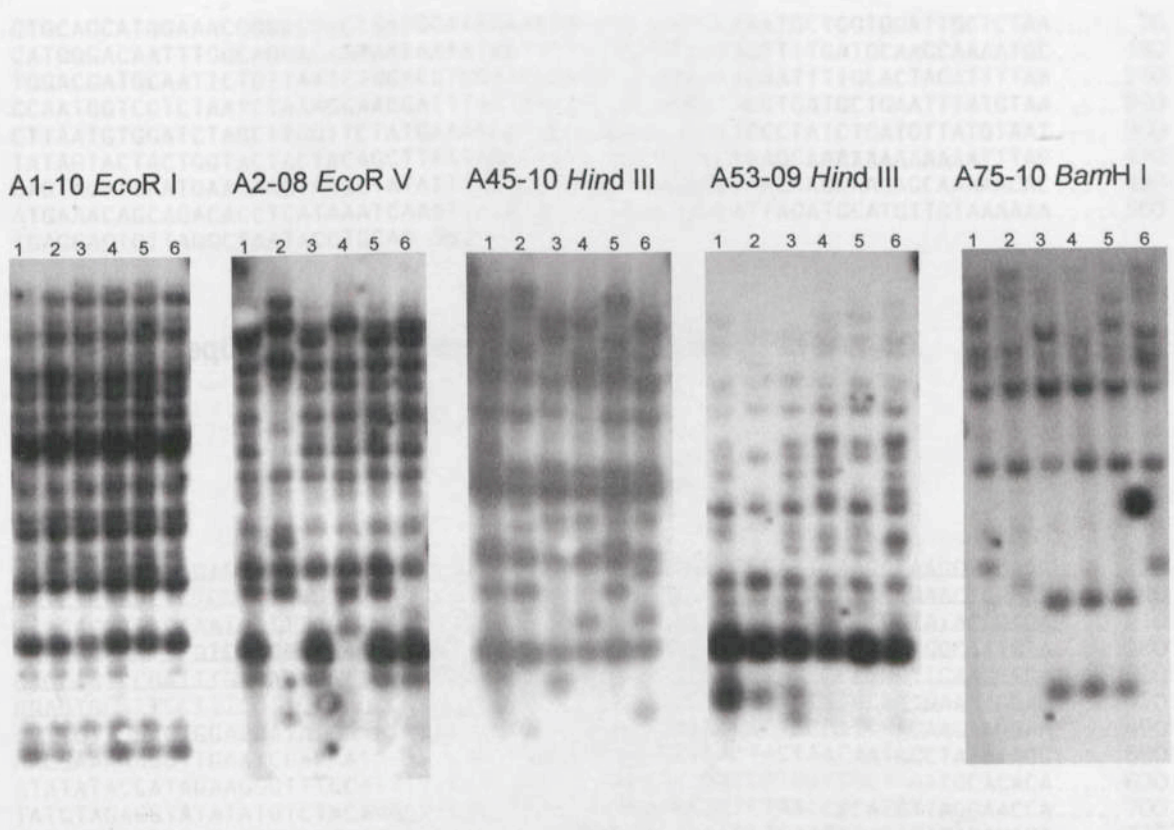


Figura 2 - "Southern blot" de DNA genômico de soja. O DNA genômico de seis genótipos de soja foi digerido com as enzimas de restrição *EcoR* I, *EcoR* V, *Hind* III e *BamH* I, separado por eletroforese em gel de agarose 0,7% e transferido para uma membrana de náilon. As membranas foram hibridizadas com as sondas A1-10, A2-08, A45-10, A53-09 e A75-10, em condições de alto rigor. Os números na parte superior equivalem às amostras de DNA genômico usadas (1 - Bônus, 2 - PI 81762, 3 - PI416937, 4 - N85-2176, 5 - PI 153293 e 6 - PI230970). Cada painel é identificado pela sonda utilizada na hibridização e pela enzima de restrição com a qual os DNAs genômicos foram digeridos.

CLONE A1-10

CTGCAGCATGGAAACGGGACTGCTGATGGAAAGAAGTCCCCAAGATGGAAATGCTGGTGGATTGCTCTAA..... 70  
 CATGGGACAATTTGGCAGCACAGAAATAAAATAGTTTTCCAAAATGCAAGTTTTGATGCAAGCAAAATGC..... 140  
 TGGACGATGCAATTCTCTTAATCTGGACGTGGATCAAATCCATGGAGAAGGATTTGCACTACATTTTAA..... 210  
 CCAATGGTCTCTAATCTAAAGGAAGGATTTAGTAAGTGGAGGGGATGGGTGATGCTGAATTTATGTAA..... 280  
 CTTAATGTGGATCTAGCTTGGTTCTATGAAAACTGGAACCAGATGGATCCCTATCTCATGTTATGTAA..... 350  
 TATAGTACTACTGGTACTACTACAGCTTAATAAAATATATTTTTGCTGAGCAAAAAAAAAAATTTAG..... 420  
 CAGTACATCAATGAACAAGTAAGCTATATTGAGAAGTCCACTTACATGCAACAAGCAAGAGCAAAAACAC..... 490  
 ATGAAACAGCAGACACCTCATAAATCAAATTAATGCAATTTATAAACATTAGATGCATGTTGTAAAAA..... 560  
 TGAGCACTGTTAGGCTAATACCTGCAG 587

Figura 3 - Seqüência de nucleotídios do clone de RFLP A1-10.

CLONE A53-09

CTGCAGGTTTIGAGAGGTTTCTAGTTGATGAGGGATTAATCCCATGAACCCACTTAAGCTAAGGTCCAAG..... 70  
 TATCTCAGACTCTCCAATGAGCCTAGGAAGCTTGGTATTGGAGTAAGAACAAAATAAATTTGAACCTCAAGT..... 140  
 CCAAACGATTCAAATATTTAGTTCAAGCAAAGAAGGACTAATCTCTCCACTCAACTCCCTATAGGGAGA..... 210  
 GCCCGCAGGTGTGTCAAGATTGATTTCCATGACTTTACCTGTGTGTGTCAGTGAACCCCTGGCCATGTA..... 280  
 CAGCAATCCGATTTGTGACACCACGATGAAAGCCTGTTTGAAGGGTCTGCTAGTCCATGCTTGAAGCTGA..... 350  
 GGAGTGCAATTCCTTTCTCTCGCTGCAAGTCATATTGAGCCTTGTCTGCTTGTCTTGCCTGCAAGTGGAG..... 420  
 AGTTGTTGCAAGTGGACAATATCAAGAGAAGAACAATGAGTTGCATATAGCACATGCCATGAAGAAGAAGGAA..... 490  
 AACGAAAAGGTTGAATCCACAATTTCTGGTTCACGGTGGTATTTACTACTAACAATGCCTATATAGG..... 560  
 ATATATACCATAGAAGGTTTGCATTTAAAATTTTATTCCACTAATGGTGGTGGTTGCAAAATGCACACA..... 630  
 TATGTAGAGGTATATATGTCTACAAGGAGACATTCAGGATTAAGTGTCTAACCACATGTTAGGAACCA..... 700  
 AGCTTGAGAACGTGAACCACCTAGTTTCGAGGGCTAGGAACCTCCATAGAAGAAGAGGAAGAGAGAGACT..... 770  
 AAGGGAGAAGTGGCTGTTATTTTCTGATATCAAAGCGCTTAATTCATCTACATATATAGCTGTTATG..... 840  
 CAGGGAACAAGGCCCAAGCAAAAACAGAAAAGATAAAATGGAATATCCTAACAACCATATTCCTTAT..... 910  
 TCCACTACCTACATATCTTCTAGAAAAGCAGATGATATCTTCTTACAGATTATACCAGCTGCATGGGTC..... 980  
 TCATGCTTTTTCTTCTGCAACCCCTTACTCCCTGCAG 1019

Figura 4 - Seqüência de nucleotídios do clone de RFLP A53-09. As seqüências sublinhadas correspondem a seqüências de leitura aberta (ORF), que são similares a genes de resistência a doenças. As seqüências marcadas de vermelho correspondem aos sítios de pareamento dos "primers" usados nas reações de PCR, seqüenciamento e RT-PCR.

CLONE A45-10

CTGCAGGTATTAATCTTGGAGCTAATTCTTGACTGGTGATGTGCCGTAACTCTTGGAACTCTCAA... 70  
 ATTTAGTGACGTTAGACCTTTTCATCTAATTTGTTAGAAGGATCTATAAAAAGAGTCAAATTTTGTAAAAC... 140  
 TTTTACATTAAGGAACCTACGTTTGTCTTGGACAAACTTGTTCCTCAGTGTCAACTCTGGATGGGCTCCT... 210  
 CCTTTTCAGCTTGAATATGTTTTACTGAGCTCCTTTGGAAATAGGTCTTAAGTTTCCAGAATGGCTAAAA... 280  
 GGCAAAGTCTGTGAAGGTTTTGACAATGTCTAAGGCAGGTATTGCAGACTTGGTTCCAAGTTGGTTTTG... 350  
 GATTGGACTTTGCAAATTTGAATTCCTGGATCTGTCAAACAATCTGTTAAGGGGAGACCTATCTAACATC... 420  
 TTTCTAAATTCAGCGTGATAAAATTTGAGTTCTAATTTGTTCAAGGGTAGATTGCCAAGTGTGCTGCAA... 490  
ATGTTGAAGTGCTGAATGTTGCCAATAACTCAATTTCCGGAACAATTTCTCCCTTTTTATGTGGAAATCC... 560  
 AAATGCTACTAATAAGTTAAGTGTGCTTGTATTTTTCAAATAATGTCTTGTCTGGTGATCTTGGTCACTGT... 630  
 TGGGTGCATTGGCAAGCTTTGGTGCATGTGAACCTGGGCAGTAACAATTTGCTCTGGTGAATTTCCAAACT... 700  
 CCATGGGGTATCTGTCTCAACTTGGAGCTTTTGCCTTGGTGTAGACGACAACCCGCTTCTCAGGATATAITCCTC... 770  
 AACACTGCAAAATTTGCTCTACAATGAAATTCATAGACATGGGAAAATAACCAACTTTCTGACACAATACCA... 840  
 GATTGGATGTGGGAAATGCAATATCTAATGGTTCCTCGTCTAAGATCCAACAATTTCAATGGCAGCATTG... 910  
 CTCAAAAGATGTGTCAACTTTCTTCCCTTATAGTGTGGATCTTGGCAATAACAGCCTGTGAGGATCCAT... 980  
 TCCAAATTTGTTGGATGACATGAAAACAATGGCGGGTGAAGATGACTTCTTTGCAAACCTTTCAAGTAT... 1050  
 TCATATGGCTCTGACTTTCAGTTATAACCACTACAAGGAAACTCTTGTCTTAGTTCCTCAAAAAAGATGAGT... 1120  
 TAGAGTACAGAGACAATCTGATTTGGTGAGAATGATGATCTTTCAAGTAATAAGCTGTCTGGAGCAAT... 1190  
 ICCATCTGAAATTTCCAAGCTATTTGCTTTGCGGTTTTTGAACCTGTCTAGAAATCATCTGTCTGGAGAG... 1260  
 ATACCAAAATGACATGGGAAAAATGAAATTTGTTAGAATCCCTTGATCTCTCGCTAAACAACATTTTCAGGTC... 1330  
 AAATCCCTCAAAGCTTATCTGATTTGTCTTCCCTCAGTTTCCCTGAATCTATCATACCACAACCTTATCAGG... 1400  
 CAGAATTTCCACGAGCACCCAACCTCAGAGCTTTGATGAACTTAGCTACACTGGAAATCTTGAGCTTTGT... 1470  
 GGTCTCTCTGTAACAAAAAATTCACAAAACAAGGAATGGTGTGAGGGAGAGTGTCTCTGTTGGACACGGTG... 1540  
 ATGGTAATTTCTTTGGAACATCAGAGTTTTATAICGGTATGGGAGTTGGATTTGCAGCAGGATTTTGGGG... 1610  
 TTTTTCAGTGTGTTTTCTTCAACAGAACTTGGAGGCTTGCCTATTTTTCAATTATCTTGACCACCTTGAGA... 1680  
 GATCTGATTTATGTGATGATAGTTTTGAAGGTAAGAAGGTACTTGGGAAATATGAGTTTACCTGTITAA... 1750  
 CTGTTAAGAAGTGAGACATTACAACGTGTTAGTCTAGCCATTAATAATGTGAAAAGGTTTACCATACTAT... 1820  
 GTACCTGTTTCACTGTGGTGTCTGTCTGTTTTGCTTTCATTGGTATTAATGTCTTAATGTTAAGAGA... 1890  
ACTAGACGAGCTTGAGACTTGAATATCAAATTCGTTTTGTATTTAATGACGCATTATTTGCTTCTACTAT... 1960  
 CTGTGCACTAACAAATGATTAGCATTAAAAATAACTGAAATGACACACTCGGTGATTTGGATTTAAGAAGAT... 2030  
 GTTTACACAACATGTAACATAAGAAATTTGCCCTGAATGAATTTAATTAGTGTAGAAAAGTTGGAGCTTGCAT... 2100  
AGGATTCTCAG 2112

Figura 5 - Seqüência de nucleotídios do clone de RFLP A45-10. As seqüências sublinhadas correspondem a uma ORF, que é similar à seqüência de genes de resistência a doenças. As seqüências marcadas de vermelho correspondem aos sítios de pareamento dos "primers" usados nas reações de PCR, seqüenciamento e RT-PCR.

