

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**AVALIAÇÃO DA BASE GENÉTICA DE GENÓTIPOS DE SOJA
RESISTENTES AO NEMATÓIDE DE CISTO (RAÇA 3) POR
MARCADORES MICROSSATÉLITES**

Daniela Garcia Penha Sarti

Bióloga

JABOTICABAL – SÃO PAULO - BRASIL

2007

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**AVALIAÇÃO DA BASE GENÉTICA DE GENÓTIPOS DE SOJA
RESISTENTES AO NEMATÓIDE DE CISTO (RAÇA 3) POR
MARCADORES MICROSSATÉLITES**

Daniela Garcia Penha Sarti

Orientador: Prof. Dr. Antonio Orlando Di Mauro

Co-orientadores: Dra. Sandra Helena Unêda-Trevisoli

Prof. Dr. José Baldin Pinheiro

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia (Genética e Melhoramento de Plantas).

Jaboticabal - SP

Julho de 2007

DADOS CURRICULARES DA AUTORA

DANIELA GARCIA PENHA SARTI – Nasceu em 08 de novembro de 1981, no município de Orlândia- São Paulo. Em março de 2003 concluiu a graduação em Ciências Biológicas, modalidade Licenciatura, pela Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas de São José do Rio Preto – São Paulo. Em julho de 2005, obteve o título de Bacharel em Ciências Biológicas pela Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas de São José do Rio Preto – São Paulo, realizando estágio curricular para Monografia no Programa de Melhoramento Genético da Soja do Departamento de Produção Vegetal da Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal – São Paulo. Obteve o Título de Mestre em Agronomia (Genética e Melhoramento de Plantas) em 29 de junho de 2007, pela Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal – São Paulo.

“Deus não está afastado no espaço incomensurável e desconhecido do homem, mas, imanente na própria Natureza, como uma Luz eterna e transcendente no processo evolutivo e criativo de tudo que existe no Universo.”

Hammed

Aos meus avós

Joaquina e Celestino

Elza e José Anésio (in memoriam)

À nossa prima

Miriam (in memoriam)

DEDICO

Aos meus pais

Tânia e Zito

Ao meu irmão

Bruno

OFEREÇO

AGRADECIMENTOS

À Deus,

À Universidade Estadual Paulista - UNESP/FCAV - Câmpus de Jaboticabal, pela oportunidade oferecida de desenvolver este trabalho.

À CAPES, pelo apoio financeiro no projeto.

Ao Prof. Dr. Antônio Orlando Di Mauro, pela orientação e confiança.

À Dra. Sandra Helena Unêda-Trevisoli, pela co-orientação, pela amizade, atenção e ensinamentos proporcionados.

Ao Prof. Dr. José Baldin Pinheiro pela co-orientação, por valiosas contribuições e trocas de informações.

A todos os professores da FCAV/UNESP pelo companheirismo e grande sabedoria, em especial aos Professores Alberto Cargnelutti Filho, Maria Aparecida Pessôa da Cruz Centurion, João Carlos de Oliveira, Domingos Fornasieri Filho, Rinaldo César de Paula, Dilermando Perecin, Rubens Sader e Manoel Victor Franco Lemos.

Aos amigos-irmãos Daniela Abreu, Marcelo Marchi, Cláudia Demétrio e Melina Mancini, pelo apoio, conselhos e, acima de tudo, amizade verdadeira.

Aos amigos Ivana Bárbaro, Lizandra Catelli, Franco Muniz, Gustavo Silveira, Antonio Morceli, Thaiza Morceli e Nair Arriel, pela amizade e incentivo em uma nova fase em minha vida, de aprendizado e crescimento.

Ao Júlio Quierati, pelo bom humor e companheirismo.

Ao Wellington Zanni e todos os estagiários e amigos de Pós-Graduação, pela ajuda na condução dos experimentos, pela confiança e apoio.

Aos funcionários do Departamento de Produção Vegetal da FCAV/UNESP: Mauro, Nice, Luís, Rubens, Osmar, Sebastião, Gabi, Marisa, Mariângela e, especialmente Geraldo, por além da amizade, o auxílio fundamental na condução do experimento.

Aos queridos amigos do IBILCE, em especial à Joana, Flávia, Cristiane, Ana Flávia, Sabrina, Marquito, Otávio, Gustavo e Renato, por todos os momentos inesquecíveis que passamos juntos.

Às amigas de sempre Susana, Roberta, Flávia, Maísa, Ana Clara, Thaís e Mariela, pelo incentivo, carinho e por fazerem parte de todas as minhas boas lembranças.

A toda minha família, meus pais Tânia e Zito, meu irmão Bruno, vó Elza, Ângela, meus adorados vô Celé e vó Quininha, pela força e confiança, me ajudando a ultrapassar os obstáculos e viver sempre com otimismo.

A todos que direta ou indiretamente contribuíram para realização deste trabalho.

SUMÁRIO

	Página
LISTA DE TABELAS.....	ix
LISTA DE FIGURAS.....	xi
RESUMO.....	xiv
SUMMARY.....	xvi
1. INTRODUÇÃO.....	1
2. REVISÃO DE LITERATURA.....	4
2.1. A Soja.....	4
2.2. Importância Econômica.....	5
2.3. Nematóide de Cisto da Soja.....	6
2.4. Genética da Resistência da Soja ao Nematóide de Cisto.....	8
2.5. Marcadores Moleculares.....	11
2.6. Divergência Genética.....	14
2.6.1. Análise de agrupamento.....	17
3. MATERIAL E MÉTODOS.....	21
3.1. Local de Condução do Experimento.....	21

3.2. Identificação e Coleta do Material.....	21
3.3. Atividades em Laboratório.....	23
3.3.1. Extração de DNA.....	23
3.3.2. Identificação de marcadores SSR.....	24
3.3.3. Reações de PCR e visualização dos resultados.....	24
3.3.4. Montagem da matriz de dissimilaridade.....	26
3.4. Análises Estatísticas.....	27
4. RESULTADOS E DISCUSSÃO.....	29
4.1. Coeficientes de Similaridade por Coincidência Simples.....	32
4.2. Coeficientes de Similariadade por Rogers e Tanimoto.....	50
4.3. Correlação entre Matrizes.....	70
4.4. Considerações Finais.....	70
5. CONCLUSÕES.....	73
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	74

LISTA DE TABELAS

	Página
Tabela 1. Cultivares e genótipos F8 de soja resistentes ao NCS, utilizados no estudo de divergência genética, com suas respectivas fontes de resistência.....	22
Tabela 2. Grupo de ligação, seqüência e número de alelos encontrados para os 12 marcadores SSR utilizados nas análises dos 40 genótipos de soja.....	25
Tabela 3. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Coincidência Simples.....	33
Tabela 4. Formação de grupos pelo método de otimização de Tocher, utilizando a dissimilaridade através do complemento aritmético do Coeficiente de Coincidência Simples dos 40 genótipos de soja resistentes ao NCS	37
Tabela 5. Dez genótipos mais similares, em ordem decrescente, em relação a cada um dos 40 genótipos estudados com fonte de resistência ao nematóide de cisto da soja, através do complemento aritmético do Coeficiente de Similaridade de Coincidência Simples.....	51
Tabela 6. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Similaridade de Rogers e Tanimoto.....	52
Tabela 7. Formação de grupos através do método de otimização de Tocher, utilizando complemento aritmético do Coeficiente de Rogers e	

Tanimoto dos 40 genótipos de soja resistentes ao NCS estudados.....	57
---	----

Tabela 8. Dez genótipos mais similares, em ordem decrescente, em relação a cada um dos 40 genótipos estudados com fonte de resistência ao nematóide de cisto da soja, através do complemento aritmético do Coeficiente de Rogers e Tanimoto.....	69
---	----

LISTA DE FIGURAS

	Página
Figura 1. Amostras dos fragmentos produzidos nas reações de amplificação dos 40 genótipos de soja utilizando o iniciador Satt308.....	29
Figura 2. Amostras dos fragmentos produzidos nas reações de amplificação dos 40 genótipos de soja utilizando o iniciador Satt152.....	30
Figura 3. Amostras dos fragmentos produzidos nas reações de amplificação dos 40 genótipos de soja utilizando o iniciador Satt154.....	31
Figura 4. Representação do dendrograma obtido através do Método de Ligação Simples – Vizinho Mais Próximo, a partir do complemento aritmético do Coeficiente de Coincidência Simples.....	41
Figura 5. Representação do dendrograma obtido através do Método de Ligação Completa – Vizinho Mais Distante, a partir do complemento aritmético do Coeficiente de Coincidência Simples.....	42
Figura 6. Representação do dendrograma obtido através do Método de Ward, a partir do complemento aritmético do Coeficiente de Coincidência Simples.....	43
Figura 7. Representação do dendrograma obtido através do Método de Ligação Média dentro do grupo, a partir do complemento aritmético do Coeficiente de Coincidência Simples.....	44
Figura 8. Representação do dendrograma obtido através do Método	

UPGMA, a partir do complemento aritmético do Coeficiente de Coincidência Simples.....	45
Figura 9. Representação do dendrograma obtido através do Método da Mediana - WPGMA, a partir do complemento aritmético do Coeficiente de Coincidência Simples.....	46
Figura 10. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Coincidência Simples, no espaço bidimensional.....	47
Figura 11. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Coincidência Simples, no espaço tridimensional (3D com rotação).....	48
Figura 12. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Coincidência Simples, no espaço tridimensional (3D sem rotação).....	49
Figura 13. Representação do dendrograma obtido através do Método de Ligação Simples – Vizinho Mais Próximo, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.....	60
Figura 14. Representação do dendrograma obtido através do Método de Ligação Completa - Vizinho mais Distante, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.....	61
Figura 15. Representação do dendrograma obtido através do Método de Ward, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.....	62

Figura 16. Representação do dendrograma obtido através do Método de Ligação Média Dentro de Grupo, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.....	63
Figura 17. Representação do dendrograma obtido através do Método de Agrupamento UPGMA, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.....	64
Figura 18. Representação do dendrograma obtido pelo Método da Mediana - WPGMA, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.....	65
Figura 19. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Rogers e Tanimoto, no espaço bidimensional.....	66
Figura 20. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Rogers e Tanimoto, no espaço tridimensional (3D com rotação).....	67
Figura 21. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Rogers e Tanimoto, no espaço tridimensional (3D sem rotação).....	68
Figura 22. Correlação entre as matrizes de dissimilaridade de Rogers e Tanimoto e de Coincidência Simples.....	71

AVALIAÇÃO DA BASE GENÉTICA DE GENÓTIPOS DE SOJA RESISTENTES AO NEMATÓIDE DE CISTO (RAÇA 3) POR MARCADORES MICROSSATÉLITES

RESUMO - O presente trabalho objetivou a avaliação da divergência genética entre linhagens e cultivares de soja com resistência ao nematóide de cisto (NCS) através de marcadores moleculares do tipo microssatélite, a fim de identificar genótipos similares e grupos heterogêneos. Foram utilizadas 30 cultivares de soja e 10 genótipos em geração F8, todos resistentes ao NCS. As análises foram realizadas com 12 iniciadores, a saber: Satt185, Satt308, Satt187, Satt070, Satt009, Satt129, Satt154, Satt152, Satt175, Satt045, Satt162 e Satt163. As análises de dissimilaridade foram realizadas através do complemento dos coeficientes de Coincidência Simples e de Rogers e Tanimoto. Os agrupamentos foram realizados pelos Métodos de Otimização de Tocher, Vizinho mais Próximo, Vizinho mais Distante, Ward, Ligação Média Dentro de Grupo, UPGMA e WPGMA. Em adição, foram feitas as projeções no plano bi e tridimensional. Os resultados mostraram coeficientes de dissimilaridade de Coincidência Simples e de Rogers e Tanimoto com médias de 0,36 e 0,5, respectivamente. Foram observados genótipos mais semelhantes e grupos mais divergentes, na maioria das vezes concordando com as respectivas fontes de resistência. Os resultados dos diferentes métodos aglomerativos confirmaram a existência de uma estrutura natural de agrupamento em função da similaridade genética entre os genótipos. Entretanto, o emprego de mais de um método de agrupamento pode ser recomendado devido às diferenças entre as metodologias. Alguns métodos mais comumente utilizados como o Vizinho Mais Distante, Ward, WPGMA e UPGMA foram considerados adequados ao agrupamento de genótipos no presente estudo, apresentando coincidência de resultados, inclusive em relação às genealogias dos genótipos analisados. As projeções no plano não mostraram resultados

satisfatórios, com elevada distorção e estresse. A combinação de resultados pode ser útil na escolha de parentais de soja para cruzamentos com fonte de resistência ao NCS, dando continuidade ao Programa de Melhoramento da Soja.

EVALUATION OF GENETIC BASIS OF SOYBEAN GENOTYPES RESISTANT TO CYST NEMATODE (RACE 3) BY MICROSATELLITE MARKERS

SUMMARY – The present work aimed the evaluation of the genetic divergence among soybean lines and cultivars, resistant to cyst nematode (NCS). Therefore microsatellite molecular markers were used, with the objective of identify similar genotypes and heterogeneous groups. There were used 30 soybean cultivars and 10 genotypes in F8 generation, all of them resistant to NCS. The analyses were made using 20 primers (Satt185, Satt308, Satt187, Satt070, Satt009, Satt129, Satt154, Satt152, Satt175, Satt045, Satt162 and Satt163). The dissimilarity analyses were made through the complement of Simple Coincidence and Rogers and Tanimoto Coefficients. Clusters were performed by the following methods: Tocher Optimization, Single Linkage, Complete Linkage, Ward, Average Link Within a Group, UPGMA and WPGMA. In addition, the bi and tridimensional plan projections were made. The results showed average values of 0,36 and 0,5 in the Simple Coincidence and Rogers and Tanimoto dissimilarity coefficients, respectively. Most similar genotypes and most divergent groups were found, usually according to the respective resistant sources. Results of the different cluster methods confirmed a group's structure according to genotypes similarity. However, the use of many clustering methods should be recommended, because of their methodology differences. Some of the most used methods, like Complete Linkage, Ward, WPGMA and UPGMA, were considered adequate for clustering the genotypes in this case. There were observed coincidence of results confirming the genotypes genealogies. The plan projections were not satisfactory for these analyses, showing high distortion and stress. The results combination can be useful in the parental choice for crosses of Soybean Breeding Programs.

1. INTRODUÇÃO

A soja (*Glycine max* (L.) Merrill) é uma espécie vegetal originária de clima temperado, com ampla adaptação nos climas subtropicais e tropicais. É considerada uma das mais importantes leguminosas, que pode crescer em quase todas as latitudes nas quais as sementes são cultivadas, sendo grande geradora de divisas aos países produtores (BLACK, 2000).

Segundo a Companhia Nacional de Abastecimento (CONAB), os maiores produtores e exportadores de soja no mundo são os Estados Unidos com cerca de 86 milhões de toneladas, vindo a seguir o Brasil com 58 milhões de toneladas na safra 2006/07 (CONAB, 2007). A produção mundial concentra-se entre Estados Unidos, Brasil, Argentina e China, que respondem por cerca de 90% da produção total (BLACK, 2000). A produtividade média das principais regiões produtoras, no Brasil, está estimada em 2809 kg/ha, com algumas regiões do centro-oeste brasileiro podendo ultrapassar 3000 kg/ha (CONAB, 2007).

Entretanto, existem fatores que limitam a obtenção de maiores rendimentos da soja, entre eles estão mais de 100 tipos de doenças e, dentre essas, ao menos 35 são de considerável importância econômica (YORINORI, 1994).

O nematóide de cisto da soja (NCS), *Heterodera glycines*, representa um dos principais problemas da soja no mundo. Tal cultura é hospedeira preferencial deste patógeno, que possui sua grande severidade aliada ao número de raças existentes e grande potencial para surgimento de outras. No Brasil, na safra 1999/2000, as doenças da soja foram responsáveis por prejuízos estimados em U\$1,39 bilhão. Desse montante, o NCS contribuiu com U\$133,2 milhões, com área infestada chegando a 22% da área total de soja no país (BORÉM, 1997; YORINORI, 2000, 2002).

Os programas de melhoramento genético da soja estão voltados à criação de cultivares mais produtivas, mais estáveis e mais bem adaptadas às diversas regiões ecológicas e aos vários sistemas de cultivo existentes no Brasil, incluindo a resistência às possíveis doenças e pragas presentes.

Embora o número de cultivares disponíveis no Brasil seja grande, a base genética do germoplasma de soja brasileiro está extremamente reduzida, aumentando, dessa forma, a vulnerabilidade para perdas com doenças e pragas (BECELAERE, et al. 2005). BONETTI (1983) estimou que cerca de 70% das cultivares desenvolvidas para o Rio Grande do Sul, naquela data descendiam das cultivares americanas Hill, Hood ou ambas. HIROMOTO & VELLO (1986) informaram que todas as cultivares recomendadas para plantio naquele ano agrícola descendiam de 26 cultivares, sendo que deste total, apenas quatro eram responsáveis por cerca da metade daquele conjunto gênico.

Dessa forma estudos de divergência genética e relações filogenéticas entre espécies vegetais de importância econômica têm merecido atenção cada vez maior, uma vez que provê informações sobre recursos genéticos disponíveis e auxilia na localização e no intercâmbio desses (CRUZ et al., 2004). Diversas técnicas de biologia molecular estão disponíveis atualmente para detecção de polimorfismo genético diretamente ao nível de DNA, auxiliando na avaliação das relações genéticas entre genótipos. Os marcadores moleculares fornecem um alto grau de polimorfismo distribuído aleatoriamente ao longo de todo o genoma e, por serem independentes de efeitos ambientais e estádios fisiológicos das plantas, permitem uma identificação precoce e precisa dos genótipos de interesse.

Os marcadores moleculares têm sido amplamente utilizados na caracterização da diversidade genética entre e dentro de genótipos de uma população, geralmente por meio da similaridade ou dissimilaridade genética. Nesses trabalhos, os pesquisadores têm interesse em agrupar os indivíduos semelhantes, de forma que as maiores diferenças ocorram entre os grupos formados a fim de caracterizar o germoplasma e, ainda, na escolha de parentais mais divergentes para realização de cruzamentos no melhoramento de plantas (CORRÊA et al., 1999; MUHLEN et al., 2000; PRIOLLI et al., 2002; MEYER et al., 2004; ARRIEL et al., 2006).

Dentre os marcadores moleculares disponíveis, os microssatélites, ou SSR ("Simple Sequence Repeat") têm sido considerados ideais para a caracterização e avaliação da variabilidade genética, pois são codominantes, multialélicos, reprodutíveis

e amplificados via reação de polimerase em cadeia (PCR – “Polymerase Chain Reaction” - MULLIS & FALOONA, 1987). Com o uso desses marcadores é possível identificar o padrão molecular de genótipos com potencial comercial, facilitando estudos futuros.

Métodos estatísticos de análise podem ser aplicados para auxiliar em estudos de divergência genética. As matrizes de similaridade, ou distâncias entre genótipos podem ser calculadas de diversas formas, sendo que diferentes propostas são encontradas na literatura (SNEATH & SOKAL, 1973; JOHNSON & WICHERN, 1988; WEIR, 1996). Considerando que os resultados dos agrupamentos podem ser influenciados pela escolha do coeficiente de similaridade (GOWER & LEGENDRE, 1986; JACKSON et al., 1989; DUARTE et al., 1999), estes coeficientes precisam ser mais bem estudados, de forma que os mais eficientes em cada situação específica possam ser empregados (DUARTE et al., 1999).

Visto que a escolha de genitores e o planejamento dos cruzamentos são etapas de fundamental importância para o sucesso de um programa de melhoramento, a avaliação da diversidade genética em soja resistente ao NCS se torna útil no âmbito de fornecer um padrão de bandas moleculares que fornecerão aos melhoristas informações genéticas adicionais e mais detalhadas dos genótipos, aumentando a probabilidade de obtenção de cultivares superiores a serem lançados no mercado nacional. Dessa forma, o presente trabalho objetivou:

- Estimar coeficientes de similaridade entre genótipos e cultivares com resistência ao nematóide de cisto da soja (NCS) através de marcadores moleculares do tipo microssatélite;
- A obtenção de dendrogramas e agrupamentos dos genótipos estudados através de diferentes critérios aglomerativos (hierárquicos e otimização) e projeções das distâncias no plano, a fim de identificar genótipos similares e grupos heterogêneos;
- Comparar diversos métodos de agrupamentos, a partir de dados moleculares, identificando o mais adequado para o estudo de divergência genética destes genótipos.

2. REVISÃO DE LITERATURA

2.1. A Soja

A soja é uma leguminosa anual de caule ereto, com grande diversidade quanto ao ciclo de vida. É uma planta milenar, sendo difícil estabelecer com clareza sua origem e história. O centro de origem do grão parece estar localizado nas partes central e ocidental da China (SLEPER & SHANNON, 2003), sendo a Manchúria um importante centro de diversificação (origem secundária) (CARUSO, 1997). HYMOWITZ (1970), com base em evidências históricas, geográficas e achados arqueológicos, concluiu que a soja foi domesticada por volta do século XI a.C. Há 3000 anos a soja se espalhou pela Ásia, onde começou a ser utilizada como alimento (CARUSO, 1997).

Na Europa esta cultura tornou-se conhecida no século XVII, quando pela primeira vez foi plantada no Jardim Botânico de Paris. Em 1765, foi introduzida no Norte da América e Sul dos Estados Unidos, sendo no último, muito utilizada para feno e forragem verde (COSTA & MANICA, 1996).