CLONE A2-08

CTGCAGCAAGAGCCAGTGCCGTGT**ACCCTCTTGTGTACGTT**CCAACAAAAGCCGTTTGGGCACCTTTGTT.....70  
TACTAACTTTCCACCATGTCTTCTGTGCTGTGAGAACTGCAACATGAACAACCGTCTTGCCAGTTTT.....140  
GAAGTTTCATACATTGCCTTACGATTGTTGTTCAAAAAGTGCGTGCAGTGTCCCAATCACCTCTTTGTA....210  
CACATCTGTGTAGTAATGGAAGACCTGCCATTTTGTGATTAA**CCATAGAGCATGTGTGATGAGA**AATTTT....280  
TTTTGCTTAATTAATTTTTTTCATATATATGGAATATAAATCGTTTTTAAAAATAACTATATTTTATATAT....350  
GAAAAACTTAATTAACCTTTTTTATGACCTAACAATTAGAGATAGTAATAATCAAATTGAAAAAATTT....420  
AAATTTAAATAAGTATTTAGACTATACAAATATACTAGTTTTAATTTTTCTAAAAAAAATAGAACTTATA....490  
AGAGATGCACATTTAATATAAGATATTTAATAAAACT**AACAACTAACGTCCGTGAAAA**AAAAAACTGT....560  
TAATATAAAATTTTTAATAGGATCAAACAATAAAAAAATCAATTTTAATTGATATAATTTAAGATAAT....630  
TAAATATAATATATTACAAAAATTAACCTTATCATGCAAAAATTTAAATACATGATAGTGTAAAAA....700  
TATTTTACTTATTAGTACATATATTACATATTAATAATATATCTAGCATGAGAAAATACTCAATTTATTA....770  
AGCAAAATATATTAATTGTTATAAACTTCTGATAATTATCTTTAATAACTATGAATAAAAAAAGATATC....840  
TATTTTTTACTTAAAAATCACAATATTTTTAAAAAAGAAAAAA**CCCATTTTAAAACTCAATGGC**TAAA....910  
TGGTATAGATACAATTAATTACATTAATATAAAACAAGATAAAAAAAGAAAAGGAAAAGGGGATTAGAA....980  
TTGGATTGGATTTACCTTGTCTATATCAGCTTTAGCTGCAAGTCTAGACTTGTACAAAGATGCCAAGGT....1050  
ATTCCATGCCCTCTGGCTGTGGTAGTTTCTTAACTGAGAGAAAGTTTCAGATCCACAAGACATCTGA....1120  
ATAATATGCAAAAATAACTTTTAAATGTATCCATTCAAATAAGCTTTATCCGATAAATAACAATCTACCCT....1190  
TTACATTTAAATCACCTAAATCAAAATGTTAAAAAATCAATTCCTACCATGCATGTAGCTTATCTTTTAA....1260  
CTTAAAACTCTATATTAACCTTCATATTTGAAAAAACTAAAAACTTTTCTATACTAAACACTCCTTAA....1330  
AAGTCAAATCTAAGATTA**TTGCTAATTGTTTGACAATAAAAAATCT**TAGTTTACTTAAAAGTTAACTATA....1400  
TTTTCCCTTTTTAAATAAATATTAGAGAATTTAATTAATAATTTCTGAAAATGTAACATGTAATCATAAA....1470  
TTATCATATAATTAACCTGAATTATATAGTAACTAATTAATTAATGTCAATTAAGAACTGTAAAATAGC....1540  
ATTTAAACCGTGTATGTTTAAAAACAAATACACTGTTAGACATGCACTAAAAAACATTTTTAAAAAATAATT....1610  
AATTA**TTTTATCATTTCCTTTTCCTTTT**AATAATATAACCAAAAAATAATTAGTATATTTTTTTGAC....1680  
TGGGTATAAAATAAATAATATTTAAACCTATTAAAGTTGTATATGTATATCCTATTTAGCAAATATTAA....1750  
AGAATTTAATTGGAATAAATTTGCACCATTATTAATTATGAATTTTCATATAAGCAAAATTTGTTTTTT....1820  
AATTTTCATATATTAAGCAAAATTTGTTTTTTAATTTCTATTTATTGTGATTTTGCATTCCATGAAAGCA....1890  
TACTTTACCTTTTTTTTATCAAAACAAAAGTAATCATTTGTCAATTACTTCCCATATGTAATTTAGTT....1960  
AACTATTCGACACGAGATCTTATAGAATTAATGTTTTCTAGT**CCAAGAGTTAGTTCTTTATTCAGGAAGA**....2030  
TATATTAAGTACTATCTATCTTCTTGTCTTCTATTTCTTTCTTTTCCCTCTTGGTGATTACTTCAATTCA....2100  
CTACTTTTCATATATGAACAACCTGGTCTTGCATCTTGGGTCTTTTATTGGGCTGATTGGCTCTGCATGAG....2170  
TAGGATATCCCTGATGGTGACTGGTTTTGTCCATCTTGTGCTTGCGGGATTTGCAGACAAACCAAAATAG....2240  
**AAGGGACTGAGGATGGTGA**TTTACTTGTCTGCATTCAATGTCAACATAAATGTAAGTTGATTTGCATATT....2310  
TGCTTCAATATTTTCATCCATTCTTTTTGTGAGTTGTTTTGCAGATGACTGCAG 2365

Figura 6 - Sequência de nucleotídeos do clone de RFLP A2-08. As seqüências marcadas de vermelho correspondem aos sítios de pareamento dos "primers" usados nas reações de PCR e seqüenciamento.

Clone	Gene	Sequência	Localização	Vetor
A75-10	Prd	ATCAACATTGTA...	Prd	26-10
	Prd	ATCAACATTGTA...	Prd	26-10
	Prd	ATCAACATTGTA...	Prd	26-10
	Prd	ATCAACATTGTA...	Prd	26-10
	Prd	ATCAACATTGTA...	Prd	26-10

CLONE A75-10

CTGCAGGTA AATGTTTCAGCTGTTT ATTATTAGTTCTAAAATTAGGGAATAGATTAATGGAGTTTCTGTA..... 70  
 ATCAACATTGTA AACTTAAAACCCTGTACTTAAATTACTTGAAATAGTAACAGGTTAATAGCCCTTAAGC..... 140  
 TGTCTGATTCTTTTTCTTGTTATTTCCATAAA TGGTTAGAGGAGGCATGACTTCTGTCCCTGTGTTCA..... 210  
 TCAGCTTGAATGTCAAAGGAACAGTATGGCCTTGGAAAGATAAACAAGGCAAAGGTTTTTTTAGTTTGAA..... 280  
 CCAGAATGTTTGATTTTGTGAAATTTAAAAGTTTAAATTTCTTTTTAAAAAATATTTTGATCATACAAGGA..... 350  
 CTA AATTACAAGATATAAAAAATTATCATTAAATTTAAAAATAACAATATAATAATACAACCTTAATTCCT..... 420  
 ATTAGTTGAAAAAATAACAGTATCAACTTCTAAATTAACATGGATATGGTGGAAACTTTTTTAATTTT..... 490  
 TAACCCATAAAAAAATGTATTTAAAATAGCGTGTGAGATAATATCTTCCAACATTTCTGGCTAATCCAA..... 560  
 TACTTGCTTATGTTTATAACTAATTTTTGACAGAAAAAATAAATCATTAAAGTCAATACCTACCGCTCGT..... 630  
 AAAGAAAAGTTGCAACAAAAGGCTGAAAAGTATTTAAATTTGCTTTTGATGAATTATAAATGCCTTATAATT..... 700  
 TGTGTATGATGGATACATATAAGCTTAAATTAGAAGATAATCCAAATTAAGACCTTGCAAAATAATATAA..... 770  
 AATATTTATACATTTTAAAAAATAAAAAATAGA AACTCAATAAATAAAAAAGTAAGTCTACTACTTTTATA..... 840  
 ATTTTTTTGATAAAATTTAGTTAGTAAAAATGTGTTTATAATTTTAGTTAA TCATGCTTCTTCTCCAGT..... 910  
 GTCTTCTCTCTTGTCATCACTTATCAGACATCTACCAAGCTTCTCACTACGGTGAACCTCAAGTCAAGCTC..... 980  
 AGACATGCTGTCCCGACAAGAAGTTGAGATGCAATGAAGAGGCAGATGAACACAAGAGGGTGAGCGAAGG..... 1050  
 GTAAGACGGAGAGGGAAAAAGGGTAAGACGGAGATATTTATTTTTCTGTGGTGCATCAAAACAAAAA..... 1120  
 GACAAGCAATTGCCAATACAATTAACATAACAACACATGACCTTATTTTATTGTCACCAACACACCCAAAA..... 1190  
 AGCTAATGACTTGTGGGTGTGTGAGTCTATTACTAATGTTTTGTTATAAGTCAAAACATATAAGATA..... 1260  
 AAAAAATTAATTTCCAAAAAATTCATCCGCTTAAATGTATGGTATGTCTATTAATTTGAGAAAATGCTATAA..... 1330  
 CCACCTAAAAGAATACCAGAAAAATGGCTGTCCCATGCAACTA AACTATTAATTAATTTCTACTTTTTTA..... 1400  
 AAATGTTTTTCAGATTAGAGAATAAATAAATTAATAAAGATATATATAAAAAATAACTAATGCATCTATAA..... 1470  
 ATTAATTTATGGATCCTTTTGTTCATATAGTTGGTTGGTCCAATAAATAGTTAGTTGGTGCAATAAAGGAT..... 1540  
 ACACCCGTGTTCACTTTACTCCTGCACCTTCATATTTGAACGAGTCAAAGTTTCATGGATCCAATTCAA..... 1610  
 AGCCATCCCAAATCATACTAAGTTTCCAGTTTTCTTCTAGGAAGTTTCGTTGGAAAAA TGAACCATAAT..... 1680  
 TCCAGCAAAATAAGATCGATTAATAATGATTCAAAAAATAAATACGGATTACACATACTCTTTAGGATA..... 1750  
 AAAAAATCTAACAAATTTTTTGAAGTCTTGAATAAAACTTATCAGCTTACTTTTTAGTTAACATTCACC..... 1820  
 AGTATTTTTATTTATCGTAACAATTAACTTTTTGGGGATGTACCATTATCAATAGTCAATTATCTTAA..... 1890  
 TATGGTGACAATTATAAAAATTGACTGTATACTTTATTAATTTGTAATGATTATCAAAATAAAAAAGTAT..... 1960  
 TTTTTGTAAGAGTTAAATAGTGCGGTGTAGAACGTATGCCACGGGTATGCACAGCCAAAGGTTTTAAC..... 2030  
 TTCTAAGAAAGAGAGACGAAACAGAGAGAGATTTCCATAAAGTTGATATGTCCATGAAAATCTAATAAT..... 2100  
 ATATATGTTTGGTCTGAGTGTAATTAACACTTATCAGTTTCATACCTACTCAAGTTTGATTTTGGAGAC..... 2170  
 CCAACAAGCAGCCATCCA AATTCATACTAAGATTTTATATCATGGCTTGGGGATCACGTTCTCTTCTC..... 2240  
 TTCTCCAATTACGATGTGTTCTCAGCTTCAGGGGTGAAGACACACGTAGTGTCTTTACTGGCCATCTC..... 2310  
 TACAACACTCTTCAGAGCAAGGGAATCCACACTTTTATTGATGATGAGAAGCTTCAAGAGGGGGAGCAAA..... 2380  
 TAAACACCAGCACTTATGAAGGCAATTGAAGATTCAGAGTTGCCATCACTGTGCTCTCTGAACACTATGC..... 2450  
 TTCTTCTCGTTTTGTTTAGATGAACITGCAACCATCTTCACTGTGACCGAGGAAACCGGTTGTTGGTT..... 2520  
 ATACCGTCTTTTATAAGGTGGATCTTCTGATGTGAGACACCAGAAAAGGTAGTTATGGAGAAGCATTTGG..... 2590  
 CAAAGCTTGAGAGGAGGTTCCAGCATGACCCGGAGAAGTTGCA GAACTGGAAGATGGCTCTGCAG 2655

Figura 7 - Sequência de nucleotídeos do clone de RFLP A75-10. As seqüências sublinhadas correspondem à ORF, que é similar a seqüências de genes de resistência a doenças. As seqüências marcadas de vermelho correspondem aos sítios de pareamento dos "primers" usados nas reações de PCR, seqüenciamento e RT-PCR.

Tabela 3 - Análise de similaridade entre as seqüências deduzidas de aminoácidos das sondas A45-10, A53-09 e A75-10 e as seqüências depositadas no "GenBank"

Clone	Clone do "GenBank"	Acesso no "GenBank"	Espécie	Valor P
A45-10	Proteína AWJL218 Cf-2.2	X81369	<i>Triticum aestivum</i>	2e <sup>-70</sup>
		U42445	<i>Lycopersicon pimpinellifolium</i>	2e <sup>-64</sup>
	Cf-2.1	U42444	<i>Lycopersicon pimpinellifolium</i>	2e <sup>-64</sup>
		U15936	<i>Lycopersicon pimpinellifolium</i>	2e <sup>-59</sup>
A53-09	Proteína inibidora de poligalacturonase Proteína LRR	AF020785	<i>Prunus armeniaca</i>	1e <sup>-12</sup>
		X95269	<i>Lycopersicon esculentum</i>	1e <sup>-12</sup>
	Proteína precursora de Cf-9	U15936	<i>Lycopersicon pimpinellifolium</i>	2e <sup>-12</sup>
	Proteína inibidora de poligalacturonase	L26529	<i>Lycopersicon esculentum</i>	9e <sup>-12</sup>
A75-10	Proteína N (resistência ao TMV)	U15605	<i>Nicotiana glutinosa</i>	7e <sup>-35</sup>
	Proteína L6	U27081	<i>Linum usitatissimum</i>	1e <sup>-29</sup>
	Proteína M (resistência à ferrugem)	U73916	<i>Linum usitatissimum</i>	1e <sup>-26</sup>
	Proteína Rpp5 (resistência ao míldio)	U97106	<i>Arabidopsis thaliana</i>	8e <sup>-23</sup>

polimorfismos dentro dessas regiões, em comparação com outros alelos (JONES e JONES, 1997). Como pode ser visto nas Figuras 8 e 9, esses clones (A45-10, A53-09 e A75-10) possuem esses motivos arranjados em suas seqüências de modo repetido.

Tem sido proposto que os domínios LRRs estão associados com interações proteína:proteína e que tais interações seriam responsáveis pelo reconhecimento do patógeno pela célula vegetal, sendo esta a região envolvida na especificidade das interações gene-a-gene (BAKER et al., 1997; ELLIS et al., 1999). Vários genes de proteínas de resistência clonados até hoje possuem seqüências que codificam LRRs (JONES e JONES, 1997), o que vem ao encontro dessa hipótese. Os genes de resistência evoluem mais rapidamente do que outros genes, devido à forte pressão de seleção exercida sobre eles. Alguns pesquisadores acreditam que as regiões LRRs seriam uma das regiões da molécula que teriam evolução mais acelerada (MEYERS et al., 1998a).

Nesse caso, o polimorfismo e o padrão hipervariável observados podem ter ocorrido em virtude da hibridização destas sondas (A45-10, A75-10 e A53-09) com outros membros da sua família gênica ou, ainda, com outras regiões LRRs presentes em outros genes.

#### **4.2. Caracterização dos polimorfismos por PCR**

Devido ao fato de toda a seqüência do clone A45-10 ser uma ORF e ter apresentado padrão polimórfico entre os genótipos de soja analisados (Figura 2), foram desenhados os "primers" 4510T7F e 4510T3F (Tabela 2), com o objetivo de verificar se eles amplificavam regiões polimórficas que pudessem ser usadas para identificação molecular de variedades.

DNA genômico de três variedades (Bônus, Hardin e Noir), de quatro acessos de *G. max* (PI 230970, PI 416977, N 85-2176 e PI 51293) e de dois acessos de *G. soja* (PI 440913 e PI 81.762) foi amplificado com o par de "primers" 4510T7F e 4510T3F, produzindo de dois a quatro fragmentos por reação, que variavam entre 200 e 2.000 pb. Os produtos de amplificação na faixa de 2.000 a 1.500 pb eram bastante polimórficos, sendo o perfil de

Clone A45-10 (ORF +1)

LQVLNLGANSLTGDPVPTLGTLSNIVTLDLSSNLEGSIKESNEVKLFTLKEFLRLSWTNLFLSVNSGWAPPFQLEYVLLSSFGIGPKFPEWVKRQSSVKVLTMSKAGIADLVPSWFIWTLQIEFLDLSSNLLRGDSLNIFLNSSVINLSSNLFKGRPSVSANVEVLNVANNISGTISPF LCGNPATNKL SVLDFSNVLSGDLGHCWVHWOALVHVNLGSNNLSGFI PMSMGYLSQLESLLDDNRFSGYIPSTLQNCSTMKFIDMGNNQLSDTIPDWMMWEMQYLMVLRNSNNFNGSIAOKMCOLESSLIVLDLGNNSLSGSIPNCLDDMKTAGEDDDFFANPSSYSYGSDFSYNHYKETLVI VPKKDFLEYRDNILVRMIDLSSNKLSGAIPSEISKLFALRFLNLSRNHLSGFI PNDMGKMLLESLDLSLNNISGQIPOSLSLSFLSFLNLSYHNLSGRIPTSTQLQSFDELSYTGNPFLCGPPVTKNCTNKEWLRRESASVGHGDGNFFGTSEFYIGMGVGF AAGFWGFC SVVFNRTWRLAYFHLYDHLRDLIYVMIVLKVRRLLGKL

A53-09 (ORF -3)

MAVLYATHVLLLILSTATLHESASKAARLNMTCSFKERNALLSEKHGLADPSNRLSSWSDKSDCCTWPGVHCNNTGKVMFINLDTAGSPYRELSGFISPSLLEFLKYLNRDLSSNYFVLTPIPSFLGSLFSLRYLDLSLSEGMGLIPHQLGNLSNLQ

A75-10 (ORF +1)

MAWGSRSSSSSSNYDVELSFRGEDTRSFTGHLNYTLQSKGIHTFIDDEKLRQGEQITPALMKAIEDSRVAITVLSHFYASSSEFLDELATILHCDQRKRLVIPVFYKVDPSDVRHQKGSYGEALAKLERRFQHDPEKLNQWKMALQ

Figura 8 - Sequências de leitura aberta (ORF) encontradas nos clones A45-10, A53-09 e A75-10. Os resíduos de leucina estão destacados de vermelho.