No Brasil, a soja foi introduzida na Bahia, em 1882, por Gustavo D'Utra. Da Bahia foi levada para São Paulo onde se obteve os primeiros resultados experimentais, relatados em 1892 pelo Instituto Agrônomo de Campinas (IAC). No Rio Grande do Sul, a soja foi cultivada em 1900, mas só no final dos anos 40 é que o cultivo comercial iniciou-se no Rio Grande do Sul e em São Paulo (COSTA & MANICA, 1996).

Na década de 60, a cultura se expandiu inicialmente, em latitudes entre 30°S e 20°S, principalmente nos Estados do Rio Grande do Sul, Santa Catarina, Paraná e São Paulo, e com a identificação de cultivares mais bem adaptadas às condições ambientais, os melhoristas passaram a combinar as características destas por meio de hibridação (BONETTI, 1983).

A expansão da cultura para médias e baixas latitudes foi possível pelo desenvolvimento de cultivares locais, onde a estratégia de obtenção consistiu no

desenvolvimento de plantas com crescimento determinado, semelhante às utilizadas no sul dos Estados Unidos, e com altura e ciclo adequados a estas condições (KIIHL et al., 2000).

Por ser uma das espécies mais cultivadas no Brasil a partir dos anos 70, a soja ocupa posição de destaque na economia agrícola do país. No período entre 1986 e 2003, a Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), em parceria com outras empresas, listaram mais de 200 cultivares recomendadas para plantio na maioria das regiões sojícolas do Brasil, representando uma maior produção ano após ano em função do desenvolvimento de novas cultivares com resistência a fatores bióticos e abióticos (EMBRAPA, 2003).

2.2. Importância Econômica

A soja tem apresentado, nas últimas cinco décadas, uma taxa de crescimento superior à taxa de crescimento populacional, ocupando lugar de destaque na alimentação humana e animal dos cinco continentes. Atualmente é uma das mais importantes oleaginosas cultivadas no mundo.

Levantamentos da Companhia Nacional de Abastecimento do Ministério da Agricultura mostram que de toda área cultivada com grãos no país, a soja ocupa 44,8%, ou 20,65 milhões de hectares. A cultura também avança com muita força em novas fronteiras agrícolas, o que estabelece perspectivas de continuidade no crescimento verificado nos últimos anos. Deve-se ressaltar ainda a importância da soja no mercado externo, onde as exportações do complexo agroindustrial do grão representam em torno de 24% das exportações do agronegócio brasileiro, e em torno de 10% de nossas exportações totais, segundo estatísticas da CONAB (2007).

Em 2004, o Brasil produziu cerca de 50 milhões de toneladas de soja, ou 25% da safra mundial, montante menor que o de 2003, quando o País produziu 52 milhões de toneladas e participou com quase 27% da safra mundial, estimada em cerca de 200

milhões de toneladas para 2004. Calcula-se que, aproximadamente 10 milhões de toneladas, ou 20% da safra brasileira de 2004, tenham sido perdidos devido às condições climáticas que não foram favoráveis para o desenvolvimento da cultura e pela falta de controle da ferrugem asiática. Na Região Sul a perda ocorreu pela estiagem e, na Região Centro-Oeste pelo excesso de chuvas, impedindo o desempenho sustentável da produção (EMBRAPA, 2004).

No Brasil, a cadeia agroindustrial da soja, gera mais de cinco milhões de empregos e participa com pelo menos 16% dos mais de 40% do Produto Interno Bruto (PIB) brasileiro gerado pela agroindústria (EMBRAPA, 2004). Considerando o PIB de 2005, conforme dados do IBGE (Instituto Brasileiro de Geografia e Estatística), isso significa um montante de cerca de R\$ 34 bilhões anuais de ganhos com a cultura. Esses números mostram a importância econômica da soja para o país. Além disso, vários setores da sociedade têm a renda diretamente relacionada aos empregos gerados pelo complexo de produção, transporte e comercialização da soja.

O processamento da soja na indústria dá origem, inicialmente, a dois produtos de grande consumo: o farelo e o óleo. O farelo é utilizado principalmente na elaboração de rações, sendo a principal fonte de proteína na nutrição animal. Dos óleos comestíveis consumidos no Brasil, o consumo do óleo de soja é o mais expressivo, correspondendo a mais de 90%. A partir deles podem ser feitos dezenas de outros produtos (BLACK, 2000). Na safra 2006/2007, o Brasil produziu 22,8 milhões de toneladas de farelo de soja e 5,6 milhões de toneladas de óleo de soja (CONAB, 2007).

2.3. Nematóide de Cisto da Soja

O nematóide de cisto da soja (*Heterodera glycines* Ichinohe) é uma das principais pragas da cultura. Ele penetra nas raízes das plantas de soja ainda jovens, dificultando a absorção de água e nutrientes. Os sintomas na lavoura se caracterizam pela ocorrência de manchas em reboleira, onde as plantas se apresentam cloróticas,

com redução do porte e do número de vagens e muitas vezes morrem prematuramente. O sistema radicular das plantas afetadas fica reduzido e apresenta minúsculas fêmeas do nematóide, com formato semelhante a um limão. Inicialmente de coloração branca, a fêmea, posteriormente adquire a coloração amarelada. Após ser fertilizada pelo macho, cada fêmea produz de 100 a 250 ovos, armazenando a maior parte delas em seu corpo. Quando a fêmea morre, seu corpo se transforma em uma estrutura dura, de coloração marrom escuro, cheia de ovos, altamente resistente à deterioração e à dessecação e muito leve, denominada cisto, que se desprende da raiz e permanece no solo (YORINORI, 1994; YORINORI, 1998; EMBRAPA, 1999).

O ciclo de vida varia de 21 a 25 dias, sendo muito influenciado pela temperatura, entre 23°C e 25°C. Dessa forma, é possível até seis gerações do nematóide em um ciclo de cultivares tardias de soja. O cisto pode sobreviver no solo, na ausência de planta hospedeira, por mais de oito anos, sendo quase impossível eliminar o parasita nas áreas onde ele ocorre (ALMEIDA et al., 1997).

Os ovos no interior do cisto, já fertilizados, sofrem embriogênese, dando origem ao juvenil de primeiro estágio (J1). Esse tem sua ecdise, ou troca de cutícula, dentro do ovo e torna-se o juvenil de segundo estágio (J2), que eclode, migra no solo e invade as raízes da planta hospedeira. Após a penetração, o J2 induz a modificação de um conjunto de células da soja no local de penetração, estabelecendo o sítio de alimentação, os sincitos, que passam a fornecer alimento para o nematóide. O J2 continua a se desenvolver, sofre mais três ecdises e, finalmente, atinge a fase adulta, de macho ou fêmea. As fêmeas aumentam de volume, assumem o formato de limão, de coloração branca amarelada e permanecem fixadas à raiz, com o corpo do lado de fora e a parte anterior internamente nos tecidos radiculares. Os machos têm corpo alongado, passam para o solo e, após a fertilização das fêmeas, morrem (TAYLOR, 1971; SCHMITT & BARKER, 1985).

A fecundação cruzada desta espécie de nematóide é a razão da sua elevada variabilidade genética, determinando a existência de raças. No Brasil a raça predominante é a 3, embora as raças 1, 2, 4, 5, 6, 10, 14, 4⁺ e 14⁺ também já tenham sido identificadas. As raças 4⁺ e 14⁺ possuem genes adicionais de parasitismo, o que

possibilita a quebra da resistência da cultivar Hartwig, anteriormente resistente a todas as raças do nematóide de cisto (DIAS et al., 1999).

Constatado pela primeira vez no Brasil na safra 1991/92, na região do cerrado, rapidamente atingiu municípios localizados em Mato Grosso, Mato Grosso do Sul, Goiás, Minas Gerais, São Paulo e Rio Grande do Sul. (EMBRAPA, 1999).

A cada safra, verifica-se uma disseminação elevada desse nematóide, principalmente através da intensa movimentação de máquinas, implementos agrícolas e veículos, que carregam solo contaminado aderido. O nematóide também pode ser disseminado pelo vento, enxurrada, animais e sementes mal beneficiadas, contendo partículas de solo. Além do uso de genótipos resistentes, recomenda-se a semeadura direta como forma de cultivo, dificultando a dispersão dos cistos (YORINORI, 1998).

O método mais eficiente de controle desse patógeno é o uso de variedades resistentes, acompanhados da rotação com culturas não hospedeiras, além da prevenção de entrada do NCS em áreas ainda isentas do mesmo. O planejamento da rotação de culturas é relativamente simples, em função da limitada gama de hospedeiros do nematóide. O milho, a cana-de-açúcar, o algodão, o sorgo, o arroz, o girassol, a mamona, o trigo, assim como as demais pastagens, podem ser cultivadas em áreas infestadas. O NCS não se reproduz nas plantas daninhas mais comuns nas lavouras de soja, no Brasil (ALMEIDA ET AL, 1997; YORINORI, 1998; EMBRAPA, 1999).

2.4. Genética da Resistência da Soja ao Nematóide de Cisto

Estudos da base genética das cultivares brasileiras de soja revelaram que o germoplasma utilizado em programas de melhoramento genético, no país como um todo, pode ser considerado como uma população de baixo tamanho efetivo ($N_e = 11$ a 15) (VELLO et al., 1988). Muitos germoplasmas resistentes têm sido identificados e

usados com sucesso em programas de melhoramento. Entretanto, a base genética de cultivares de soja resistente ao NCS é relativamente estreita (YUE et al., 2000).

Durante o ano agrícola 1983/1984, foram traçadas as genealogias de 74 cultivares de soja recomendadas para o cultivo no Brasil, onde se derivou somente 26 ancestrais. Destes, onze contribuíram com 89% do conjunto gênico das cultivares brasileiras de soja; sendo que seis desses ancestrais são os mais freqüentes nas genealogias das cultivares americanas (HIROMOTO & VELLO, 1986). Desse modo, o patógeno pode facilmente superar a resistência dessas cultivares. A coleta e a caracterização genética de novas fontes de resistência têm que ser atividades constantes (DIAS, 2003).

A base genética da resistência em soja ao NCS, raça 3, é complexa e ainda não bem entendida, sendo estudada por diversos autores. Pode estar envolvido um bloco de genes e/ou um ou poucos genes com vários alelos (CALDWELL, 1960; RAO-ARELLI & ANAND, 1988; RAO-ARELLI ET AL., 1992). Uma das hipóteses mais aceita indica a reação de resistência à raça 3 do NCS controlada por um gene dominante e dois genes recessivos (MYERS & ANAND, 1991; RAO-ARELLI et al., 1992; MAURO et al., 1999).

WEBB et al. (1995) mapearam três QTLs nos grupos de ligação A, G e M, que conferem completa resistência à raça 3 do NCS na PI 437654, ancestral resistente de Hartwig. Estudos de YUE et al. (2000) concluíram que a resistência a várias raças do NCS na PI 438489B é controlada por três ou quatro genes de resistência.

Para desenvolver cultivares de soja resistente ao NCS é necessário o conhecimento do número de genes relacionados à resistência, bem como o tipo de ação gênica e a herdabilidade do caráter (MAURO et al., 1999). Muitos genes envolvidos na resistência a raças específicas de NCS já foram identificados. O gene de resistência *Rhg4* foi identificado e mapeado no grupo de ligação A (KEIM et al., 1990; WEBB et al., 1995), em cruzamentos envolvendo a PI 290136 (WEISEMANN et al., 1992) e Peking (MAHALINGAN & SKORUPSKA, 1995). Segundo WEBB et al. (1995), a PI 437654 e Peking têm os mesmos locos para resistência à raça 3.

O mecanismo de resistência é baseado na hipersensibilidade, a qual determina a morte de células infectadas e bloqueia o ciclo de vida do nematóide (SCHIMIDT &

NOEL, 1984). Também é freqüente a ocorrência de ligação entre os alelos de resistência às diferentes raças (HARTWIG & YOUNG, 1986) e entre os locos de resistência (*Rhg4*) e o loco *i*, que confere cor preta ao tegumento da semente da soja (MATSON & WILLIAMS, 1965; YUE et al., 2000). Nessa mesma região, uma família de genes que codificam a síntese de chalconas (CHS) foi encontrada (TODD & VODKIN, 1998). Deleções nessa região aumentam níveis de mRNA de chalcona sintase, que é uma enzima chave envolvida em respostas das plantas a ataque de patógenos e importante na síntese de fitoalexinas e flavonóides (WELLE et al., 1991).

O gene de resistência *rhg1* está localizado no grupo de ligação G e tem demonstrado controlar mais de 50% da variação para resistência, controlando efetivamente várias raças do NCS (WEBB et al., 1995; CONCIBIDO et al., 1996, 1997). MUDGE et al. (1996) sugerem em seus estudos a presença de um gene responsável pela redução da produtividade na mesma região de *rhg1*, o que deve ser levado em consideração na seleção de genótipos, concomitantemente à redução do nível de ligação gênica. PRABHU et al. (1999), em testes da resistência de Hartwig, verificaram ainda, que os índices médios de parasitismo pelo NCS somente foram baixos naqueles genótipos contendo ambos os alelos de resistência (*Rhg4* e *rhg1*).

Das fontes de resistência ao NCS já identificadas, a PI 437654 é a única com resistência a todas as raças (ANAND & GALLO, 1984; ANAND, 1985; RAO-ARELLI et al., 1992). Como essa planta introduzida (PI) tem sementes com tegumentos pretos e apresenta estabilidade e produtividade baixas, foi melhorada através do seu cruzamento com Forrest, sendo produzida Hartwig (ANAND, 1992). Essa cultivar não chegou a ser muito cultivada por ser pouco produtiva, mas, por apresentar resistência a todas as raças do NCS conhecidas, passou a ser utilizada em larga escala nos programas de melhoramento genético da soja.

Segundo HARTWIG (1985) e RAO-ARELLI (1994), a maioria dos materiais de soja resistentes tem genes de Peking (raças 1, 3 e 5), PI88788 (raças 3 e 14) ou de ambas. Porém, o uso contínuo de cultivares resistentes derivadas de Peking resultou no aparecimento de populações do NCS capazes de se multiplicar nesses genótipos (quebra de resistência). Do cruzamento de Peking com Lee (suscetível) resultou Pickett,

a primeira cultivar de soja com semente amarela e resistente às raças 1 e 3 do NCS (BRIM & ROSS, 1966).

No Brasil, são utilizados como fontes de resistência genótipos originados de Peking, PI 88788, PI 90763 ou da PI 437654. Os primeiros resultados desse trabalho possibilitaram o lançamento das cultivares BRSMG Renascença e BRSMG Liderança, resistentes à raça 3, para Minas Gerais; BRSMG Preciosa, resistente à raça 3 para Minas Gerais, São Paulo, Goiás e Distrito Federal; BRSMG Robusta, resistente à raça 3, para Minas Gerais, Goiás, Distrito Federal, Mato Grosso e Bahia; BRSMT Pintado, BRSMT Tucunaré, BRSMT Caxara, BRSMT Matrinxã e BRSMT Piraíba, resistentes às raças 1 e 3 para Mato Grosso; BRSGO Chapadões, resistente às raças 1, 3, 4 e 14 e BRSGO Ipameri, resistentes às raças 3 e 14, para Goiás, dentre outras. Diferentemente do ocorrido nos Estados Unidos, as primeiras cultivares resistentes brasileiras têm alcançado altas produtividades e, conseqüentemente, vêm tendo grande aceitação pelos sojicultores. Na safra 2001/02, cerca de 2 milhões de hectares foram semeados com soja resistente ao NCS (DIAS, 2003).

2.5. Marcadores Moleculares

Com os avanços de técnicas modernas de biologia molecular, surgiram diversos métodos de detecção de polimorfismo genético diretamente ao nível de DNA (FERREIRA & GRATTAPAGLIA, 1998). Nos últimos anos, novas técnicas, que auxiliam na avaliação das relações genéticas entre genótipos, foram desenvolvidas, acompanhando o rápido e constante avanço na área da genética molecular. O desenvolvimento e aplicação de tecnologias baseadas em marcadores moleculares, fornecem ferramentas únicas, capazes de revelar polimorfismos ao nível da sequência do DNA, suficientes para discriminar a variação genética entre indivíduos e dentro de populações (KRESOVICH et al., 1995).

No melhoramento genético de plantas, tem-se buscado cada vez mais a manipulação assistida por marcadores moleculares, visando obter maior eficiência na transferência de fatores genéticos. Marcadores moleculares ligados a diferentes características de importância econômica têm sido desenvolvidos, permitindo a seleção indireta de caracteres desejáveis em gerações segregantes precoces. Esta estratégia reduziria tempo, recursos e energia, necessários não só para desenvolver grandes populações segregantes por várias gerações, mas também para estimar parâmetros usados na seleção indireta (CAIXETA et al., 2006).

Os marcadores moleculares também podem auxiliar em estudos de divergência genética. Estes consistem em métodos bioquímicos que identificam diferenças diretamente no material genético e podem revelar, para uma dada região do DNA, uma banda ou marca que permite comparar os indivíduos em estudo quanto à sua presença, ou ausência. As bandas reveladas são codificadas pelo número 1 e as não reveladas pelo número 0. Dessa forma os dados aos quais serão aplicados métodos estatísticos provêm de uma matriz binária, constituída de uns e zeros (MEYER, 2002).

Os primeiros marcadores de DNA a serem utilizados foram os fragmentos produzidos pela digestão do DNA com enzimas de restrição. A variação do tamanho dos fragmentos obtidos de diferentes indivíduos, após a digestão com as enzimas de restrição, deu origem à classe de marcadores denominados RFLP ("Restriction Fragment Length Polymorphism"; BOTSTEIN et al., 1980). Mais tarde, surge uma nova técnica denominada Minissatélites ou locos de Sequências Adjacentes que se Repetem em Número Variável (VNTR – "Variable Number of Tandem Repeats"; JEFFREY et al., 1985).

Com o surgimento da técnica de Reação de Polimerase em Cadeia (PCR – "Polymerase Chain Reaction"; MULLIS & FALOONA, 1987), aliou-se o poder da informação gerada por marcadores moleculares e a rapidez da técnica baseada em ciclos contínuos de desnaturação e amplificação das fitas de DNA mediados pela enzima DNA polimerase em pontos específicos do genoma, determinados pelo pareamento de iniciadores ou "primers" (seqüência de nucleotídeos de tamanho pequeno, variando geralmente entre 20 a 30 bases) específicos de seqüências

complementares a este ponto. Várias outras técnicas utilizam o princípio da técnica de PCR, tais como os marcadores do tipo: Polimorfismo de DNA Amplificado ao Acaso (RAPD – “Random Amplified Polymorphic DNA”; WILLIAMS et al., 1990); Regiões Amplificadas Caracterizadas por Seqüências (SCAR – “Sequence Characterized Amplified Regions”; PARAN & MICHELMORE, 1993); Microssatélites ou Seqüências Simples Repetidas (SSR – “Simple Sequence Repeats”; Litt & Luty, 1989); Polimorfismo de Comprimento de Fragmentos Amplificados (AFLP – “Amplified Fragment Length Polymorphism”; VOS et al., 1995) e os Polimorfismos de Base Única (SNP’s – “Single Nucleotide Polymorphisms”; COLLINS et al., 1998).

A técnica RFLP tem sido utilizada em um grande número de estudos de caracterização de cultivares, o que se deve à sua alta consistência e repetibilidade na obtenção dos resultados. Os RFLP’s cobrem amplamente o genoma da espécie estudada, pois as sondas utilizadas podem ser obtidas de regiões transcritas (cDNA). Outra vantagem é sua expressão co-dominante, o que possibilita distinguir os indivíduos homozigotos dos heterozigotos, gerando volume maior de informações para estudos genéticos e aplicações no melhoramento. Os RFLPs ainda têm outras utilidades como a identificação de germoplasma, a identificação de variedades, o controle de qualidade na produção de sementes híbridas, a caracterização genética de populações, o monitoramento nos retrocruzamentos e auxílio na identificação e clonagem de genes, entre outras. Como limitações, pode-se dizer que o desenvolvimento de marcadores RFLP é um processo caro e trabalhoso, tornando difícil a geração de grande volume de dados (CAIXETA et al., 2006).

A técnica de Minissatélites é baseada na clivagem do DNA com enzimas de restrição, separados eletroforeticamente, imobilizado em membrana e a detecção é feita através da hibridização com sondas. Uma vantagem adicional dos minissatélites é o alto grau de polimorfismo apresentado, decorrente da variação na distribuição dos sítios de restrição, das sondas utilizadas e do número e tipos das seqüências repetitivas (CATELLI, 2005).

Na técnica de RAPD são amplificados fragmentos de DNA distribuídos ao acaso no genoma, sem a necessidade do conhecimento prévio da seqüência do DNA. Os

marcadores RAPD apresentam como característica básica a dominância, ou seja, este tipo de marcador não permite distinguir indivíduos homozigotos dominantes de heterozigotos em uma população (CAIXETA et al., 2006).

Os Microssatélites são seqüências curtas repetidas em tandem de 1 a 4 nucleotídeos, bem distribuídos em todo genoma eucarioto; são codominantes, multialélicos, reprodutíveis e amplificados via reação de polimerase em cadeia (PCR) utilizando-se um par de iniciadores específicos (de 20 a 30 pb) e complementares que flanqueiam as regiões de microssatélite (FERREIRA & GRATTAPAGLIA, 1998). Tendo em vista a expressão codominante e o multialelismo, os marcadores microssatélites são os que possuem o mais elevado conteúdo de informações de polimorfismo, ou seja, PIC ("Polymorphism Information Content"). Por causa disso, toda e qualquer população segregante pode ser utilizada como população referência para estudos de ligação e mapeamento genético (FERREIRA & GRATTAPAGLIA, 1998).