LQVLNL---GANSLTGD-VPVT--LGTLSN  
LVTLDL---SSNLEGS- IKESNEVKLFT  
LKEFLRL---SWTNLFLS-VNSGWAPPFQ  
LEYVLL---SSFGIGPK-FPEW--LKRQSS  
VKVLTM---SKAGIADL-VPSWFIWTLQ  
IEFLDL---SNNLLRGD-LSN---IFLN  
SSVINL---SSNLFKGR-LPS---VSN  
VEVLNVA---NNSISGT-ISPFL--LCGNPNATNK  
LSVLDF---SNNVLSGD-LGHCW-VHWQA  
LVHVNLG--SNN-LSGE-IPNS--MGYLSQ  
LESLLL---DDNRFSGY-IPST--LQNCST  
MKFIDM---GNNQLSDT-IPDW--MWEMQY  
LMVLRRL--SNN-FNGS-IAQK--MCQLSS  
LIVLDL---GNNSLSGS-IPNC--LDDMKTAGEDDDFFANPSSYSYGSDFSYNHYKET  
LVLV-----PKK---DE-LEY---RDNLIL  
VRMIDL---SSNKLGA-IPSE--ISKLFA  
LRFLNL---SRNHLSGE-IPND--MGKMKL  
LESLDLSL--NN-ISGQ-IPQS--LSDL-SF  
LSFLNL---SYHNLSGR-IPTST---QLQSF  
-DELSYT--GNPLCGPPVTKNCTNKEWLR  
SASVGHGDGNFFGTSEFYIGMGVGF AAGFWGF  
CSVVFNRTWRLAYFHLYDHLRDLIYVMIVLKVRRLLGKL

Figura 9 - Sequência do clone A45-10 arranjado em um padrão de LRR. Os resíduos de leucina estão destacados de vermelho.

amplificação capaz de distinguir os diferentes genótipos de *G. max*, com exceção das variedades Hardin e Noir e do acesso PI 230970, que apresentaram o mesmo padrão de bandas (Figura 10). Portanto, esses “primers”, em conjunto com outros marcadores, podem ser usados como “fingerprinting” para distinguir cultivares de soja. No entanto, a presença do fragmento de 200 pb em todas as variedades e acessos não era esperada, já que a região delimitada pelos “primers” corresponde, no clone A45-10, a aproximadamente 2.000 pb. Esse dado indica que os “primers” desenhados poderiam estar pareando em mais de uma região do genoma.

Com o objetivo de obter maiores informações a respeito da natureza do polimorfismo e da presença da banda monomórfica de 200 pb, os diferentes produtos de amplificação, dos diferentes genótipos, foram clonados, seqüenciados e comparados. Os fragmentos de 200 pb foram totalmente seqüenciados, porém somente cerca de 400 nucleotídios em cada uma das extremidades dos produtos de PCR, de 1.500 a 2.000 pb, foram seqüenciados.

Como pode ser observado na Figura 11, o polimorfismo de tamanho entre os fragmentos de PCR de 1,5 a 2 kb deve-se a pequenas deleções e inserções; além disso, foram observadas mutações de ponto nas seqüências de nucleotídios.

Os fragmentos monomórficos de 200 pb aparentemente são do mesmo tamanho, mas no seqüenciamento foi observada pequena diferença, que pode ser atribuída a pequenas inserções, ou deleções, de um a quatro nucleotídios, principalmente no acesso N852176-10 (Figura 12). Este acesso foi o mais divergente entre todos os genótipos quando analisado pelo método de alinhamento de Jotun Hein, com o uso do programa Megalign - DNASTAR (Figura 13). Apesar de os fragmentos terem sido delimitados pelos “primers” derivados do clone A45-10, a percentagem de similaridade entre os fragmentos e a sonda de RFLP ficou em torno de 30%. Esses dados, juntamente com os resultados de hibridização (Figura 2), indicam que a sonda A45-10 pertence a uma família multigênica e os “primers” estariam pareando com mais de um membro dessa família.

### 4.3. Amplificação do DNA genômico de FT-Cristalina com “primers” específicos para os clones de RFLP

Pelo fato de as três sondas de RFLP A45-10, A53-09 e A75-10 apresentarem similaridade de seqüência com genes de resistência e serem altamente polimórficas entre as variedades analisadas, foi feita uma amplificação do DNA genômico de FT-Cristalina, com a finalidade de identificar seqüências homólogas às sondas de RFLP nesse cultivar, já que essa variedade é resistente a algumas doenças. A partir das seqüências dos três clones (Figuras 4, 5 e 7) foram construídos “primers” específicos para os clones A45-10 (4510T7B1 e 4510T3C1), A75-10 (7510C e 7510B) e A53-09 (5309D e 5309A), apresentados na Tabela 2. Tais pares de “primers” amplificam as pressupostas regiões codificadoras.

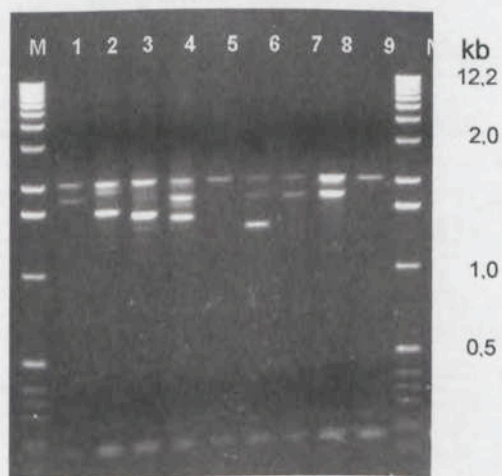


Figura 10 - Produtos da reação de amplificação do DNA genômico de diferentes genótipos de soja amplificado com os “primers” 4510T7F e 4510T3F. Os produtos de amplificação foram separados por eletroforese em gel de agarose 1%. Os números na parte superior se referem às amostras de DNA de soja usadas (1 - PI 230970, 2 - PI 416937, 3 - N85-2176, 4 - PI 51293, 5 - PI 440913, 6 - Bônus, 7 - Hardin, 8 - Noir e 9 - PI 81.762), e M refere-se ao marcador de peso molecular (“1 Kb DNA Ladder” - Life Technologies). Os números à direita correspondem ao tamanho das bandas em kb.

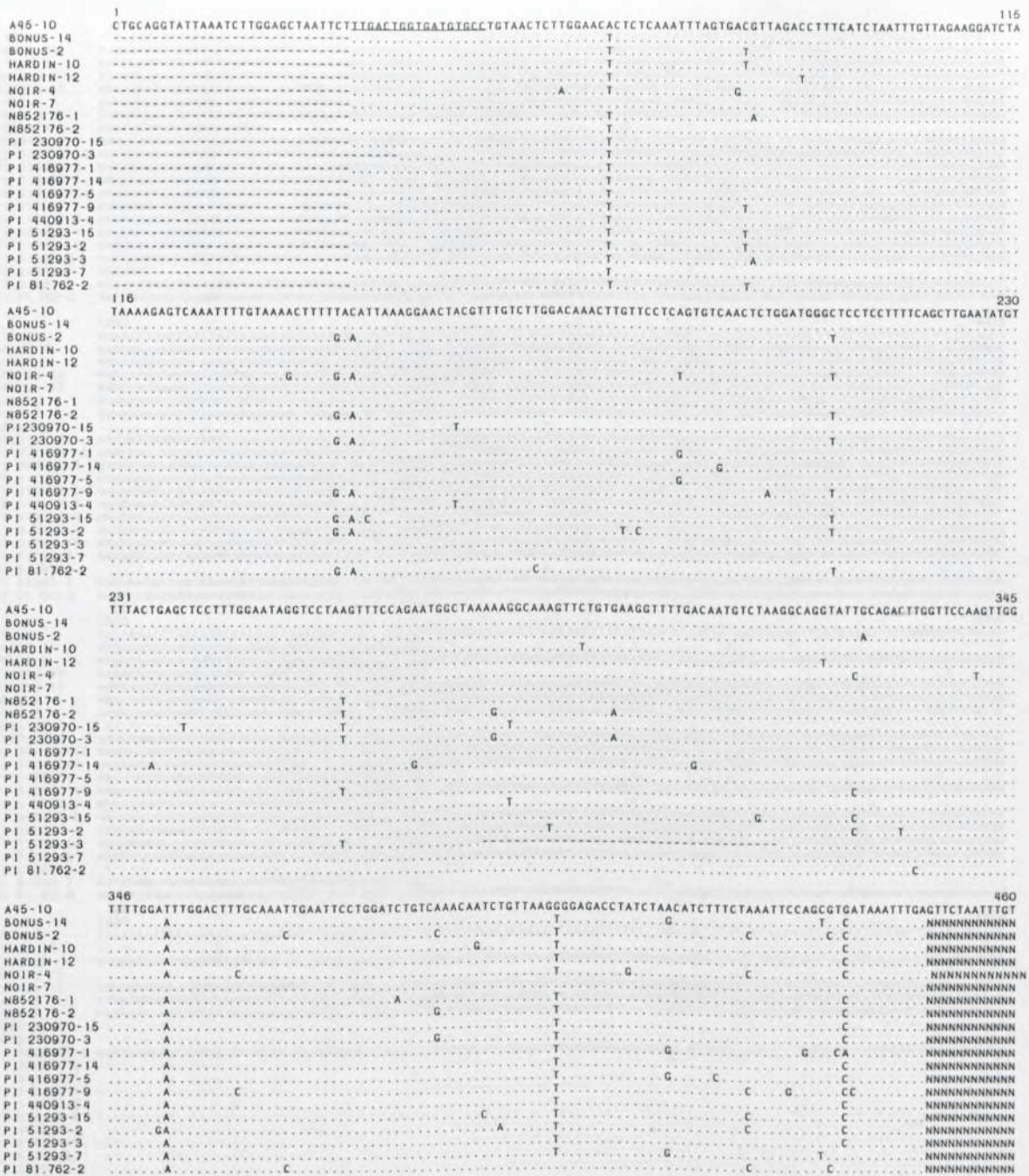


Figura 11 - Comparação das seqüências dos produtos de PCR de 1,5 a 2 kb amplificados com os "primers" 4510T7F e 4510T3F a partir de DNA genômico de diferentes genótipos de soja. Apenas cerca de 400 nucleotídios em cada uma das extremidades dos clones foram seqüenciados. A parte intermediária dos clones que não foram seqüenciados está representada por N. Os números que seguem os nomes dos genótipos de soja correspondem aos diferentes clones analisados. As seqüências dos "primers" utilizados estão sublinhadas, os pontos indicam identidade com a seqüência do clone de RFLP original e os traços designam deleções ou inserções introduzidas no alinhamento.

1181  
A45-10 ATATTGGGTGAGAAATGATTCATCTTCAAGTAATAAGCTGTGGAGCAATCCATCGAAATTCC-AAGCTATTTCCTTCGGGTTTTCAGACTGTGTAGAAATCATCTGCT  
BONUS-14  
BONUS-2  
HARDIN-10  
HARDIN-12  
N01R-4  
N01R-7  
N852176-1  
N852176-2  
P1 230970-15  
P1 230970-3  
P1 416977-1  
P1 416977-14  
P1 416977-5  
P1 416977-9  
P1 440913-4  
P1 51293-15  
P1 51293-2  
P1 51293-3  
P1 51293-7  
P1 81.762-2

1257  
GGAGAGATACCAAATGACATGGGAAATAAGAAATGTTAGAAATCCCTTCATCTCTCGCTAAACAACTTTTCAGGTCAAAATCCCTCAAAGCTTATCGATTGTCTTCTCCAGCTT  
A45-10  
BONUS-14  
BONUS-2  
HARDIN-10  
HARDIN-12  
N01R-4  
N01R-7  
N852176-1  
N852176-2  
P1 230970-15  
P1 230970-3  
P1 416977-1  
P1 416977-14  
P1 416977-5  
P1 416977-9  
P1 440913-4  
P1 51293-15  
P1 51293-2  
P1 51293-3  
P1 51293-7  
P1 81.762-2

1373  
TCTGAATCTATCATACCAACTTATCAGCGAATTTCCACGAGACCCCACTTCAGAGCTTTGATGAACCTTAGCTA-CACGGAAATCCTGAGCTTTGGTCCCTCTGTAA  
A45-10  
BONUS-14  
BONUS-2  
HARDIN-10  
HARDIN-12  
N01R-4  
N01R-7  
N852176-1  
N852176-2  
P1 230970-15  
P1 230970-3  
P1 416977-1  
P1 416977-14  
P1 416977-5  
P1 416977-9  
P1 440913-4  
P1 51293-15  
P1 51293-2  
P1 51293-3  
P1 51293-7  
P1 81.762-2

1488  
TCTGAATCTATCATACCAACTTATCAGCGAATTTCCACGAGACCCCACTTCAGAGCTTTGATGAACCTTAGCTA-CACGGAAATCCTGAGCTTTGGTCCCTCTGTAA  
A45-10  
BONUS-14  
BONUS-2  
HARDIN-10  
HARDIN-12  
N01R-4  
N01R-7  
N852176-1  
N852176-2  
P1 230970-15  
P1 230970-3  
P1 416977-1  
P1 416977-14  
P1 416977-5  
P1 416977-9  
P1 440913-4  
P1 51293-15  
P1 51293-2  
P1 51293-3  
P1 51293-7  
P1 81.762-2

1489  
CAAAAATATGCAACAACAAGCAATCGTTGAGGGACAGCTGCTCTGTGGACACGGTGATGGTAAATTCCTTGGAAACATCAGAGTTTTATATCGGTATGGCAGTTGGATTGCGAC  
A45-10  
BONUS-14  
BONUS-2  
HARDIN-10  
HARDIN-12  
N01R-4  
N01R-7  
N852176-1  
N852176-2  
P1 230970-15  
P1 230970-3  
P1 416977-1  
P1 416977-14  
P1 416977-5  
P1 416977-9  
P1 440913-4  
P1 51293-15  
P1 51293-2  
P1 51293-3  
P1 51293-7  
P1 81.762-2

1805  
AGGATTTTGGGGTTTTTTCAGTGTGTTTCTTCAACAGAACTGAGAGGCTTCCTATTTTCATTTACTTGGACACTTCAGAGACTGATTTATGTCATGATGCTTTT-GAAGGT  
A45-10  
BONUS-14  
BONUS-2  
HARDIN-10  
HARDIN-12  
N01R-4  
N01R-7  
N852176-1  
N852176-2  
P1 230970-15  
P1 230970-3  
P1 416977-1  
P1 416977-14  
P1 416977-5  
P1 416977-9  
P1 440913-4  
P1 51293-15  
P1 51293-2  
P1 51293-3  
P1 51293-7  
P1 81.762-2

1820  
AGGATTTTGGGGTTTTTTCAGTGTGTTTCTTCAACAGAACTGAGAGGCTTCCTATTTTCATTTACTTGGACACTTCAGAGACTGATTTATGTCATGATGCTTTT-GAAGGT  
A45-10  
BONUS-14  
BONUS-2  
HARDIN-10  
HARDIN-12  
N01R-4  
N01R-7  
N852176-1  
N852176-2  
P1 230970-15  
P1 230970-3  
P1 416977-1  
P1 416977-14  
P1 416977-5  
P1 416977-9  
P1 440913-4  
P1 51293-15  
P1 51293-2  
P1 51293-3  
P1 51293-7  
P1 81.762-2

Figura 11, Cont.

```

1721
A45-10 AACAAAGGTTACTTGGGAAA-TTATGAGTTTA-CCTGTTAA-CTGTTAAGAA-GTGAGACATTACAACGTTGTTAGCTAGCCATTA AAAAATGTGAAAAGGTTACCACATATGTA 1836
BONUS-14 .....G.....A.....G.GT.A.....T.A.....
BONUS-2 .....G.....A.....G.GT.A.....T.A.....
HARDIN-10 .....G.....A.....G.GT.A.....T.A.....
HARDIN-12 .....G.....A.....G.GT.A.....T.A.....
NOIR-4 .....A.G.....A.A.....A.....G.GT.A.....T.A.....
NOIR-7 .....A.G.....A.A.....A.....G.GT.A.....T.A.....
NB52176-1 .....G.....A.....G.GT.A.....TG.CA.....T.....
NB52176-2 .....G.....A.....G.GT.A.....TG.CA.....T.....
PI 230970-15 .....G.....A.....G.GT.A.....TG.CA.....T.....
PI 230970-3 .....G.....A.....G.GT.A.....TG.CA.....T.....
PI 416977-1 .....G.....A.....G.GT.A.....TG.CA.....T.....
PI 416977-14 .....A.....G.....G.....G.....T.....
PI 416977-5 .....A.....G.....G.....G.....T.....
PI 416977-9 .....A.....G.....G.....G.....T.....
PI 440913-4 .....A.....G.....G.....G.....T.....
PI 51293-15 .....G.....A.....G.GT.A.....T.A.....
PI 51293-2 .....G.....A.....G.GT.A.....T.A.....
PI 51293-3 .....G.....A.....G.GT.A.....T.A.....
PI 51293-7 .....G.....A.....G.GT.A.....T.A.....
PI 81.762-2 .....G.....A.....G.GT.A.....T.A.....

1837
A45-10 CCTGTTCACTGGGGTTCGTGCTGTTTTGCTTTTCATGGTATTAATGTTCTTAATGTTAAGAGAAGCTAGACGAGCTTGAGACTTGCAGTATCAAAATTCGTTTTGTATTATG 1952
BONUS-14 .....A.....A.G.....T.....A.....C.....
BONUS-2 .....A.....A.G.....T.....A.....C.....
HARDIN-10 .....A.....A.G.....T.....A.....C.....
HARDIN-12 .....A.....A.G.....T.....A.....C.....
NOIR-4 .....A.....A.G.....T.....A.....C.....
NOIR-7 .....A.....A.G.....T.....A.....C.....
NB52176-1 .....A.....A.G.....T.....A.....C.....
NB52176-2 .....A.....A.G.....T.....A.....C.....
PI 230970-15 .....A.....A.G.....T.....A.....C.....
PI 230970-3 .....A.....A.G.....T.....A.....C.....
PI 416977-1 .....A.....A.G.....T.....A.....C.....
PI 416977-14 .....A.....A.G.....T.....A.....C.....
PI 416977-5 .....A.....A.G.....T.....A.....C.....
PI 416977-9 .....A.....A.G.....T.....A.....C.....
PI 440913-4 .....A.....A.G.....T.....A.....C.....
PI 51293-15 .....A.....A.G.....T.....A.....C.....
PI 51293-2 .....A.....A.G.....T.....A.....C.....
PI 51293-3 .....A.....A.G.....T.....A.....C.....
PI 51293-7 .....A.....A.G.....T.....A.....C.....
PI 81.762-2 .....A.....A.G.....T.....A.....C.....

1953
A45-10 ACGCATTATTGTCTT-----CTACTATCTGGCACTAACAATGATTAGCATTAAATAACTGAAATGACACACTCGGTGATTGGATTAAGAAGATGTTTTACACAA 2068
BONUS-14 .....A.....A.....C.....
BONUS-2 .....A.....A.....C.....
HARDIN-10 .....A.....A.....C.....
HARDIN-12 .....A.....A.....C.....
NOIR-4 .....A.....A.....C.....
NOIR-7 .....A.....A.....C.....
NB52176-1 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
NB52176-2 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 230970-15 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 230970-3 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 416977-1 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 416977-14 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 416977-5 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 416977-9 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 440913-4 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 51293-15 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 51293-2 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 51293-3 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 51293-7 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....
PI 81.762-2 .....C.....CAGTATCAAAATG.....GT.T.G.....C.....

2069
A45-10 CATGTAACATAAGAATTTGCCTGAATGAATTAATTAGTGTAGAAAAGTTGGACCTTGCATAGGATTCG 2137
BONUS-14 .....C.....
BONUS-2 .....C.....
HARDIN-10 .....C.....
HARDIN-12 .....C.....
NOIR-4 .....T.....
NOIR-7 .....T.....
NB52176-1 .....T.....T.....
NB52176-2 .....T.....T.....
PI 230970-15 .....T.....T.....
PI 230970-3 .....T.....T.....
PI 416977-1 .....T.....T.....
PI 416977-14 .....T.....T.....
PI 416977-5 .....T.....T.....
PI 416977-9 .....T.....T.....
PI 440913-4 .....T.....T.....
PI 51293-15 .....T.....T.....
PI 51293-2 .....A.....T.....G.....TA.....
PI 51293-3 .....A.....T.....G.....TA.....
PI 51293-7 .....A.....T.....G.....TA.....
PI 81.762-2 .....A.....T.....G.....TA.....

```

Figura 11, Cont.

```

1                                     60
BONUS   CAGAATCCTATGCAAGCTCCACTATATATACAAAATGTTGAAGCTAAGCTATGACTGCAA
HARDIN  -----
NOIR    -----
N852176 -----
PI 416977 -----
PI 51293 -----
PI 81.762 -----

61                                     120
BONUS   AACTTGTTTACTTTGTTTCCTC---ACTGAATTCAATGGC---AACCAATACCTCAACT
HARDIN  -----
NOIR    -----
N852176 -----
PI 416977 -----
PI 51293 -----
PI 81.762 -----

121                                     180
BONUS   -TTATCTCATGTTGCAGCACTCAGGCCTTGTCTGCATCTGCAATTTGTCACAGGCACA
HARDIN  -----
NOIR    -----
N852176 -----
PI 416977 -----
PI 51293 -----
PI 81.762 -----

181                                     191
BONUS   ICACCAGTCAA
HARDIN  -----
NOIR    -----
N852176 -----
PI 416977 -----
PI 51293 -----
PI 81.762 -----

```

Figura 12 - Comparação dos produtos de amplificação de cerca de 200 pb produzidos a partir dos "primers" 4510T7F e 4510T3F, amplificados a partir de DNA genômico de diferentes genótipos de soja. As seqüências dos "primers" utilizados estão sublinhadas, os pontos indicam identidade com a seqüência do clone derivado do cultivar bônus e os traços designam deleções ou inserções produzidas no alinhamento.

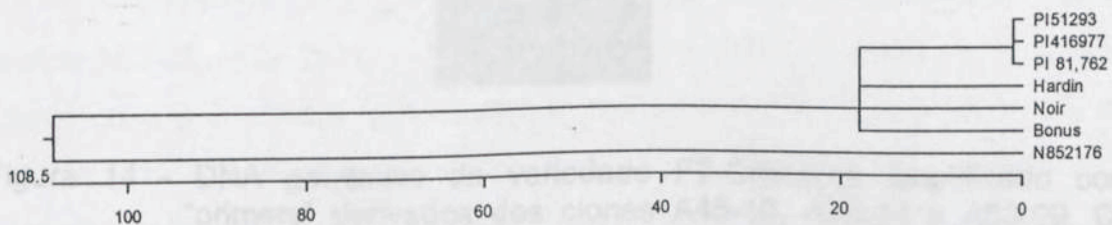


Figura 13 - Árvore filogenética construída a partir da matriz de distância, usando-se seqüências de nucleotídios dos produtos de amplificação de cerca de 200 pb, pelo método de alinhamento de Jotun Hein.

Primeiramente, as reações foram feitas, usando-se uma temperatura de anelamento de 55°C, mas nenhuma banda era visualizada com os “primers” desenhados com base na seqüência do clone A75-10 (dados não mostrados). Desse modo, usou-se uma temperatura de anelamento de 50°C em todas as reações. A amplificação com os “primers” desenhados para a ORF do clone A75-10 resultou em pelo menos seis bandas, que variavam de 900 a 200 pb (Figura 14). Já com os “primers” para o clone A45-10, foram produzidos dois fragmentos, um de 900 e outro de 564 pb. Para o clone A53-09, verificou-se a presença de uma única banda de aproximadamente 250 pb. Tais resultados indicam que as três sondas possuem seqüências homólogas na variedade FT-Cristalina e que, assim como nos outros genótipos analisados (Figura 2), a sonda A53-09 apresentou padrão menos hipervariável que as outras.

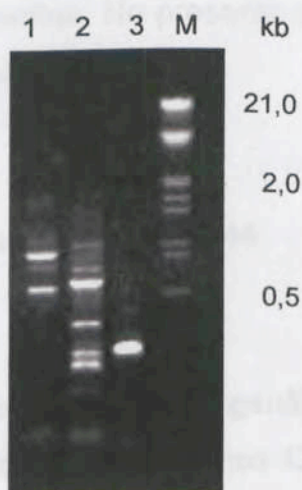


Figura 14 - DNA genômico da variedade FT-Cristalina amplificado com “primers” derivados dos clones A45-10, A75-10 e A53-09. Os produtos de amplificação foram separados por eletroforese em gel de agarose 1%. Os números na parte superior se referem aos “primers” usados (1 - A45-10, 2 - A75-10 e 3 - A53-09), e M refere-se ao marcador de peso molecular (DNA de  $\lambda$  digerido com *EcoR* I e *Hind* III). Os números à direita correspondem ao tamanho das bandas em kb.

#### 4.4. Isolamento de clones genômicos homólogos às sondas de RFLP

Uma vez detectadas seqüências homólogas aos três clones de RFLP na variedade FT-Cristalina, foi construída uma biblioteca genômica, com as finalidades de melhor caracterizar regiões hipervariáveis em uma variedade de soja brasileira e isolar possíveis candidatos a genes de resistência.

A sonda A53-09 foi produzida pela amplificação do clone de RFLP A53-09 com os "primers" T3 e 5309A (Tabela 2); a sonda A75-10 foi gerada pela amplificação do clone de RFLP A75-10, com os "primers" T7 e 7510A (Tabela 2); e a sonda A45-10 correspondia ao fragmento de *Pst* I, do clone de RFLP A45-10. As três sondas constituíam possíveis regiões codificadoras dos três clones. A temperatura de hibridização usada para as sondas dos clones A53-09 e A45-10 foi de 55°C; para o clone A75-10, foi de 50°C, posto que nenhum clone foi identificado quando se usou temperatura de hibridização de 55°C. Três clones foram identificados, usando-se a sonda A53-09 de  $3 \times 10^5$  fagos recombinantes analisados; dois clones isolados, utilizando-se a sonda A45-10 de  $9 \times 10^5$  fagos recombinantes; e somente um clone, usando-se a sonda A75-10 de  $9 \times 10^5$  fagos recombinantes. No presente trabalho, só foi seqüenciado um clone homólogo à sonda A53-09.

#### 4.5. Caracterização molecular do clone CR44

Dos três clones isolados da biblioteca genômica de FT-Cristalina com a sonda A53-09, dois clones, designados clones CR44 e CR45, pareciam ser iguais, já que aparentemente tinham o mesmo tamanho, cerca de 16 kb, e o mesmo padrão de restrição quando os seus DNAs foram submetidos à digestão com enzimas de restrição (dados não mostrados). O terceiro fago recombinante, chamado de clone CR40, é menor que os clones CR44 e CR45, tendo um tamanho estimado de cerca de 10 kb. Como o clone CR40 era menor que os outros dois e os clones CR44 e CR45 pareciam ser iguais, foi escolhido o clone CR44 para caracterização molecular.

O DNA do fago recombinante, clone CR44, foi submetido à digestão com as enzimas *EcoR* I, *Sal* I e *BamH* I, para construção do mapa de restrição. Os produtos de digestão foram analisados em géis de agarose 0,7% e seus tamanhos em pares de bases, estimados em relação a padrões de DNA, usando-se o programa Onedscan (Stratagene). A ordem dos fragmentos e a respectiva localização dos sítios de restrição foram confirmadas por meio de digestões duplas e triplas do DNA do clone CR44 (Figura 15). Concomitantemente à construção do mapa de restrição do clone CR44, foi feita uma análise de "Southern blot" (Figura 16) para identificar qual (ou quais) era(m) o(s) fragmento(s) homólogo(s) à sonda A53-09, para posteriores subclonagem e seqüenciamento.

Como pode ser visto na Figura 16, dois fragmentos menores hibridizaram com a sonda A53-09: um fragmento de cerca de 2,4 kb, produzido pela clivagem do DNA do fago recombinante CR44 com a enzima *EcoR* I; e um segundo fragmento de cerca de 2,0 kb, produzido pela digestão do clone CR44 com *Pst* I. Esses dois fragmentos foram subclonados no vetor pUC18 e denominados clone A53-09 E2 (fragmento de 2,4 kb de *EcoR* I) e clone A53-09 P4 (fragmento de 2,0 kb de *Pst* I).

O DNA dos dois subclones foi isolado e submetido à clivagem com as enzimas de restrição *BamH* I, *EcoR* I, *Hind* III e *Pst* I, com a finalidade de construir um mapa físico deles (Figuras 17 e 18). A partir desse mapa, foram determinadas as estratégias de subclonagem e de seqüenciamento dos dois subclones. O fragmento *Pst* I de 1,0 kb foi isolado do subclone A53-09 E2 e ligado ao vetor pUC18, previamente digerido com *Pst* I, resultando no subclone 21. A ordem dos fragmentos e os sítios de restrição foram, depois, confirmados por seqüenciamento. Traçada essa estratégia de subclonagem, os clones foram seqüenciados e após a etapa de subclonagem, desenhados "primers", para finalizar o seqüenciamento (Tabela 2).