Nos últimos anos, os Microssatélites têm sido considerados ideais para a caracterização e avaliação da variabilidade genética, sendo uma das ferramentas mais utilizadas no melhoramento de plantas. Com o uso desses marcadores é possível organizar o germoplasma ativo de um programa de melhoramento em conjuntos gênicos, facilitando a escolha e diminuindo o número de combinações a serem feitas pelo melhoristas.

2.6. Divergência Genética

A análise da diversidade genética se destina à identificação de genitores adequados à obtenção de híbridos com maior efeito heterótico e que proporcionem maior segregação em recombinações, possibilitando o aparecimento de transgressivos (CRUZ & CARNEIRO, 2006).

A similaridade genética (ou divergência genética) entre os genótipos é muito útil na seleção de combinações entre parentais para formação de populações segregantes,

bem como para a manutenção da diversidade genética em programas de melhoramento. Essas estimativas revelam disponibilidade de alelos substitutos para caracteres de interesse, que são a base para ganhos de seleção em longo prazo. Cruzamentos entre parentais geneticamente divergentes proporcionam uma variabilidade genética entre progênes bem maior do que as obtidas em cruzamentos entre plantas com maior grau de parentesco, e com isso aumentam as oportunidades do melhorista obter progênes superiores (MESSMER et al., 1993).

Segundo FALCONER (1981), a heterose manifestada em híbridos é função dos efeitos da dominância dos genes para o caráter em questão e do quadrado da diferença das freqüências gênicas de seus genitores, além de efeitos epistáticos que geralmente são negligenciados.

Existem duas maneiras básicas de se inferir a diversidade genética, sendo a primeira de natureza quantitativa e a outra de natureza preditiva. Entre os métodos de natureza quantitativa de avaliação da diversidade, ou da heterose manifestada nos híbridos, citam-se as análises dialélicas, que, por serem necessárias avaliações de p genitores e de todas (ou amostras de) as suas combinações híbridas, resultam dificuldades de avaliação e execução (CRUZ & CARNEIRO, 2006).

Entre os métodos preditivos da heterose citam-se aqueles que tomam por base as diferenças morfológicas, fisiológicas ou moleculares, quantificando-as em alguma medida de dissimilaridade que expressa o grau de diversidade genética entre os genitores (CRUZ & CARNEIRO, 2006).

A avaliação por meio de estudos genealógicos, normalmente, é dificultada pela falta de registro da genealogia das populações. A diversidade geográfica tem sido utilizada frequentemente como indicação da divergência genética, embora a origem geográfica não quantifique a divergência genética entre populações, fato este que tem sido constatado para muitas espécies, como melão, tomate, grão-de-bico, batata, algodão, mostarda e cana-de-açúcar. Por outro lado, a deriva genética e a seleção em diferentes ambientes podem causar maior divergência que a distância geográfica. Outro fato a ser considerado, é que para, muitas espécies, o intenso intercâmbio de materiais entre pesquisadores e instituições de diferentes regiões, tem provocado uma

recombinação geral de todo germoplasma, com o conseqüente desaparecimento de tipos particulares (PEREIRA, 1989).

Atualmente os marcadores moleculares têm sido empregados como forma de avaliar a similaridade genética de maneira mais precisa (CAIXETA et al., 2006). O desenvolvimento e aplicação de tecnologias baseadas em marcadores moleculares, fornecem ferramentas únicas, capazes de revelar polimorfismos ao nível da sequência do DNA, suficientes para discriminar a variação genética entre indivíduos e dentro de populações (KRESOVICH et al., 1995). Dessa forma, a diversidade genética entre acessos é avaliada por meio de variáveis binárias, a partir das quais é avaliada a presença ou ausência de marcas (bandas), codificadas em zeros e uns. Nessa situação, os coeficientes de similaridade entre pares de genótipos podem ser obtidos levando-se em conta os valores:

- a) Valor que quantifica o número de coincidência de tipo 1-1 para cada par de genótipo;
- b) Valor que quantifica o número de discordância do tipo 1-0 para cada par de genótipo;
- c) Valor que quantifica o número de discordância o tipo 0-1 para cada par de genótipo;
- d) Valor que quantifica o número de coincidência do tipo 0-0 para cada par de genótipo.

Com base nesses valores podem ser obtidas as medidas de similaridade. Apesar de existirem vários coeficientes de similaridade e dissimilaridade, na prática tem sido de uso mais rotineiro o coeficiente de Coincidência Simples, o de Jaccard e o de Nei e Li. Por ser objeto de estudo desse trabalho, abordaremos apenas os Coeficientes de Similaridade de Coincidência Simples e o de Rogers e Tanimoto. O Coeficiente de Coincidência Simples é igual ao quadrado da Distância Euclidiana Média, medida bastante utilizada em características quantitativas, além de ter a característica de considerar como fator de similaridade o número de coincidência de ausência da marca molecular (coincidência tipo "0-0") (CRUZ & CARNEIRO, 2006). O Coeficiente de Rogers e Tanimoto segue a mesma linha, incluindo o d , ou seja, considerando como

válida a coincidência “0-0”; somente diferindo do método de coincidência simples pelo fato de dobrar a importância de b e c no denominador, ou seja, as discordâncias “1-0” e “0-1”.

2.6.1. Análise de agrupamento

A análise de agrupamento permite classificar n itens (populações, clones, variedades, indivíduos, etc.) avaliados por um conjunto p de variáveis, cujo objetivo é identificar e separar os itens em grupos, de forma que os mais semelhantes permaneçam no mesmo grupo. Em todos os casos se desconhece “a priori” o número e a composição dos diferentes grupos ou “clusters” a serem formados.

Atualmente, a técnica é muito usada principalmente por pesquisadores da área de melhoramento genético, em estudos de divergência e estudos evolutivos. Esse tipo de análise tem por finalidade reunir, por algum critério de classificação, qualquer tipo de unidade amostral em vários grupos, de tal forma que exista homogeneidade dentro do grupo e heterogeneidade entre grupos (BUSSAB et al., 1990; DIAS, 1998; MEYER, 2002; CRUZ et al., 2004).

Inicia-se o processo definindo-se os indivíduos e os objetivos desejados para a aplicação da análise, além dos critérios que irão definir as similaridades entre eles. Obtidos esses dados, eles são dispostos na forma de uma matriz, que representam os indivíduos de interesse e as linhas representam as variáveis. A matriz de distância é gerada a partir de amostras de n itens, totalizando $n(n-1)/2$ pares de distâncias. A seguir um algoritmo de agrupamento é aplicado sobre essa matriz, de modo a identificar e conectar grupos homogêneos. Tais grupos são representados graficamente em um diagrama de árvore denominado dendrograma (DIAS, 1998; MEYER, 2002; ARRIEL, 2004).

Existem inúmeros métodos de agrupamento, que se distinguem pelo tipo de resultado a ser fornecido e pelas diferentes formas de definir a proximidade entre um

indivíduo e um grupo já formado ou entre dois grupos quaisquer. Cada critério de análise de agrupamento impõe certo grau de estrutura nos dados, e para se assegurar que o resultado obtido não é um artefato da técnica utilizada, é recomendado que diferentes critérios de agrupamento sejam aplicados e se aceite a estrutura resultante da maior parte deles. (DIAS, 1998; CRUZ & CARNEIRO, 2006).

Dos métodos de agrupamento, os mais utilizados são os hierárquicos e os de otimização. Nos métodos hierárquicos, os genótipos são agrupados por um processo que se repete em vários níveis, até que seja estabelecido o dendrograma ou o diagrama de árvore. As delimitações podem ser estabelecidas por um exame visual do dendrograma, em que se avaliam pontos de alta mudança de nível, tomando-os em geral como delimitadores do número de genótipos para determinado grupo (CRUZ & CARNEIRO, 2006).

No agrupamento hierárquico denominado Método do Vizinho Mais Próximo ou Método de Ligação Simples ("Single Linkage Method"), forma-se um grupo inicial de indivíduos mais similares e em seguida são calculadas as distâncias daquele grupo em relação aos demais indivíduos usando-se a menor distância do conjunto. Neste método, as conexões entre indivíduos e grupos são feitas por ligações simples entre pares de indivíduos, ou seja, a distância entre os grupos é definida como sendo aquela entre os indivíduos mais parecidos entre esses grupos (CRUZ & CARNEIRO, 2006). Este método leva a grupos longos e os dendrogramas resultantes deste procedimento são geralmente pouco informativos, devido à informação dos indivíduos intermediários que não são evidentes (CARLINI-GARCIA, 1998).

Quando a semelhança entre dois grupos é definida como a máxima distância entre pares de indivíduos tem-se o Método do Vizinho Mais Distante ou Método de Ligação Completa ("Complete Linkage Method"). A similaridade entre dois grupos é definida como aquela apresentada pelos indivíduos de cada grupo que menos se parecem, formando-se todos os pares com um membro de cada grupo. Este método, geralmente leva a grupos compactos e discretos, tendo os seus valores de similaridade relativamente pequenos (MEYER, 2002).

O Método de Ward considera, para a formação inicial do grupo, aqueles indivíduos que proporcionam a menor soma de quadrados dos desvios entre acessos, admitindo perdas de informações em razão do agrupamento realizado (CRUZ & CARNEIRO, 2006).

Se a distância entre os pares de grupos for estimada com base na média aritmética dos indivíduos tem-se o UPGMA (“Unweighted Pair-Group Method Using Arithmetic Averages”), ou Ligação média entre grupo. Este último evita caracterizar a dissimilaridade por valores extremos entre os genótipos, não considera a estrutura de subdivisão do grupo, dando pesos iguais a cada indivíduo do grupo calculando a similaridade média de um indivíduo que pretende se juntar ao grupo existente. É um dos mais empregados nos estudos de divergência genética (CRUZ et al., 2004; CRUZ & CARNEIRO, 2006).

O Método da Ligação Média Dentro de Grupos, também se baseia em agrupamentos em função da média aritmética, calculando os coeficientes de similaridade médios entre o indivíduo que se pretende agrupar e os indivíduos do grupo já existente (MEYER, 2002).

O Método da Mediana ou WPGMA (“Weighted Pair-Group Method Using Arithmetic Averages”) difere do UPGMA pelo fato de atribuir aos indivíduos admitidos mais recentemente a um grupo, peso igual aos dos demais indivíduos já pertencentes ao grupo (ROMESBURG, 1984).

Nos métodos de otimização, os grupos são formados pela adequação de algum critério de agrupamento, ou seja, o objetivo é alcançar uma partição dos indivíduos que otimize (maximize ou minimize) alguma medida predefinida. Um dos métodos mais comumente utilizados no melhoramento genético é o proposto por Tocher, citado por RAO (1952). Ele identifica o par de indivíduos mais similares para formação do grupo inicial e a partir daí avalia a possibilidade de inclusão de novos indivíduos, adotando-se o critério de que a distância média intragrupo deve ser menor que a distância média intergrupo (CRUZ & CARNEIRO, 2006).