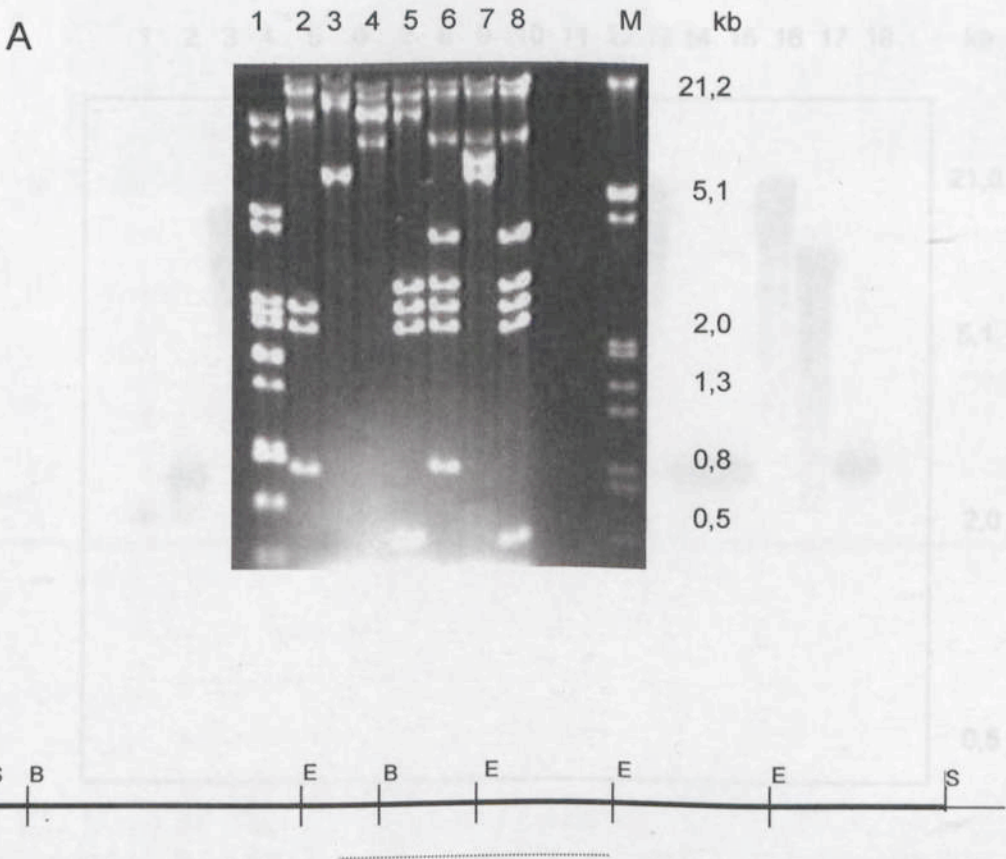


Figura 15 - Mapa de restrição do clone CR44. A - Digestão do clone CR44 com as enzimas *Pst* I (1); *EcoR* I (2); *BamH* I (3); *Sal* I (4); *EcoR* I e *BamH* I (5); *EcoR* I e *Sal* I (6); *BamH* I e *Sal* I (7); *EcoR* I, *BamH* I e *Sal* I (8); e DNA de fago  $\lambda$  clivado com *EcoR* I e *Hind* III (M). As setas indicam o tamanho das bandas em kb. B - Esquema do clone CR44. A linha grossa representa o inserto do fago recombinante, e as linhas fracas indicam as regiões adjacentes do vetor  $\lambda$ EMBL3. A linha pontilhada corresponde à região seqüenciada. As posições dos sítios das enzimas de restrição estão indicadas. As letras correspondem às iniciais das enzimas de restrição (S - *Sal* I, B - *BamH* I e E - *EcoR* I).

Figura 15 - Mapa de restrição do clone CR44. A - Digestão do clone CR44 com as enzimas *Pst* I (1); *EcoR* I (2); *BamH* I (3); *Sal* I (4); *EcoR* I e *BamH* I (5); *EcoR* I e *Sal* I (6); *BamH* I e *Sal* I (7); *EcoR* I, *BamH* I e *Sal* I (8); e DNA de fago  $\lambda$  clivado com *EcoR* I e *Hind* III (M). As setas indicam o tamanho das bandas em kb. B - Esquema do clone CR44. A linha grossa representa o inserto do fago recombinante, e as linhas fracas indicam as regiões adjacentes do vetor  $\lambda$ EMBL3. A linha pontilhada corresponde à região seqüenciada. As posições dos sítios das enzimas de restrição estão indicadas. As letras correspondem às iniciais das enzimas de restrição (S - *Sal* I, B - *BamH* I e E - *EcoR* I).



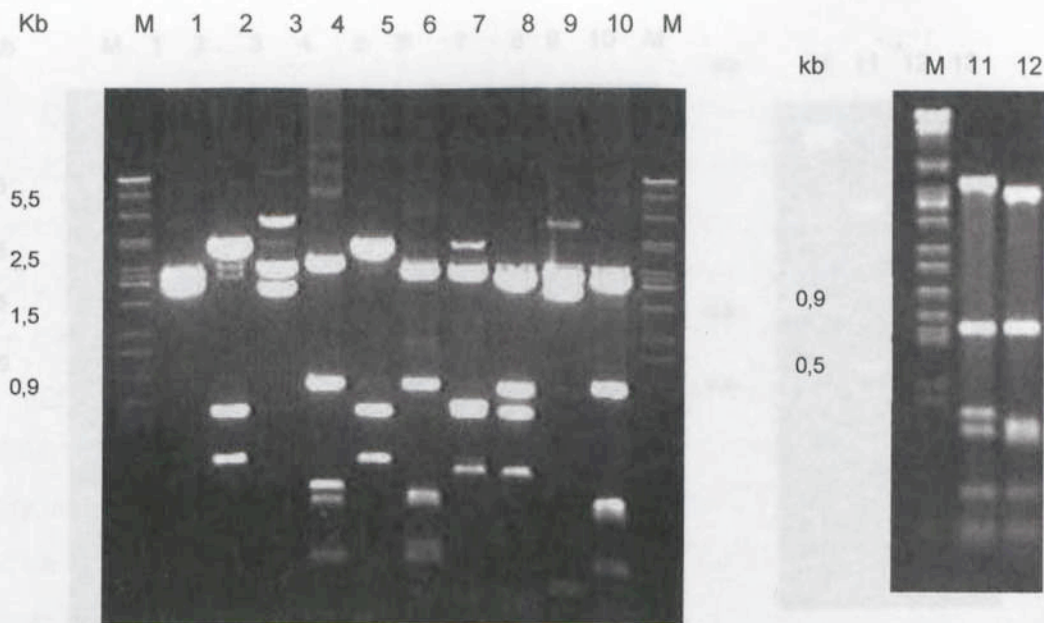
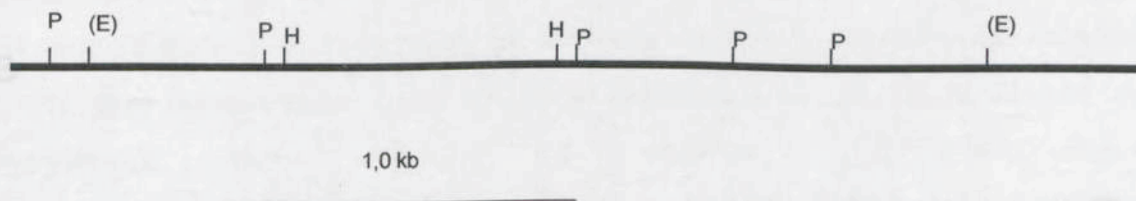
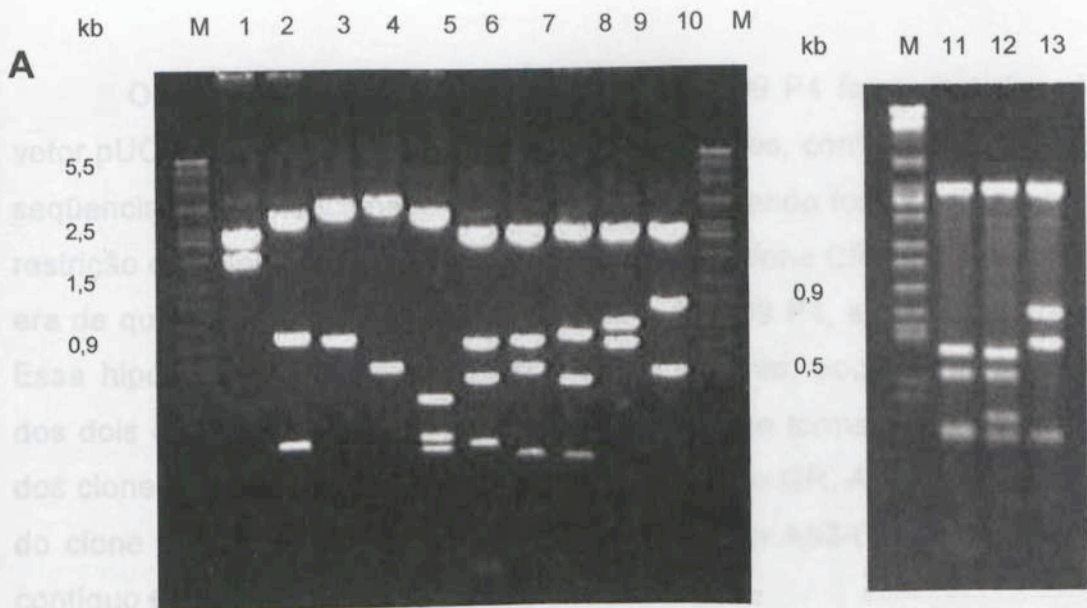
**A****B**

Figura 17 - Mapa de restrição do clone A53-09 E2. A - Padrão de restrição do clone A53-09 E2. Os números correspondem à digestão com enzimas de restrição (1 - *EcoR* I; 2 - *Hind* III; 3 - *Sal* I; 4 - *Pst* I; 5 - *Pst* I e *Hind* III; 6 - *Pst* I e *Sal* I; 7 - *Sal* I e *Hind* III; 8 - *EcoR* I e *Hind* III; 9 - *EcoR* I e *Sal* I; 10 - *EcoR* I e *Pst* I; 11 - *Hind* III e *Pst* I; e 12 - *Hind* III, *Pst* I e *EcoR* I), e M corresponde ao marcador de peso molecular (DNA de  $\lambda$  clivado com *EcoR* I, *Hind* III e *Bam*H I). As setas indicam o tamanho das bandas em kb. B - Esquema do mapa de restrição. Na primeira linha está esquematizado o clone A53-09 E2. As posições dos sítios das enzimas de restrição estão indicadas, estando o sítio de clonagem *EcoR* I entre parênteses. A linha inferior representa o subclone 21, fragmento de 1,0 kb *Pst* I. O número acima do fragmento indica o seu tamanho em kb, e as letras correspondem às iniciais das enzimas de restrição (E - *EcoR* I, P - *Pst* I e H - *Hind* III).



A sequência do clone A53-09 P4 foi analisada pelo programa ORF-Finder, para detectar possíveis ORFs. Foram identificadas neste clone possíveis ORFs; a maior delas potencialmente codifica um polipeptídeo de 262 aminoácidos (Figura 20). A seguinte sequência pode ser visualizada na Figura 21. Uma busca por domínios conservados nesta proteína potencial mostrou a presença de seis repetições de leucina conectada, de modo, 24 aminoácidos cada uma. Esses domínios LRRs têm semelhança com LRRs extracitoplasmáticas de proteínas de resistência, que possuem a seguinte sequência conservada:

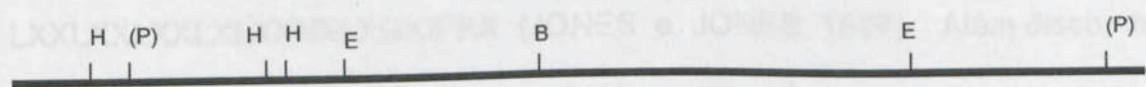


Figura 18 - Mapa de restrição do clone A53-09 P4. A - Padrão de bandas do clone A53-09 P4. Os números correspondem à reação de clivagem do DNA com as enzimas de restrição (1 - *Pst* I; 2 - *EcoR* I; 3 - *BamH* I; 4 - *Hind* III; 5 - *EcoR* I e *BamH* I; 6 - *EcoR* I e *Hind* III; 7 - *BamH* I e *Hind* III; 8 - *EcoR* I e *Pst* I; 9 - *Pst* I e *BamH* I; 10 - *Pst* I e *Hind* III; 11 - *EcoR* I, *BamH* I e *Hind* III; 12 - *EcoR* I, *BamH* I e *Pst* I; e 13 - *BamH* I, *Hind* III e *Pst* I), e M corresponde ao DNA do fago  $\lambda$  clivado com *EcoR* I, *Hind* III e *BamH* I. B. Esquema do mapa de restrição do clone A53-09 P4. Na primeira linha está esquematizado o clone A53-09 P4. As posições dos sítios das enzimas de restrição estão indicadas, estando o sítio de clonagem *Pst* I entre parênteses. As letras correspondem às iniciais das enzimas de restrição (E - *EcoR* I, P - *Pst* I, H - *Hind* III e B - *BamH* I). *Hind* III e *BamH* I estão indicadas.

#### 4.6. Análise da seqüência dos subclones A53-09 E2 e A53-09 P4

Os DNAs dos clones A53-09 E2 e A53-09 P4 foram subclonados no vetor pUC18 e seqüenciados em ambas as direções, conforme a estratégia de seqüenciamento mencionada no item anterior. Quando foram feitos o mapa de restrição e o "Southern blot" (Figuras 15 e 16) do clone CR44, a hipótese inicial era de que os dois subclones, A53-09 E2 e A53-09 P4, estariam sobrepostos. Essa hipótese foi confirmada pelo seqüenciamento, podendo o alinhamento dos dois clones ser visto na Figura 19. O contíguo formado pelo alinhamento dos clones A53-09 E2 e A53-09 P4 foi denominado GR. A seqüência completa do clone A53-09 E2 contém 2.475 pb; a do clone A53-09 P4, 2.102 pb; e o contíguo GR, 4.198 pb (Figura 20).

A seqüência do contíguo GR foi analisada pelo programa ORF Finder, para detectar possíveis ORFs. Foram identificadas quatro possíveis ORFs; a maior delas potencialmente codifica um polipeptídeo de 262 aminoácidos (Figura 20). A seqüência predita pode ser visualizada na Figura 21. Uma busca por domínios conservados nesta potencial proteína identificou seis repetições de leucina contendo, em média, 24 aminoácidos cada uma. Esses domínios LRRs têm semelhança com LRRs extracitoplasmáticas de proteínas de resistência, que possuem a seguinte seqüência-consenso: LXXLXXLXXLXXNXLXGXIPXX (JONES e JONES, 1997). Além disso, foi

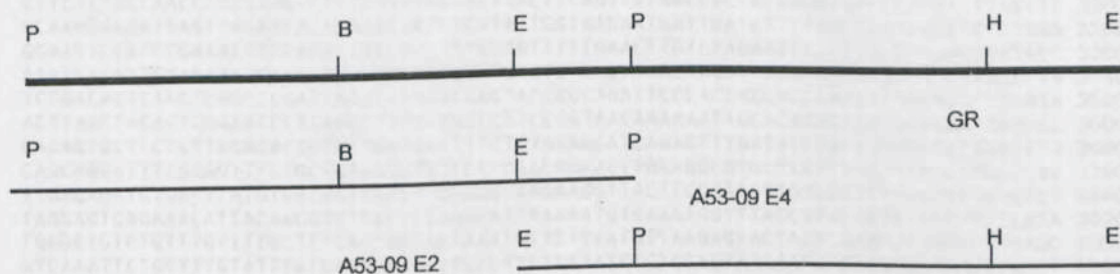


Figura 19 - Mapa de restrição de GR, contíguo de DNA formado pela sobreposição dos clones A53-09 E2 e A53-09 P4. As posições dos sítios das enzimas de restrição (E - *EcoR* I, P - *Pst* I, H - *Hind* III e B - *Bam*H I) estão indicadas.