Ressalta-se que o agrupamento conduz à perda de informações ao nível de indivíduos, restando apenas informações sobre grupos similares. Por esse motivo, o

estudo da divergência pode ser conduzido também simultaneamente, por componentes ou coordenadas principais, isto é particularmente importante quando o número de indivíduos é grande (CRUZ & CARNEIRO, 2006).

Uma alternativa de análise da diversidade genética foi proposta por CRUZ & VIANA (1994) a qual apresenta características similares às técnicas de agrupamento e de coordenadas principais, sendo denominada de Projeção de Medidas de Dissimilaridade no Plano bi ou tridimensional. Neste procedimento, as medidas são convertidas em escores relativos às variáveis X e Y, que quando representadas em gráficos irão refletir, no espaço bi ou tridimensional, as distâncias originalmente obtidas a partir do espaço n-dimensional (n = número de caracteres utilizados para obtenção das distâncias). Inicialmente, é necessário estabelecer uma ordem decrescente de dissimilaridade entre os pares de indivíduos. Estimam-se coordenadas para cada indivíduo, a partir da matriz de dissimilaridades, por procedimentos estatísticos que minimizam as diferenças entre as distâncias originais e as obtidas no espaço bidimensional. A viabilidade do uso desta técnica pode ser avaliada pela correlação entre as distâncias, pelo grau de distorção (α) ou pela adequação da representação gráfica (coeficiente de estresse – s).

3. MATERIAL E MÉTODOS

3.1. Local de Condução do Experimento

Os experimentos foram conduzidos na casa de vegetação do Departamento de Produção Vegetal e no Laboratório de Biotecnologia Aplicada ao Melhoramento de Plantas da Faculdade de Ciências Agrárias e Veterinárias, UNESP – Campus de Jaboticabal – SP.

3.2. Identificação e Coleta do Material

No presente estudo, foram analisadas 10 linhagens de soja em geração F_8 de endogamia, desenvolvidas pelo Programa de Melhoramento da Soja da UNESP-FCAV, e 30 cultivares disponíveis no mercado. Tais genótipos foram selecionados por possuírem resistência ao nematóide de cisto da soja, raça 3 (Tabela 1).

Os genótipos foram semeados em vasos de plástico em casa de vegetação, com condições de luz e umidade controladas. As amostras foram coletadas quando as plantas apresentaram o segundo par de trifólios desenvolvidos (estádio V2), com coleta de um ou dois folíolos jovens por planta, de cinco plantas por genótipo. A seguir, foram armazenadas, etiquetadas e depositadas em gelo. Na sequência, foram levadas ao laboratório para armazenamento em freezer a -80°C para posterior extração de DNA.

Tabela 1. Cultivares e genótipos F₈ de soja resistentes ao NCS, utilizados no estudo de divergência genética, com suas respectivas fontes de resistência.

N°	Cultivares/Genótipos	Fontes de Resistência envolvidas na genealogia^{1/}
1	FMT Matrinxã	PI 437654 e PI 88788 e Peking
2	BRS Jiripoca	Peking e PI 437654
3	BRS Piraíba	PI 437654 e PI 90763, PI 88788 e Peking
4	BRS Tambaqui	PI 437654 e Peking
5	FMT Tucunaré	Peking e PI 437654
6	BRSGO Ipameri	PI 88788 e Peking
7	BRSGO Chapadões	PI 437654
8	BRS 231	Peking e PI 437654
9	FMT Cachara	Peking e PI 437654
10	BRSMG Liderança (BR91-10569)	Peking
11	BRSMG 250 (Nobreza)	Peking
12	BRSMG 251 (Robusta)	Peking
13	M-SOY 7901	Peking
14	M-SOY 8001	Peking
15	M-SOY 8200	Kirby
16	M-SOY 8400	Peking
17	M-SOY 8329	Peking
18	M-SOY 8757	Peking
19	BRSMT Pintado (MTBR95-123247)	PI 437654
20	BRSMG Renascença	Kirby e Forrest (Peking)
21	Hartwig	PI 437654
22	MGBR 9520937	Forrest (Peking)
23	98N41	Peking
24	98N71	Peking
25	98N82	Peking

Continua...

Tabela 1. Cultivares e genótipos F₈ de soja resistentes ao NCS, utilizados no estudo de divergência genética, com suas respectivas fontes de resistência (Continuação).

26	Foster	Peking
27	Fayette	PI 88788
28	Egyptiam	-
29	Leflore	Peking
30	Jeff	PI 88788
31	JAB 99102/4.5	BR92-15440 - Peking
32	JAB 99102/8.1	BR92-15440 - Peking
33	JAB 99103/8.2	BR92-15440 - Peking
34	JAB 99174/9.1	PI 437654
35	JAB 99174/9.15	PI 437654
36	JAB 99101/2.5	BR92-15440 - Peking
37	JAB 99102/4.2	BR92-15440 - Peking
38	JAB 99174/9.10	PI 437654
39	JAB 994311/15.20	PI 437654
40	JAB 994012/1.2	BR92-1554 - Peking

^{1/}espaços com hífen indicam falta de informação sobre a fonte de resistência da cultivar.

3.3. Atividades em Laboratório

3.3.1. Extração de DNA

O DNA genômico foi extraído seguindo o protocolo de extração descrito em FERREIRA & GRATTAPAGLIA (1998), sendo conhecido como método CTAB. A seguir,

realizou-se a quantificação do DNA extraído, mediante leitura da densidade ótica a 260 e 280 nm em aparelho biofotômetro BioPhotometer Eppendorf AG. Posteriormente à quantificação, a integridade do DNA foi observada pela visualização em gel de agarose 1,5%, corado com brometo de etídeo. Por fim, alíquotas das amostras de DNA foram diluídas para a concentração desejada e parte foi estocada a -20° C para uso posterior.

3.3.2. Identificação de marcadores SSR

Foram selecionados 12 pares de iniciadores SSR ou microssatélites, que flanqueiam regiões repetitivas dispersas no genoma da soja, em diferentes grupos de ligação: Satt185 (E), Satt308 (M), Satt187 (A2), Satt070 (B2), Satt009 (N), Satt129 (D1a), Satt154 (D2), Satt152 (N), Satt175 (M), Satt045 (E), Satt162 (A2) e Satt163 (G) (Tabela 2).

3.3.3. Reações de PCR e visualização dos resultados

As reações de amplificação por PCR, foram realizadas para cada um dos 40 genótipos, utilizando os diferentes iniciadores SSR. A reação foi preparada com um volume total de 20 μ L, contendo 300 ng de DNA genômico de soja, 1,5 mM de $MgCl_2$, Tampão PCR 1X (10 mM Tris-HCl, pH 8,0 e 50 mM KCl), 0,2 mM de dNTP's, 1U de Taq-DNA polimerase e 10 μ M de cada iniciador "Forward" e "Reverse".

Tabela 2. Grupo de ligação, sequência e número de alelos encontrados para os 12 marcadores SSR utilizados nas análises dos 40 genótipos de soja.

Iniciador	GL		Seqüências	T	NA
Satt308	M	F	GCGTTAAGGTTGGCAGGGTGGAAAGTG	55 °C	3
		R	GCGCAGCTTTATACAAAAATCAACAA		
Satt154	D2	F	AGATACTAACAAGAGGCATAAACT	55 °C	2
		R	AAAGAAACGGAATAATACTACATT		
Satt152	N	F	GCGCTATTCCTATCACAACACA	55 °C	3
		R	TAGGGTTGTCACTGTTTTGTTCTTA		
Satt070	B2	F	TAAAAATTAAAATACTAGAAGACAAC	50 °C	3
		R	TGGCATTAGAAAATGATATG		
Satt129	D1a	F	TTCAGTACAAGTCGGGTGAATAATAATA	59 °C	3
		R	TCACATGTTCTGGGACTTAAGGTAT		
Satt175	M	F	GACCTCGCTCTCTGTTTCTCAT	60 °C	3
		R	GGTGACCACCCCTATTCCTTAT		
Satt185	E	F	GCGCATATGAATAGGTAAGTTGCACTAA	55 °C	3
		R	GCGTTTTCTCTACAATAATATTTTCTCAT		
Satt045	E	F	TGGTTTCTACTTTCTATAATTATTT	50 °C	3
		R	ATGCCTCTCCCTCCT		
Satt009	N	F	CCAAGTTGAAATTACTAGAGAAA	55 °C	5
		R	CTTACTAGCGTATTAACCCCTT		
Satt187	A2	F	GCGTTTAAATTTATGATATAACCAA	55 °C	3
		R	GCGTTTTATCTCTTTTTCCACAAC		
Satt162	A2	F	GCGTGGTTTTTCGCTGGATATA	46-57 °C	2
		R	GCGCATTTTCGTAACATATTTTTTCTCAC		
Satt163	G	F	GCGGTATATATGTTTGCAAGACATATT	46-57 °C	3
		R	GCGGAATCTCGCCCAGGAGGAAGCTT		

GL: grupo de ligação; **T:** temperatura de pareamento; **NA:** n° de alelos encontrados; **F:** Forward; **R:** Reverse

O programa de termociclagem foi composto por um ciclo inicial de desnaturação a 95°C por 30 s, seguido de 94°C por 30 s, 22 ciclos para o anelamento do iniciador, onde a temperatura variou de acordo com cada iniciador (Tabela 2) e 68°C durante 30 s, para extensão pela Taq Polimerase. Em seguida finalizou-se o programa com 4°C, até que os microtubos fossem retirados do aparelho. Após a retirada, adicionou-se 3 µL de tampão de carregamento (5,75 mL de glicerol; 2 mL de EDTA 0,5 M, pH 8,0; 0,5 mL de dodecilsulfato de sódio - SDS; 1,75 mL de água e 0,01% de azul de bromofenol) em cada microtubo, imediatamente antes da aplicação da amostra no gel.

Os produtos amplificados foram separados por eletroforese em gel de agarose de alta resolução (agarose 1000), a uma concentração de 3,0% para um melhor padrão de diferenciação de bandas, em tampão TAE (2 M Tris, 100 mM EDTA, 1M ácido acético) e corados com brometo de etídeo. A eletroforese procedeu-se a 80 V por cerca de duas horas. A visualização foi feita em luz UV através de transluminador e fotografada pelo sistema Kodak EDAS 290. A amplitude de variação do tamanho de cada loco foi estimada por meio de comparação com o marcador de tamanho 100 pb e também através de das ferramentas digitais do Programa Kodak EDAS 290.

3.3.4. Organização da matriz de dissimilaridade

A partir dos resultados moleculares observados nos géis, foi organizada uma matriz binária baseada na presença (1) ou ausência (0) de marcas (bandas) para cada alelo encontrado. Com base nessas observações, foram obtidas as estimativas de similaridade entre todos os pares de acessos.