GAATTCAAAGATTTACACGCTAGTCCATTGTTTAGAATAAACTGCTACATATATAAACCCTGTTTACAAAAAATAA 80  
 AAAATAAAAAGTTTAAATTATCTTTTGGCCCTTAAAGTTACGACAATATTTTTTTAGTCCTTATAAATTAATCATTCT 160  
 ATTTTAGTCCCAGTTGAATTTTTTTTTACTCTTTTGATTTCTCACCATCAAGTTTCGACTAACGTCATTGTTGAACTA 240  
 ATGTTGTGGAGAGGGCAAAGACACTCTCCTATGCACACACGAGACGTTGGTGATCATCTTTTCTCTCAACAACAAACCGG 320  
 CACCTGCAACATGCATCATCCCACTATGAAGGCACAACACTACCTCACCCTGTACTCATGGGACCAGAAACTCACAAT 400  
 CACACACTTTATCTGCAGCAATCAAACCATCATCCCAATTGGAACGATCAACTTAGTTCAATGCTGTGATACCACCTCAA 480  
 CTTGATGAGAGACTGTCAAGAAGGAGATTACATCATGGCAATTGGCAAAATGCTAAGCTTGCAGAGGTCGGTGGCAGACTC 560  
 AAGAATAATGGAATGTGTAGTGGCTTTGTCTCATACCTCCATTACACTTACAGTTGTACCTCAGCTTGCTTCCCTATTTT 640  
 ATTAATTTTGTTCATGTTTATGTGTGAGTTGAATCTTCCCAGCAAGGACCAAAAGATAAAAAAATGTAAATATAGGAA 720  
 CCAAAATGAATGTTTTAACTATAATGACTAAAAAATGATCACAATAAAGGACTAAAAGGATAGTTAAATCAAAT 800  
 AAAATAAACTAGTCAAAAATAATGGTGGTGGTGGTAGACACTTTTAGTCTAAATGTCTCGTTGTAGACATATATACCTC 880  
 TACATATGTGTGCATTTACCACCAATTAGTGGAAATAAAATTTAAATGCAAACCTTCTATGTTATATATCTTATATAGGC 960  
 ATTTGATGTGTAATTAACCAAGCTGAACAGAAAATTTGGATTCAACTTTTGGCGCTTCTCTTCTTCTCATGGCAGT 1040  
 GCTATATGCAACTCATGTTCTTCTCTTGTATTTGTCACGCAACAACCTCCAGTTTCAAGTCAAGCAAGGCAAGCAAGG 1120  
 TCAATATGACTTGCAGCGAGAAGGAAAGGAATGCACTCCCTCAGCTTCAAACATGGACTAGCAGACCCTTCAAACAGGCTT 1200  
 TCATCGTGGTCTGCACAAATCGGATTTGCTGTACATGGCCAGGGGTTCACTGCAACAACACAGGTAAGTCAATGGAAATCAA 1280  
 TCTTGACACACTGCGGGCTCTCCCTATAGGGAGTTGAGTGGAGAGATTAGTCTTCTTTGCTTGAACATAAAATTTTGA 1360  
 ATCGTTTGGACTTGAGTTCAAATATTTTGTCTTACTCCAATACCAAGCTTCTTAGGCTCATGGAGAGTCTGAGATAC 1440  
 TTTGGACCTTAGCTTAAAGTGGGTTTCAATGGGATTAATCCCTCATCAACTAGGAAACCTCTCAAACCTGCAGCAGCTTAACT 1520  
 TGGATCAAAATATGCTCTTCAGATAGATAACCTTAAATGGATTCAGGCTATCTTCTTGGAGTACCTTGATTTGAGTG 1600  
 GTTCAGACCTTCAATAAACCAAGGTAACCTGCTTCAAGTACTGAGTGCACCTTCCATCTCTTTTCAGAACACACTTGGAGAGC 1680  
 TGCAAAATTTGATAAATTAGGACCACCAAAAGGAAAGCCAACTTCACACATCTCCAAGTCTTGTATTTCAATTAACAA 1760  
 TCTCAATCAGAAAAATCCCTTCATGGCTATTTAACTCAGCACAACCTTGTTCCAACTTGAATTTACACAGTAACCTTTTA 1840  
 CAGGGACAAATTCACAATAATATCAAGCCTTCGAGAACATAAAAAATCTAGACCTGCAGAACCAACAGCTCAGTGGGG 1920  
 CACTTCCAGATTCATTAGGCCAGCTTAAAGCATCTTGAAGTTCTAAATCTCAGTAATAATACCTTTACTTGTCCAATTTCC 2000  
 TCACCATTTGCAAACCTTGTATCCTTGGAACTTTAAACCTGGCTCACAACCGACTGAATGGAACCATTCCTAAGAGTTT 2080  
 TGAGTTCCTCAGAAACCTGCAGGATTAATCTTGAAGTAAATCTTGGACTGGTGATATGCCTGTAACCTTTGGAACCT 2160  
 TCTCAAATTTAGTAATGTTAGACCTTTCAATTTGTTAGAAGGATCTATAAAAGAGTCAAATTTCTGTAACCTTTTGG 2240  
 AAATTAAGGAACTACGTTTGTCTTGGACAAACTTGTCTCAGTGTCAACTCAGGATGGGTTCTCTCTTTTTCAGCTTGA 2320  
 ATATGTTTACTGAGCTCCTTGGAAATAGGTCCTAACTTTCCAGAAATGGCTAAAAAGGCAAAAGTCTGTGAAGGTTTGA 2400  
 CAATGTCTAAGCAGGATCGCAGACTTGGTTCGAAGTGGTTTTGGAAATGAACTTCGCAAAATGAAATCCTGGATCTG 2480  
 TCAAACAATCTGTTAAGTGGAGCCTATCAACATCTTTCTCAATCCAGCGTCAATAAATTTGAGTCTAATTTGTTCAA 2560  
 GGGTACATTTGCCAAGTGTGCTGCAAAATGTTGAAGTGTGAATGTTGCCAATAACTCAATTTCCGGAACAATTTCTCCCT 2640  
 TTTTATGTTGGAAGGAAAATGCTACTGATAAGTTAAGTGTGCTTGATTTTCAAATAATGCTTTATATAGGTGATCTTG 2720  
 GTCAGTGGTGGTGCATTTGGCAACATTTGGTGCATTTGAACTTGGTGGTATGCAATTTGCTGGTGTATTTCAACCTCATG 2800  
 GGGTCTATCTAAGTGGATCTTGTGTTAAACAAACCTGCTCTCTGGATATATTCCTTCAACACTGCAAGTTTGAACAAAG 2880  
 GAGAATTTTCAATGAACAATGGGCAATAAACCAACTTTTCCGACGCAATAACCAGATTGGGATGGGGGAAAGGCAAAATCTA 2960  
 AATGGTCTTCTGCTAAGATCCAACAATTTCAATGGCAGCATTACTCAAAGATTTGTCAACTTCTTCCCTTATAGTG 3040  
 CTGGATCTTGGCAATAACAGCCTGTCAGGATCCATTTCAAATGTTTGGATGACATGAAGACAATGGCTGGTGAAGATGA 3120  
 CTCTTTGGCAAACCTTTAAGTTATTCATATAGCTCTGACTTCAAGTTAATAACCACTACAAGGAACTCTTGTCTTAGTCT 3200  
 CCAAAGGAGATGAGTTAGAGTACAGAGACAATCTGATATTGGTGAAGATGATTGATCTTTAAGTAATAAGCTGTCTGGA 3280  
 GCAATCCATCTGAAATTTCCAAGCTATCTGCTTTGGGTTTTTGAACCTGTCTAGAAATCATCTGCTGGAGGGATACC 3360  
 AAATGACATGGGAAAAATGAAATTTGTTAGAATCCCTTGTCTCTCACTAAACAACATTTCAAGGTCATCCCTCAAAGCTTA 3440  
 TCTGACACTCTCCTCGGTCGGATTAATCAAAACAACACTATCGGAGATTCCCAAGGACCCCAACTTCAGAGCTGGAGA 3520  
 ACTTAGCTACACTGGAATCCTGAAGCTTTTGGTCTTCTCTGTAACAAAAAATGCACAGACCAAGGAAGAGTTGACCC 3600  
 GAGAGTCTTCTGTTGGACACGGTGTAGGTAATTTCTTTGGAACATCAGAGTTGATATCGGATGGGAGTTGGATTTG 3680  
 CAGCAGGATTTGGGGTTTTTGCAGTGTGGTTTTCTTCAACAGAACTTGGAGGCGTGCTTATTTTCATTATCTTGACCAC 3760  
 TTGAGAGACTGATTTATGTGATTAAGTTTGAAGGTAAGAAGGTTACTTGGGAAATTTGTGAGTTTACATGTTAACTGT 3840  
 TAGGAGTGAGAAACATTACAACGTGTTATTCTAGACATTAATAATGTGAAAAGGTTTACCATACTATGTACATGTTTATA 3920  
 TGGGGTGTCTGTTTGTGTTTCTTTCATTGGTAATAAATGTTCTTAATGTTAAGAGAATAGACGAGCTTCAGACTTGAGC 4000  
 ATCAAATCTGTTTTGATTTATGACGCATTAATGCTTCTACTATCTGTGCCTAACAATGATTAGCATTAAAATACCT 4080  
 GGAATGACACACTCGGTGATTGGATTTAAGAAGATTGTTTTACGCAACGTGTAACCTTAGAATTTGCTGAAATGAAATTTAA 4160  
 TTAGTGTAGAAAGTTGGAGCTTGCATAGGATTCTGCAG 4198

Figura 20 - Sequência de nucleotídios do contígua GR. As ORF estão destacadas de vermelho. O sinal de poliadenilação está destacado de azul.

também encontrado o sinal clássico de poliadenilação (AATAAA) localizado na posição 2.903 a 2.908 (Figura 20).

A comparação da seqüência de aminoácidos predita em todas as possíveis ORFs com outras existentes no “GenBank” revelou homologia entre GR e outras proteínas codificadas por genes de resistência. GR apresentou similaridade de seqüência com as proteínas “Cf-2.1-like” de *Arabidopsis thaliana* (acesso no “GenBank” AC004238), proteína de resistência a doenças homóloga a Cf2/Cf-5 de *Hordeum vulgare* (acesso no “GenBank” AF166121) e outras proteínas de resistência de *Lycopersicon esculentum*. Esses dados em conjunto indicam que GR possui similaridade com genes de resistência.

#### 4.7. Análise da expressão órgão-específica de A45-10, A53-09 e A75-10

Foi feito um RT-PCR, com o objetivo de verificar se os clones de RFLP continham seqüências que eram expressas em diferentes órgãos da variedade de soja FT-Cristalina, bem como eliminar o fato de que esses clones poderiam ser pseudogenes. O RNA total de raiz, haste e folha do cultivar FT- Cristalina, estágio V3, foi isolado e amplificado, usando-se a técnica de RT-PCR descrita em “Material e Métodos”. A etapa de amplificação da primeira fita de cDNA foi

```

MAVLYATHVLLLILSTATTLQFSASKAARLNMTCSSEKERNALLSFKHGLADPSNRLSSWSDKSDCCTWPG
VHCNNTGKVMELNLDTPAGSPYRELSGEISPSLLELK
.....YLNRLDLSSNYFVLTPIPSFLGSLE
.....SLRYLDLSLGFMLIPHQLGNLS
.....NLQHLNLGYNYALQIDNLNWSRLS
.....SLEYLDLSGSDLHKQGNWLQVLS
ALP.....SLSELHLESCQIDNLGPPKGKA
NFT.....HLOVLDLSINNLNOKIPSWLFNLST
TLVPT
  
```

Figura 21 - Seqüência de aminoácidos predita pela maior ORF presente no contígua GR. A seqüência mostra seis domínios LRR alinhados entre si. Os pontos foram inseridos para facilitar o alinhamento dos domínios LRR. Os resíduos de leucina estão destacados de vermelho.

feita para o clone A45-10 com os "primers" 4510T7B1 e 4510T3C1; para o clone A53-09, usaram-se os "primers" 5309A e 5309D e para o clone A75-10, os "primers" 7510A e 7510T7F (Tabela 2). Esses "primers" foram escolhidos porque amplificavam regiões possivelmente transcritas. As amostras de RNA total também foram amplificadas com os "primers" para actina 3, a fim de verificar se havia contaminação com DNA e eliminar artefatos experimentais. A temperatura de pareamento nas reações para A45-10 e A53-09 foi de 55°C e para A75-10 e actina 3, de 50°C.

Na Figura 22, mostra-se que existem seqüências transcritas similares aos três clones analisados em todos os órgãos examinados. A amplificação com os "primers" para o clone A53-09 resultou em uma banda de aproximadamente 300 pb, em todas as amostras, porém o acúmulo de mRNA parece ser diferente. O órgão que aparenta ter maior expressão é a haste, seguida pela raiz e folha. Esse resultado indica que, apesar de ter sido detectada a expressão em todos os órgãos, deve haver controle do nível de expressão.

Os produtos de amplificação com os "primers" para A53-09 foram seqüenciados diretamente e alinhados com GR. Eles apresentam 100% de identidade entre si, nos três órgãos, e se alinham entre os nucleotídios 1.033 e 1.821, onde se localiza a ORF predita de 789 nucleotídios (Figura 23). Isso indica que a ORF deduzida é, de fato, transcrita. Para testar essa hipótese, um clone de cDNA correspondente ao contíguo GR deverá ser clonado e, assim, delimitado o gene.

Já para o clone A75-10 foi detectada a presença de seqüências similares em todos os órgãos. A amplificação com esse conjunto de "primers" resultou em mais de uma banda para cada amostra, e, ainda, as bandas apresentaram tamanhos diferentes e foram polimórficas em relação aos órgãos examinados. Isso indica que os "primers" estavam pareando com mais de um transcrito.

A amplificação com os "primers" para o clone A45-10 resultou em única banda fraca, de aproximadamente 900 pb, em todos os órgãos analisados. Esse resultado não era esperado, já que, dos três clones, esse era o que possuía ORF maior e, talvez, fosse o único com seqüências transcritas. Os resultados apresentados aqui estão de acordo com os alcançados nos



GR.....ATGGCAGTGTATATGCAACTCATGTTCTTCTCTTGATATTGTCCACTGCAACAACCTCCAGTTCAGTGAAGCAAG  
Raiz.....  
Haste.....  
Folha.....

GR.....GCAGCAAGGCTCAATATGACTTGCAGCGAGAAGGAAAGGAATGCACTCCTCAGCTTCAACATGGACTAGCAGACCCCT  
Raiz.....CAGCGAGAAGGAAAGGAATGCACTCCTCAGCTTCAACATGGACTAGCAGACCCCT  
Haste.....CAGCGAGAAGGAAAGGAATGCACTCCTCAGCTTCAACATGGACTAGCAGACCCCT  
Folha.....CAGCGAGAAGGAAAGGAATGCACTCCTCAGCTTCAACATGGACTAGCAGACCCCT

GR.....TCAAACAGGCTTTCATCGTGGTCTGACAAATCGGATTGCTGTACATGGCCAGGGGTTCACTGCAACAACACAGGTAAA  
Raiz.....TCAAACAGGCTTTCATCGTGGTCTGACAAATCGGATTGCTGTACATGGCCAGGGGTTCACTGCAACAACACAGGTAAA  
Haste.....TCAAACAGGCTTTCATCGTGGTCTGACAAATCGGATTGCTGTACATGGCCAGGGGTTCACTGCAACAACACAGGTAAA  
Folha.....TCAAACAGGCTTTCATCGTGGTCTGACAAATCGGATTGCTGTACATGGCCAGGGGTTCACTGCAACAACACAGGTAAA

GR.....GTCATGGAAATCAATCTTGACACACCTGCGGGCTCTCCCTATAGGGAGTTGAGTGGAGAGATTAGTCTCTCTTTGCTT  
Raiz.....GTCATGGAAATCAATCTTGACACACCTGCGGGCTCTCCCTATAGGGAGTTGAGTGGAGAGATTAGTCTCTCTTTGCTT  
Haste.....GTCATGGAAATCAATCTTGACACACCTGCGGGCTCTCCCTATAGGGAGTTGAGTGGAGAGATTAGTCTCTCTTTGCTT  
Folha.....GTCATGGAAATCAATCTTGACACACCTGCGGGCTCTCCCTATAGGGAGTTGAGTGGAGAGATTAGTCTCTCTTTGCTT

GR.....GAACTAAAATATTTGAATCGTTGGACTTGAGTTCAAATTATTTGTTCTTACTCCAATACCAAGCTTCTTAGGCTCA  
Raiz.....GAACTAAAATATTTGAATCGTTGGACTTGAGTTCAAATTATTTGTTCTTACTCCAATACCAAGCTTCTTAGGCTCA  
Haste.....GAACTAAAATATTTGAATCGTTGGACTTGAGTTCAAATTATTTGTTCTTACTCCAATACCAAGCTTCTTAGGCTCA  
Folha.....GAACTAAAATATTTGAATCGTTGGACTTGAGTTCAAATTATTTGTTCTTACTCCAATACCAAGCTTCTTAGGCTCA

GR.....ITGGAGAGTCTGAGATACTTGACCTTAGCTTAAGTGGGTTTCATGGGATTAATCCCTCATCAACTAGGAAACCTCTCA  
Raiz.....ITGGAG  
Haste.....ITGGAG  
Folha.....ITGGAG

GR.....AACCTGCAGCACCTTAATCTTGATACAATTATGCTCTTTCAGATAGATAACCTTAATTGGATATCAAGGCTATCTTCC  
Raiz.....  
Haste.....  
Folha.....

GR.....TTGGAGTACCTTGATTTGAGTGGTTTCAGACCTTCATAAACAAGGTAAGTGGCTTCAAGTACTGAGTGCACCTTCATCT  
Raiz.....  
Haste.....  
Folha.....

GR.....CTTTCAGAACTACACTTGAGAGCTGTCAAATTGATAACTTAGGACCACCAAAAGGAAAAGCCAACCTTCACACATCTC  
Raiz.....  
Haste.....  
Folha.....

GR.....CAAGTCTTGATCTTTCAATTAACAATCTCAATCAGAAAATCCCTTCATGGCTATTTAATCTCAGCACAACTCTTGT  
Raiz.....  
Haste.....  
Folha.....

GR.....CCAACCTGA  
Raiz.....  
Haste.....  
Folha.....

Figura 23 - Comparação da ORF de 789 nucleotídios com os produtos de RT-PCR amplificados a partir de amostras de RNA total de raiz, haste e folha da variedade de soja FT-Cristalina. As seqüências dos "primers" utilizados no RT-PCR estão sublinhadas.

laboratórios da DuPont (EUA). Em uma busca no banco de dados de EST ("Expressed Sequence Tags"), da DuPont, foram encontrados dois clones homólogos a A53-09, um a A75-10 e nenhum a A45-10 (Dr. Rafalski, comunicação pessoal).

Os dois clones de cDNA homólogos a A53-09 foram isolados de vagens em desenvolvimento de soja. O clone de cDNA *sdp2c.pk007.p18* apresentou 100% de homologia a A53-09 do nucleotídeo 49 ao 522. O clone cDNA *sdp2c.pk007.a22* é 95% homólogo a A53-09, do nucleotídeo 49 ao 526. Os dois clones cDNA são homólogos à proteína "Cf-2.1-like" de *Arabidopsis* e a outros genes de resistência de tomate. O clone de cDNA *s1s1c.pk010.j1*, isolado de soja infectada com *Sclerotinia sclerotiorum*, apresentou 83% de similaridade com A75-10. Esse clone de cDNA é homólogo à proteína N, que confere resistência ao TMV. Esses dados, em conjunto (Dr. Rafalski, dados não publicados) com os obtidos neste trabalho, indicaram que A53-09 é ativamente transcrito e que A75-10 é membro de uma família multigênica de genes com homologia com genes de resistência a doenças. Apesar de A45-10 ter mostrado alta similaridade com genes de resistência (Tabela 3) e padrão bem definido de LRR (Figura 9) não foi possível detectar expressão significativa no experimento de RT-PCR (Figura 22) e encontrar um clone de cDNA (dados da DuPont) que indicasse que ele era ativamente transcrito. Portanto, A45-10 pode ser um pseudogene nas variedades de soja analisadas nesses dois trabalhos ou, ainda, ser expresso em situações diferentes das analisadas no presente trabalho.

## 5. RESUMO E CONCLUSÕES

A soja (*Glycine max* (L.) Merrill) é uma cultura autógama que vem sendo cultivada pelo homem há muito tempo. Apesar de possuir base genética extensa, como espécie, a soja cultivada apresenta essa base genética bastante estreita. A análise da variabilidade genética de variedades de soja por meio de hibridizações com sondas de RFLP revelou grande número de bandas monomórficas e algumas bandas polimórficas de baixo número de cópias, e número muito pequeno de sondas mapeia em regiões hipervariáveis do genoma. Durante a construção de um mapa genético entre *G. max* e *G. soja*, na DuPont (EUA), usando-se sondas de RFLP de uma biblioteca genômica da variedade de soja Bônus, foram analisadas 2.000 sondas, 75% das quais eram monomórficas (RAFASLKI e TINGEY, 1993). Dentre as sondas polimórficas, 0,5% delas eram extremamente polimórficas. Cinco dessas sondas foram selecionadas e caracterizadas molecularmente.

Duas sondas (A1-10 e A2-08) não apresentaram similaridade com seqüências depositadas no "GenBank". Entretanto, esses dois clones eram ricos em seqüências A:T. Tal fato pode explicar o padrão de hibridização altamente polimórfico, revelado por essas sondas. A comparação das seqüências das outras três sondas (A45-10, A53-09 e A75-10) revelou que existia alta similaridade entre elas e outras seqüências de genes que codificam proteínas de resistência a doenças em várias espécies de plantas. Nesse caso,

o polimorfismo pode ser explicado devido à presença de domínios LRRs nesses clones.

Devido ao fato de a sonda A45-10 apresentar padrão de hibridização extremamente polimórfico entre os genótipos de soja analisados, foram construídos “primers” que amplificavam regiões polimórficas em soja com a finalidade de obter padrão de amplificação que pudesse ser usado como “fingerprint” na identificação molecular de variedades. Os produtos, na faixa de 1,5 a 2,0 kb, resultantes das reações de amplificação foram capazes de distinguir os diferentes genótipos analisados, podendo, em conjunto com outros tipos de marcadores, ser usados para identificação de variedades de soja.

A partir de uma biblioteca genômica da variedade de soja FT-Cristalina, foram isolados clones, usando-se as três sondas de RFLP (A45-10, A53-09 e A75-10), com a finalidade de obter seqüências homólogas a genes de resistência em um cultivar brasileiro resistente a algumas doenças. No presente trabalho, somente um dos clones, que apresentava homologia com a sonda A53-09, denominado clone 44, foi caracterizado molecularmente, o qual possui tamanho aproximado de 16 kb. Dois fragmentos que hibridizaram com a sonda A53-09 foram subclonados e seqüenciados. O contíguo formado por esses dois clones contém 4.198 pb e foi denominado GR.

GR possui similaridade de seqüência com genes que codificam as proteínas “Cf-2.1-like” de *A. thaliana*, proteína de resistência a doenças homóloga a Cf2/Cf5 de *Hordeum vulgare* e outras proteínas de resistência do gênero *Lycopersicum*. Quando a seqüência de GR foi analisada pelo programa ORF Finder, para predição de possíveis ORFs, encontrou-se uma ORF de 789 nucleotídios, que codificava uma potencial proteína com 262 aminoácidos. Essa proteína possui seis domínios LRRs, em um padrão semelhante ao de proteínas de resistência extracitoplasmáticas. Além disso, a seqüência de GR possui, logo em seguida à ORF de 789 nucleotídios, um sinal de poliadenilação.

Quando amostras de RNA de raiz, haste e folha foram amplificadas com “primers” específicos para as sondas A45-10, A75-10 e A53-09, detectaram-se transcritos homólogos aos três clones de RFLP. A amplificação com os “primers” para A45-10 resultou em uma banda muito fraca nos três órgãos.

Para A75-10, foi detectada a presença de várias bandas polimórficas, nos três órgãos, indicando que os "primers" pareavam com mais de um transcrito.

A amplificação para A53-09 resultou em uma banda de aproximadamente 300 pb em todos os órgãos analisados. Entretanto, a maior intensidade da banda presente na amostra da haste indicou maior nível de expressão neste órgão em relação aos outros dois analisados. Os produtos de amplificação para A53-09 dos três órgãos foram seqüenciados. O alinhamento dessas seqüências com o clone GR revelou que elas se alinhavam perfeitamente com a ORF predita de 789 nucleotídios, indicando que essa ORF é realmente transcrita nos três órgãos da variedade de soja FT-Cristalina analisados. Tal resultado, juntamente com a análise de similaridade de seqüência, evidenciou que GR codifica uma proteína homóloga a proteínas de genes de resistência a doenças.

Como continuação do presente trabalho, o clone 44 deverá ser seqüenciado completamente, com a finalidade de identificar possíveis elementos *cis* que possam estar atuando na regulação deste gene. Além disso, o clone de cDNA correspondente deverá ser isolado para delimitar possíveis introns e exons. Para determinar se esse gene codifica um gene *R*, deverão ser conduzidos experimentos que indiquem que GR está co-segregando com algum marcador de resistência e, depois, experimentos de complementariedade genética.

AUSUBEL, F. M., BRENT, R., KINGSTON, R. E., MORRIS, D. D., SEIDMAN, J. G., SMITH, J. & Current protocols in molecular biology in CD-ROM. New York: John Wiley & Sons, Inc., 1995.

BAKER, B., ZAMBRYK, P., STASKAWICZ, B., DINESH-KUMAR, S. P. Signalling in plant-microbe interactions. Science, v. 276, p. 726-733, 1997.

BENDAHMANE, A., KANYUKA, K., SAULCOMBE, D. G. The *Rx* genes from potato control resistance to late blight and cell death responses. Plant Cell, v. 11, p. 781-793, 1999.

BENT, A. P. Plant disease resistance genes: function meets structure. Plant Cell, v. 8, p. 1.757-1.771, 1996.

BENT, A., KUNKEL, B., THOMPSON, G., BROWN, K., SCHMIDT, R., GIRAUDAT, J., LELING, J., CHAMPAGNE, S. RPS2 of Arabidopsis thaliana: A leucine-rich repeat class of plant disease resistance genes. Science, v. 265, p. 1.856-1.860, 1994.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABDELNOOR, R.V., BARROS, E.G., MOREIRA, M.A. Determination of genetic diversity within Brazilian soybean germplasm using random amplified polymorphic DNA techniques and comparative analysis with pedigree data. **Braz. J. Genet.**, v. 18, p. 265-273, 1995.
- ALTSCHUL, S. F., MADDEN, T. L., SCHÄFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W., LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucl. Acid. Res.**, v. 25, p. 3.389-3.402, 1997.
- ANDERSON, P. A., LAWRENCE, G. J., MORRISH, B. C., AYLIFFE, M. A., FINNEGAN, E. J., ELLIS, J. G. Inactivation of the flax rust resistance gene *M* associated with loss of a repeated unit within the leucine-rich repeat coding region. **Plant Cell**, v. 9, p. 641-651, 1997.
- AUSUBEL, F. M., BRENT, R., KINGSTON, R. E., MOORE, D. D., SEIDMAN, J. G., SMITH, K. S. **Currents protocols in molecular biology in CD-ROM**. New York: John Wiley & Sons, Inc., 1998.
- BAKER, B., ZAMBRYSKI, P., STASKAWICZ, B., DINESH-KUMAR, S. P. Signalling in plant-microbe interactions. **Science**, v. 276, p. 726-733, 1997.
- BENDAHMANE, A., KANYUKA, K., BAULCOMBE, D. C. The *Rx* gene from potato controls separate virus resistance and cell death response. **Plant Cell**, v. 11, p. 781-791, 1999.
- BENT, A. F. Plant disease resistance genes: function meets structure. **Plant Cell**, v. 8, p. 1.757-1.771, 1996.
- BENT, A., KUNKEL, B., DAHLBECK, D., BROW, K., SCHMIDT, R., GIRAUDAT, J., LEUNG, J., STASKAWICZ, B. *RPS2* of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. **Science**, v. 265, p. 1.856-1.860, 1994.