3.4. Análises Estatísticas

As medidas de similaridade (C_{ij}) foram convertidas em dissimilaridade através do complemento aritmético (D), utilizando-se dois coeficientes:

- Coincidência Simples (Cruz & Carneiro, 2006), dado por:

$$C_{ii'} = \frac{a + d}{a + b + c + d};$$

- Rogers e Tanimoto (Cruz & Carneiro, 2006), dado por:

$$C_{ii'} = \frac{a + d}{[a + 2(b + c) + d]};$$

Sendo: $D_{ii'} = 1 - C_{ii'}$,

Onde,

a: número de concordâncias do tipo 1-1;

b: número de discordâncias do tipo 1-0;

c: número de discordâncias 0-1;

d: número de concordâncias 0-0.

Tais coeficientes foram utilizados devido ao fato de considerarem como fator de similaridade, o número de coincidências da ausência da marca molecular (tipo “0-0”), fator importante para dados obtidos a partir de marcadores SSR. Considerando a base genética de marcadores SSR, a ausência de amplificação de uma determinada banda em dois genótipos, pode indicar uma similaridade genética entre eles. Desta forma, coeficientes que incluem a co-ocorrência negativa (d: 0,0) das expressões de similaridade tornam-se mais adequados em análises que utilizam marcadores desta natureza.

Após a obtenção das informações sobre o grau de diferença apresentado entre os genótipos, foram realizados os agrupamentos de acordo com vários métodos

aglomerativos, que se distinguem pelo tipo de resultado fornecido e pelas diferentes formas de definir proximidade entre um indivíduo e um grupo já formado ou entre dois grupos quaisquer.

Foram realizados agrupamentos através do método de otimização de Tocher (RAO, 1952) e de métodos hierárquicos (CRUZ et al., 2004) como:

- Ligação simples – Vizinho mais próximo (“Single Linkage Clustering”);
- Ligação completa – Vizinho mais distante (“Complete Linkage Clustering”);
- Método de Ward;
- Ligação média dentro de grupo;
- Ligação média entre grupo - UPGMA (“Unweighted Pair-Group Method Arithmetic Average”);
- Método da Mediana - WPGMA (“Weighted Pair-Group Method Using Arithmetic Averages”).

Após os agrupamentos, foi calculada a matriz dos genótipos mais e menos similares em relação a cada um dos genótipos estudados. Foram realizadas também, as projeções das medidas de dissimilaridade em gráficos bi e tridimensionais, calculando-se a coordenada das medidas mais divergentes e, a seguir, daquelas que demonstraram, em ordem crescente, a maior diversidade com os genótipos já considerados (CRUZ & VIANA, 1994).

Depois de obtidos todos os possíveis agrupamentos, as medidas de dissimilaridade foram correlacionadas através do Teste de Mantel e tal relação foi quantificada e projetada em um gráfico de correlação, no qual cada eixo representa uma medida de similaridade. Todas as análises foram realizadas através do Programa GENES, versão 4.1 (CRUZ, 2001).

4. RESULTADOS E DISCUSSÃO

Conforme observado na Tabela 2, todos os 12 marcadores SSR apresentaram polimorfismo. O número de alelos por loco variou de dois (Satt 154 e Satt 162) a cinco (Satt 009), com número médio de três alelos por loco, obtendo um total de 36 alelos, distribuídos entre os 40 genótipos. Exemplos de padrões de bandas e polimorfismos detectados pelos iniciadores Satt 308, Satt 152 e Satt 154 são apresentados nas Figuras 1, 2 e 3.

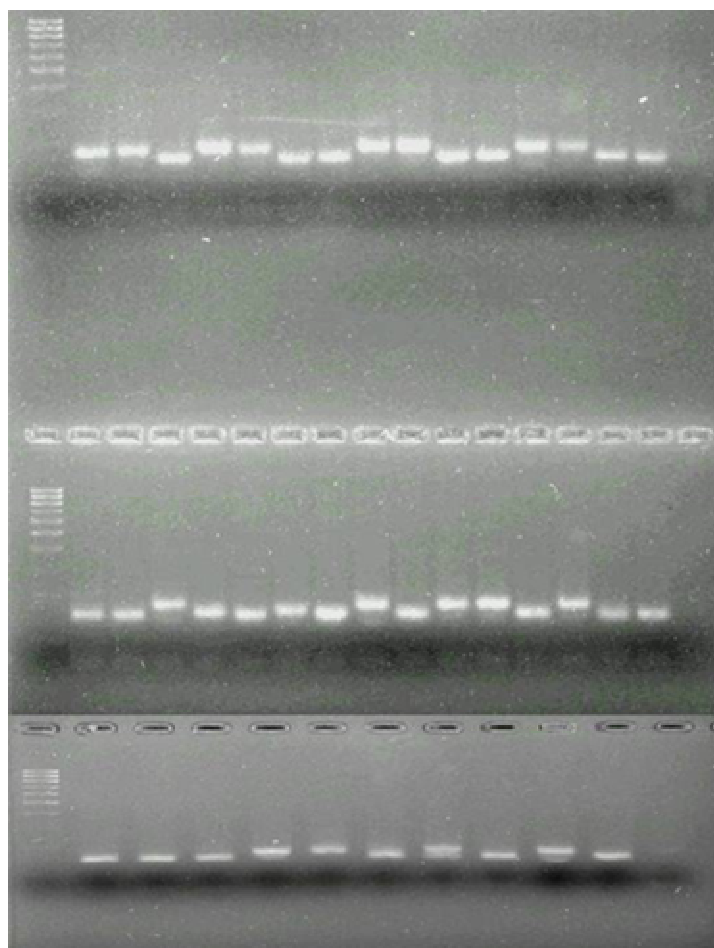


Figura 1. Amostras dos fragmentos produzidos nas reações de amplificação dos 40 genótipos de soja utilizando o iniciador Satt308.

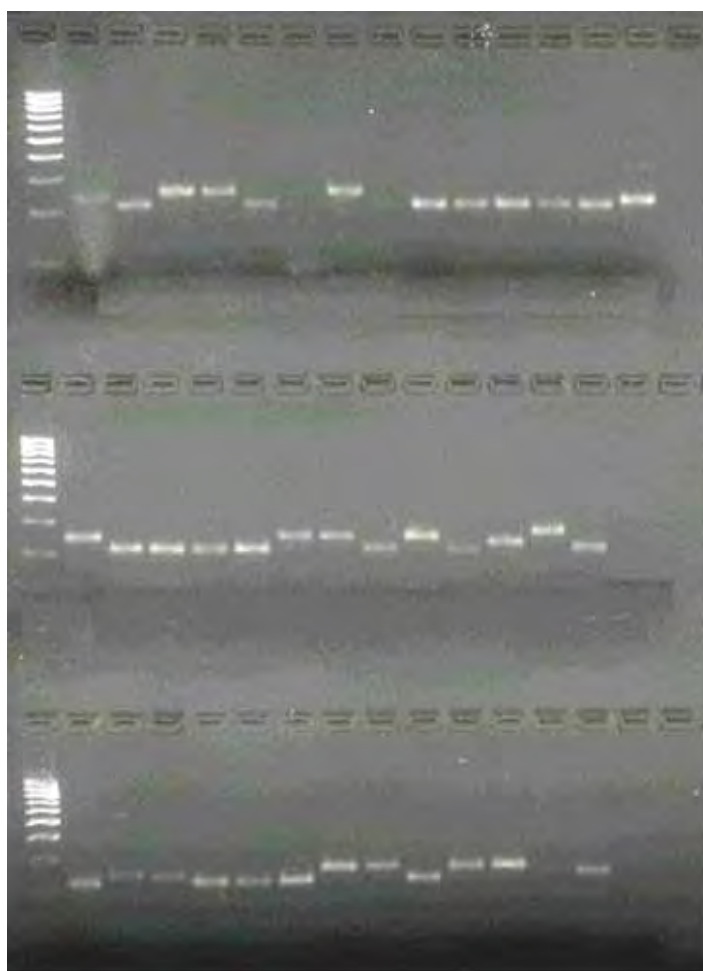


Figura 2. Amostras dos fragmentos produzidos nas reações de amplificação dos 40 genótipos de soja utilizando o iniciador Satt152.

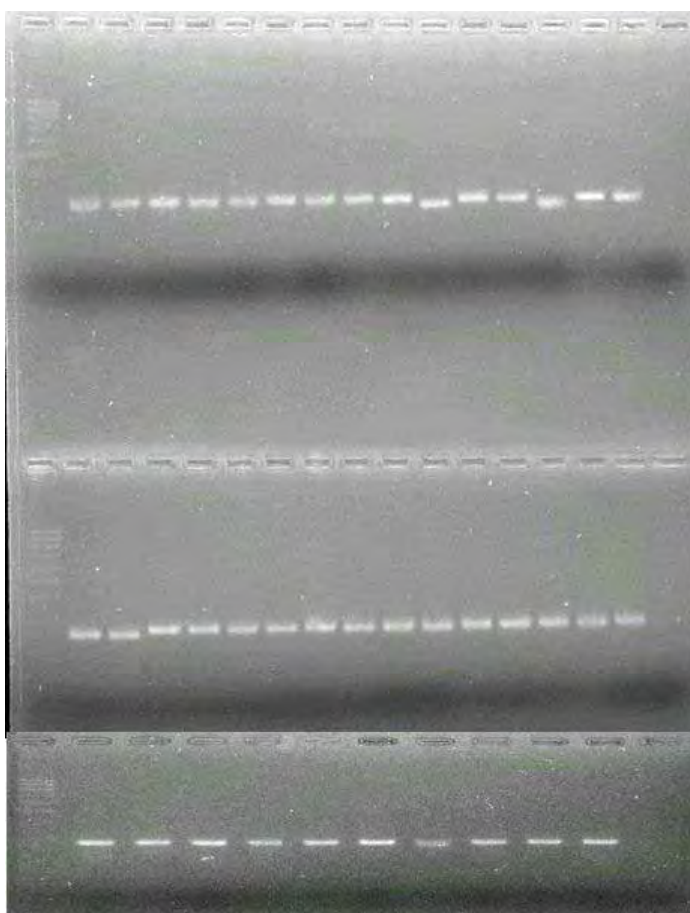


Figura 3. Amostras dos fragmentos produzidos nas reações de amplificação dos 40 genótipos de soja utilizando o iniciador Satt154.