- BOTELLA, M. A., PARKER, J. E., FROST, L. N., BITTNER-EDDY, P. D., BEYNON, J. L., DANIELS, M. J., HOLUB, E. B., JONES, J. D. G. Three genes of the *Arabidopsis* RPP1 complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. **Plant Cell**, v. 10, p. 1.847-1.860, 1998.
- CAI, D., KLEINE, M., KIFLE, S., HARLOF, H-J., SANDAL, N. N., MARCKER, K. A., KLEIN-LANKHORST, R. M., SALENTIJN, E. M. J., LANGE, W., STIEKEMA, W. J., WYSS, U., GRUNDLER, F. M. W., JUNG, C. Positional cloning of a gene for nematode resistance in sugar beet. **Science**, v. 275, p. 832-834, 1997.
- CRUTE, I. R., PINK, D. A. Genetics and utilization of pathogen resistance. **Plant Cell**, v. 8, p. 1.747-1.755, 1996.
- DANGL, J. L., DIETRICH, R. A., RICHBERG, M. H. Death don't have no mercy: Cell death programs in plant-microbe interactions. **Plant Cell**, v. 8, p. 1.793-1.807, 1996.
- DANGL, J. Plants just say NO to pathogens. **Nature**, v. 394, p. 525-527, 1998.
- DELANNAY, X., RODGERS, D.M., PALMER, R.G. Relative genetic contribution among ancestral lines to North American soybean cultivars. **Crop Sci.**, v. 23, p. 944-949, 1983.
- DELLEDONNE, M., XIA, Y., DIXON, R. A., LAMB, C. Nitric oxide functions as signal in plant disease resistance. **Nature**, v. 394, p. 585-588, 1998.
- DIXON, M. S., HATZIXANTHIS, K., JONES, D. A., HARRISON, K., JONES, J. D. G. The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. **Plant Cell**, v. 10, p. 1.915-1.925, 1998.
- DIXON, M. S., JONES, D. A., KEDDIE, J. S., THOMAS, C. M., HARRISON, K., JONES, J. D. G. The tomato *Cf2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. **Cell**, v. 84, p. 451-459, 1996.
- DIXON, R. A., HARRISON, M. J., LAMB, C. J. Early events in the activation of plant defence. **Annu. Rev. Plant Phytopathol.**, v. 32, p. 479-501, 1994.
- DOYLE, J. J., DOYLE, J. L. Isolation of plant DNA from fresh tissue. **Focus**, v. 12, p. 13-15, 1990.
- DURNER, J., WENDEHENNE, D., KLESSIG, D. F. Defence gene induction in tobacco by nitric oxide, cyclic GMP and cyclic ADP ribose. **Proc. Natl. Acad. Sci. USA**, v. 95, p. 10.328-10.333, 1998.
- ELLIS, J. G., LAWRENCE, G. J., AYLIFFE, M., ANDERSON, P., COLLINS, N., FINNEGAN, J., FROST, D., LUCK, J. PRYOR, T. Advances in molecular genetic analysis of the flax-flax rust interaction. **Ann. Rev. Phytopathol.**, v. 35, p. 271-291, 1997.
- ELLIS, J. G., LAWRENCE, G. J., LUCK, J. E., DODDS, P. N. Identification of regions in alleles of the flax rust resistance gene *L* that determine in gene-for-gene specificity. **Plant Cell**, v. 11, p. 495-506, 1999.

- FLOR, H. H. Current status of gene-for-gene concept. **Ann. Rev. Phytopathol.**, v. 9, p. 275-296, 1971.
- FLOR, H. H. The complementary genetic system in flax and flax rust. **Adv. Genet.**, v. 8, p. 29-54, 1956.
- GIZLICE, Z., CARTER, JR., T. E., BURTON, J. W. Genetic diversity in North American soybean: I. Multivariate analysis of founding stock and relation to coefficient of parentage. **Crop Sci.**, v. 33, p. 614-620, 1993.
- GOLDBERG, R.B. DNA sequence organization in the soybean plant. **Biochem. Genet.**, v. 16, p. 45-68, 1978.
- GRANT, M., GODIARD, L., STRAUBE, E., ASHFIELD, T., LEWALD, J., SATTLER, A., INNES, R., DANGL, J. Structure of the *Arabidopsis thaliana* *RPM1* gene enabling dual specificity disease resistance. **Science**, v. 269, p. 843-846, 1995.
- HAMMOND-KOSACK, K. E., JONES, J. D. G. Resistance gene-dependent plant defence responses. **Plant Cell**, v. 8, p. 1.773-1.791, 1996.
- HAMMOND-KOSACK, K. E., JONES, J. D. G. Plant disease resistance genes. **Ann. Rev. Plant Physiol. Plant Mol. Biol.**, v. 48, p. 573-607, 1997.
- HIGGINS, D. G., SHARP, P. M. Fast and sensitive multiple sequence alignments on a microcomputer. **CABIOS**, v. 5, n. 2, p. 151-153, 1989.
- HIROMOTO, D. M., VELLO, N. A. The genetic base of Brazilian soybean (*Glycine max* (L.) Merrill) cultivars. **Braz. J. Genet.**, v. 9, p. 295-306, 1986.
- INOUE, H., NOJIMA, H., OKAYAMA, H. High efficiency transformation of *Escherichia coli* with plasmids. **Gene**, v. 96, p. 23-28, 1990.
- JACKSON, A. O., TAYLOR, C. B. Plant-microbe interactions: life and death at the interface. **Plant Cell**, v. 8, p. 1.651-1.668, 1996.
- JOHAL, G. S., BRIGGS, S. P. Reductase activity encoded by the *HM1* disease resistance gene in maize. **Science**, v. 258, p. 985-987, 1992.
- JOHAL, G. S., GRAY, J., GRUIS, D., BRIGGS, S. P. Convergent insights into mechanisms determining disease and resistance response in plant-fungal interactions. **Can. J. Bot.**, s. 1, p. S468-S474, 1995.
- JONES, D. A., JONES, J. D. G. The role of leucine-rich repeat proteins in plant defences. **Adv. Bot. Res.**, v. 24, p. 89-167, 1997.
- JONES, D. A., THOMAS, C. M., HAMMOND-KOSACK, K. E., BALINT-KURTI, P. L., JONES, J. D. G. Isolation of the tomato *Cf9* gene for resistance to *Cladosporium fulvum* by transposon tagging. **Science**, v. 266, p. 789-793, 1994.
- KAMOUN, S., HUITEMA, E., VLEESHOUWERS, V. G. A. A. Resistance to oomycetes: a general role for the hypersensitive response? **Trends in Plant Sci.**, v. 4, n. 5, p. 196-200, 1999.
- KEIM, P., SHOEMAKER, R.C., PALMER, R.G. Restriction fragment length polymorphism diversity in soybean. **Theor. Appl. Genet.**, v. 77, p. 786-792, 1989.

- KIRÁLY, Z., HORNOK, L. Molecular aspects of plant-pathogen interactions in relation to novel strategies of breeding for disease resistance. *Acta Phytopath. et Entomol. Hung.*, v. 32, p. 1-28, 1997.
- KNOGGE, W. Fungal infection of plants. *Plant Cell*, v. 8, p. 1.711-1.722, 1996.
- KORNBERG, A., BAKER, T. **DNA Replication**. 2. ed. New York: W H Freeman & Co., 1991. 931 p.
- LAMB, C. J. Plant disease resistance genes in signal perception and transduction. *Cell*, v. 76, p. 419-422, 1994.
- LAWRENCE, G., FINNEGAN, E. J., AYLIFFE, M., ELLIS, J. The  $L^6$  gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. *Plant Cell*, v. 7, p. 1.995-1.206, 1995.
- MACHADO, M. A. **Integração de mapas, mapeamento de QTLs e análise da diversidade genética em soja utilizando marcadores AFLP**. Viçosa, MG: UFV, 1999. 90 p. (Tese D.S.).
- MARTIN, G. B., BROMMONSCHENKEL, S. H., CHUNWONGSE, J., FRARY, A., GANAL, M. W., SPIVEY, R., WU, T., EARLE, E. D., TANKSLEY, S. D. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science*, v. 262, p. 1.432-1.436, 1993.
- MARTIN, G. B., FRARY, A., WU, T., BROMMONSCHENKEL, S. H., CHUNWONGSE, J., EARLE, E. D., TANKSLEY, S. D. A member of tomato *Pto* gene family confers sensibility to fenthion resulting in rapid cell death. *Plant Cell*, v. 6, p. 1.543-1.552, 1994.
- MAUGHAN, P. J., SAGHAI-MAROOF, M. A, BUSS, G. R., HUESTIS, G. M. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near isogenic line analysis. *Theor. Appl. Genet.*, v. 93, p. 392-401, 1996.
- McDOWELL, J. M., DHANDAYDHAM, M., LONG, T. A., AARTS, M. G. M., GOFF, S., HOLUB, E. B., DANGL, J. L. Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the *RPP8* locus of *Arabidopsis*. *Plant Cell*, v. 10, p. 1.861-1.874, 1998.
- MEYERS, B. C., CHIN, D. B., SHEN, K., A., SIVARAMAKRISHNAN, S., LAVELLE, D. O., ZHANG, Z., MILCHELMORE, R. W. The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. *Plant Cell*, v. 10, p. 1.817-1.832, 1998a.
- MEYERS, B. C., SHEN, K., A., ROHANI, P., GAUT, B. S., MILCHELMORE, R. W. Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. *Plant Cell*, v. 11, p. 1.833-1.846, 1998b.
- MINDRINOS, M., KATAGIRI, F., YU, G. -L, AUSUBEL, F. The *Arabidopsis thaliana* disease resistance gene *RPS2* encodes a protein contain a nucleotide-binding site and leucine rich repeats. *Cell*, v. 78, p. 1.089-1.099, 1994.

- ORI, N., ESHED, Y., PARAN, I., GERNOT, P., AVIV, D., TANKSLEY, S., ZAMIR, D., FLUHR, R. The *I2C* family from the wilt disease resistance locus *I2* belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. **Plant Cell**, v. 9, p. 521-532, 1997.
- OUSBOURN, A. E. Preformed antimicrobial compounds and plant defence against fungal attack. **Plant Cell**, v. 8, p. 1.821-1.831, 1996.
- PARKER, J. E., COLEMAN, M. J., SZABÒ, V., FROST, L. N., SCHMIDT, R., van der BIENZEN, E. A., MOORES, T. DEAN, C., DANIELS, M. J., J. D. G. The *Arabidopsis* downy mildew resistance gene *RPP5* shares similarity to the Toll and interleukin-1 receptors with *N* and *L*<sup>6</sup>. **Plant Cell**, v. 9, p. 879-894, 1997.
- PARNISKE, M., HAMMOND-KOSACK, K. E., GOLSTEIN, C., THOMAS, C. M., JONES, D. A., HARRISON, K., WULFF, B. B. H., JONES, J. D. G. Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. **Cell**, v. 91, p. 821-832, 1997.
- PIETERSE, C. M. J., LOON, L. C. v. Salicylic acid-independent plant defence pathways. **Trends in Plant Sci.**, v. 4, n. 2, p. 52-58, 1999.
- POWELL, W., MORGANTE, M., DOYLE, J.J., McNICOL, J.W., TINGEY, S.V., RAFALSKI, A.J. Genepool variation in genus *Glycine* subgenus *soja* revealed by polymorphic nuclear and chloroplast microsatellites. **Genetics**, v. 144, p. 793-803, 1996.
- RAFALSKI, J.A., TINGEY, S.V.. RFLP map of soybean (*Glycine max*) 2N=40. In: O'BRIEN, S. J.(Ed.) **Genetic maps**. 6. ed. New York: Cold Spring Harbor Laboratory, Cold Spring Harbor, 1993. p. 149-156.
- ROMMENS, C. M. T., SALMERON, J. M., BAULCOMBE, D. C., STASKAWICZ, B. J. Use of gene expression system based on potato virus X to rapidly identify and characterize a tomato *Pto* homolog that's controls fenthion sensitivity. **Plant Cell**, v. 7, p. 1.537-1.433, 1995.
- SALMERON, J. M., BARKER, S. J., CARLAND, F. M., METHA, A. Y., STASKAWICZ, B. J. Tomato mutants altered in bacterial disease resistance provide evidence for a new locus controlling pathogen recognition. **Plant Cell**, v. 6, p. 511-520, 1994.
- SALMERON, J. M., OLDROY, G. E. D., ROMMENS, C. M. T., SCHOFIELD, S. R., KIM, H. S., LAVELLE, D. T., DAHLBECK, D., STASKAWICZ, F. J. Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance gene and lies embedded with the *Pto* kinase gene cluster. **Cell**, v. 86, p. 123-133, 1996.
- SAMBROOK, J., FRITSCH, E. F., MANIATIS, T. **Molecular cloning: a laboratory manual**. 2. ed. New York: Cold Spring Harbor Laboratory, 1989.
- SANGER, F., NICKLEN, S., COULSON, A. R. DNA Sequence with chain-termination inhibitors. **Proc. Natl. Acad. Sci. USA**, v. 74, p. 5.463-5.467, 1977.

- SCOFIELD, S. R., TOBIAS, C. M., RATHJEN, J. P., CHANG, J. H., LAVELLE, D. T., MILCHEMORE, R. W., STASKAWICZ, B. J. Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. **Science**, v. 274, p. 2.063-2.065, 1996.
- SHAH, D. M., HIGHTOWER, R. C., MEAGHER, R. B. Complete nucleotide sequence of a soybean actin gene. **Proc. Natl. Acad. Sci. USA**, v. 79, p. 1.022-1.026, 1982.
- SONG, W., -Y., PI, L.-Y., WANG, G.-L., GARDNER, J., HOLSTEN, T., RONALD, P.C. Evolution of the rice *Xa21* disease resistance gene family. **Plant Cell**, v. 9, p. 1.279-1.287, 1997.
- SONG, W.-Y., WANG, G.-L., CHEN, L.-L., KIM, H.-S., PI, L.-Y., HOLSTEN, T., GARDNER, J., WANG, B., ZHAI, W.-X., ZHU, L.-H., FAUQUET, C., RONALD, P. C. A receptor kinase-like protein encoded by rice disease resistance gene, *Xa21*. **Science**, v. 270, p. 1.804-1.806, 1995.
- STASKAWICZ, B. J., AUSEBEL, F. M., BAKER, B. J., ELLIS, J. G., JONES, D. G. Molecular genetics of plant disease resistance. **Science**, v. 268, p. 661-667, 1995.
- TANG, S., FREDERICK, R. D., ZHOU, J., HALTERMAN, D. A., JIA, Y., MARTIN, G. B. Initiation of plant disease resistance by physical interaction of *AvrPto* and *Pto* kinase. **Science**, v. 274, p. 2.060-2.062, 1996.
- THOMAS, C. M., JONES, D. A., PARNISKE, M., HARRISON, K., BALINT-KUTI, P. J., HATZIXANTHIS, K., JONES, J. D. G. 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**, v. 9, p. 2.209-2.224, 1997.
- VOS, P., SIMONS, G., JESSE, T., WIJBRANDI, J., HEIN, L., HOGERS, R., FRIJTERS, A., GROENENDIJK, J., DIERGAARDE, P. REIJANS, M., FIERENS-ONSTENK, J., BOTH, M., PELEMAN, J., LIHARSKA, T., HONTELEZ, J., ZABEAU, M. The tomato *Mi-1* gene confers resistance to both root-knot nematodes and potato aphids. **Nature Biotech.**, v. 16, p. 1.365-1.369, 1998.
- WADSWORTH, G. J., REDINHBAUGH, M. G., SCANDALIOS, J. G. A procedure for the small-scale isolation of plant RNA suitable for RNA blot analysis. **Anal. Bioch.**, v. 172, p. 297-283, 1988.
- WANG, G.-L., RUAN, D.-L., SONG, W.-Y., SIDERIS, S., CHEN, L., PI, L., -Y., ZHANG, S., ZHANG, Z., FAUQUET, C., GAUT, B. S., WHALEN, M. C., RONALD, P. C. *Xa21D* encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. **Plant Cell**, v. 10, p. 765-779, 1998.
- WHITHAM, S., DINESH-KUMAR, S. P., CHOI, D., HEHL, R., CORR, C., BAKER, B. The product of the tobacco mosaic virus resistance gene *N*: similarity to toll and interleucine-1 receptor. **Cell**, v. 78, p. 1.101-1.115, 1994.
- ZHOU, J., LOH, Y. T, BRESSAN, R. A., MARTIN, G. B. The tomato gene *Pti1* encodes a serine/threonine kinase that is phosphorylated by *Pto* and is involved in the hypersensitive response. **Cell**, v. 83, p. 925-935, 1995.