4.1. Coeficientes de Similaridade pelo Método de Coincidência Simples

O coeficiente de dissimilaridade médio obtido pelo complemento aritmético do Coeficiente de Similaridade de Coincidência Simples entre todos os genótipos foi de 0,356297. Esta média pode ser indicativa da estreita base genética da soja, já reportada por YUE et al. (2000), pois baixa dissimilaridade significa alta semelhança entre os genótipos. Os pares de genótipos mais divergentes foram 4 (BRS Tambaqui) e 16 (M-SOY 8400), 13 (M-SOY 7901) e 38 (JAB 99174/9.10), 15 (M-SOY 8200) e 28 (Egyptiam), 16 (M-SOY 8400) e 28 (Egyptiam), 17 (M-SOY 8329) e 28 (Egyptiam), 16 (M-SOY 8400) e 34 (JAB 99174/9.1), 16 (M-SOY 8400) e 35 (JAB 99174/9.15), com coeficiente de dissimilaridade igual a 0,64. Os pares de genótipos menos divergentes foram 11 (Nobreza) e 31 (JAB 99102/4.5), 31 (JAB 99102/4.5) e 32 (JAB 99102/8.1), 32 (JAB 99102/8.1) e 33 (JAB 99103/8.2), 16 (M-SOY 8400) e 17 (M-SOY 8329), com coeficiente de dissimilaridade de 0,06.

A partir da matriz de dissimilaridade com as distâncias genéticas entre os 40 genótipos analisados (Tabela 3), foram obtidos os agrupamentos, através de diferentes métodos aglomerativos.

O método de otimização de Tocher resultou na formação de 17 grupos, com distribuição dos genótipos em um grupo maior, além de 6 grupos unitários, formados por apenas um genótipo (Tabela 4). Os grupos 2 (genótipos 10 - Liderança, 16 – M-SOY 8400 e 17 M-SOY 8329) e 13 (genótipo 35 – JAB 99174/9.15) foram os mais divergentes, com distância média intergrupo de 0,59. Os grupos que apresentaram menor distância entre si foram o 9 (genótipos 14 – M-SOY 8001 e 36 – JAB 99101/2.5) e o 12 (genótipo 39 – JAB 994311/15.20), com média intergrupo de 0,22. Tal método apresenta a vantagem da fácil interpretação dos grupos formados, porém, não houve informações sobre a similaridade dos genótipos dentro de cada grupo. Essa desvantagem do método também foi observada em MEYER et al. (2004), em experimentos com linhagens de milho e ARRIEL et al. (2006), com gergelim.

Tabela 3. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Coincidência Simples.

Genótipo	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0,36	0,47	0,39	0,33	0,19	0,39	0,36	0,50	0,50	0,47	0,42	0,56	0,44	0,42	0,58	0,53	0,36	0,47	0,39
2		0,33	0,25	0,08	0,22	0,36	0,22	0,31	0,42	0,28	0,28	0,36	0,42	0,56	0,44	0,50	0,28	0,44	0,42
3			0,36	0,36	0,39	0,36	0,44	0,47	0,36	0,22	0,56	0,42	0,25	0,28	0,39	0,44	0,39	0,33	0,42
4				0,17	0,31	0,22	0,19	0,39	0,50	0,36	0,36	0,39	0,5	0,47	0,64	0,58	0,31	0,47	0,33
5					0,19	0,28	0,14	0,22	0,33	0,19	0,19	0,28	0,44	0,47	0,47	0,42	0,19	0,36	0,33
6						0,25	0,22	0,36	0,31	0,28	0,33	0,42	0,36	0,39	0,44	0,39	0,22	0,39	0,31
7							0,19	0,28	0,39	0,36	0,31	0,39	0,39	0,42	0,53	0,47	0,25	0,36	0,17
8								0,25	0,42	0,28	0,22	0,36	0,47	0,50	0,56	0,50	0,17	0,33	0,25
9									0,28	0,31	0,14	0,22	0,39	0,42	0,36	0,31	0,14	0,25	0,33
10										0,19	0,42	0,22	0,22	0,31	0,14	0,08	0,31	0,25	0,33
11											0,39	0,31	0,31	0,33	0,33	0,28	0,28	0,28	0,36
12												0,36	0,53	0,39	0,39	0,33	0,28	0,39	0,36
13													0,39	0,47	0,36	0,31	0,25	0,25	0,39
14														0,25	0,25	0,31	0,36	0,31	0,28
15															0,28	0,22	0,44	0,39	0,42
16																0,06	0,44	0,39	0,47
17																	0,39	0,33	0,42
18																		0,22	0,31
19																			0,25

Continua...

Tabela 3. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Coincidência Simples (Continuação).

Genótipo	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	0,31	0,47	0,28	0,36	0,44	0,39	0,22	0,5	0,33	0,25	0,42	0,36	0,42	0,28	0,33	0,36	0,42	0,22	0,36	0,28
2	0,33	0,33	0,25	0,22	0,42	0,31	0,31	0,31	0,42	0,33	0,33	0,33	0,39	0,25	0,36	0,39	0,33	0,36	0,33	0,31
3	0,33	0,39	0,47	0,22	0,42	0,42	0,47	0,53	0,19	0,28	0,22	0,17	0,22	0,36	0,53	0,28	0,28	0,42	0,39	0,42
4	0,36	0,47	0,33	0,36	0,22	0,28	0,39	0,22	0,39	0,36	0,36	0,36	0,42	0,17	0,28	0,53	0,42	0,39	0,36	0,44
5	0,31	0,31	0,22	0,19	0,33	0,22	0,28	0,28	0,33	0,31	0,25	0,25	0,31	0,22	0,33	0,42	0,42	0,44	0,36	0,39
6	0,33	0,28	0,31	0,22	0,47	0,36	0,19	0,36	0,31	0,22	0,22	0,22	0,28	0,25	0,36	0,44	0,39	0,36	0,39	0,31
7	0,36	0,31	0,33	0,36	0,44	0,22	0,33	0,28	0,28	0,36	0,36	0,36	0,42	0,17	0,28	0,53	0,47	0,39	0,36	0,44
8	0,39	0,33	0,25	0,33	0,31	0,25	0,25	0,14	0,31	0,39	0,28	0,33	0,39	0,14	0,19	0,5	0,44	0,42	0,33	0,36
9	0,36	0,19	0,28	0,31	0,44	0,22	0,33	0,33	0,33	0,36	0,36	0,36	0,36	0,33	0,39	0,36	0,36	0,5	0,36	0,44
10	0,36	0,19	0,44	0,25	0,5	0,44	0,44	0,56	0,33	0,36	0,25	0,25	0,25	0,5	0,56	0,31	0,42	0,56	0,42	0,5
11	0,28	0,22	0,42	0,11	0,31	0,42	0,42	0,42	0,25	0,28	0,06	0,11	0,17	0,42	0,47	0,33	0,44	0,58	0,5	0,47
12	0,5	0,33	0,25	0,39	0,42	0,19	0,31	0,31	0,42	0,44	0,44	0,44	0,44	0,31	0,31	0,44	0,44	0,42	0,39	0,42
13	0,31	0,25	0,39	0,31	0,44	0,33	0,39	0,44	0,39	0,36	0,31	0,31	0,25	0,44	0,44	0,47	0,42	0,61	0,47	0,56
14	0,25	0,31	0,44	0,36	0,44	0,5	0,5	0,56	0,28	0,31	0,31	0,31	0,31	0,5	0,56	0,19	0,25	0,44	0,25	0,33
15	0,39	0,39	0,47	0,33	0,42	0,42	0,47	0,64	0,31	0,33	0,33	0,28	0,28	0,42	0,47	0,28	0,22	0,36	0,39	0,47
16	0,50	0,33	0,58	0,39	0,53	0,58	0,58	0,64	0,47	0,50	0,39	0,39	0,39	0,64	0,64	0,28	0,39	0,47	0,39	0,47

Continua...

Tabela 3. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Coincidência Simples (Continuação).

Genótipo	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
17	0,44	0,28	0,53	0,33	0,47	0,53	0,53	0,64	0,42	0,44	0,33	0,33	0,33	0,58	0,58	0,33	0,44	0,53	0,44	0,53
18	0,28	0,22	0,31	0,28	0,36	0,31	0,25	0,31	0,25	0,28	0,22	0,22	0,28	0,25	0,36	0,44	0,39	0,53	0,39	0,47
19	0,28	0,17	0,36	0,33	0,47	0,36	0,25	0,47	0,25	0,39	0,28	0,28	0,22	0,42	0,42	0,33	0,44	0,53	0,44	0,42
20	0,36	0,25	0,22	0,42	0,39	0,28	0,28	0,39	0,28	0,31	0,36	0,36	0,31	0,28	0,28	0,42	0,53	0,44	0,25	0,39
21		0,28	0,36	0,22	0,36	0,47	0,36	0,53	0,25	0,22	0,28	0,28	0,33	0,36	0,47	0,33	0,33	0,47	0,44	0,36
22			0,25	0,17	0,53	0,31	0,25	0,47	0,25	0,28	0,28	0,28	0,22	0,36	0,36	0,39	0,39	0,53	0,50	0,47
23				0,31	0,44	0,11	0,17	0,39	0,28	0,25	0,47	0,42	0,36	0,17	0,22	0,36	0,36	0,39	0,31	0,33
24					0,42	0,36	0,31	0,47	0,19	0,17	0,17	0,11	0,17	0,31	0,47	0,33	0,33	0,47	0,50	0,47
25						0,50	0,56	0,33	0,39	0,36	0,31	0,36	0,42	0,39	0,39	0,47	0,53	0,56	0,42	0,5
26							0,22	0,33	0,33	0,36	0,47	0,42	0,36	0,17	0,22	0,42	0,36	0,39	0,36	0,44
27								0,39	0,33	0,25	0,36	0,31	0,25	0,22	0,28	0,42	0,42	0,33	0,42	0,33
28									0,44	0,53	0,42	0,47	0,53	0,28	0,33	0,58	0,53	0,50	0,42	0,44
29										0,19	0,25	0,19	0,25	0,28	0,44	0,31	0,36	0,50	0,42	0,39
30											0,22	0,17	0,17	0,31	0,42	0,33	0,33	0,42	0,39	0,36
31												0,06	0,11	0,42	0,47	0,39	0,44	0,58	0,50	0,47
32													0,06	0,36	0,53	0,33	0,39	0,53	0,44	0,53

Continua...

Tabela 3. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Coincidência Simples (Continuação).

[illegible]

Tabela 4. Formação de grupos pelo método de otimização de Tocher, utilizando a dissimilaridade através do complemento aritmético do Coeficiente de Coincidência Simples dos 40 genótipos de soja resistentes ao NCS.

Grupo	Indivíduos
I	11, 31, 32, 33, 24, 30, 29, 3
II	16, 17, 10
III	2, 5, 8, 4, 34
IV	23, 26, 27
V	9, 12, 18
VI	7, 20
VII	19, 22
VIII	1, 6
IX	14, 36
X	15, 37
XI	38, 40
XII	39
XIII	35
XIV	21
XV	28
XVI	25
XVII	13

Nos agrupamentos hierárquicos, a delimitação dos grupos é feita de maneira subjetiva, observando os pontos de alta mudança de nível no dendrograma, por isso, diferentes padrões de agrupamento podem ocorrer. Quando não se tem informação sobre a relação genética entre a maioria dos genótipos, não se pode determinar qual método de agrupamento é mais acurado. Assim, ao se comparar os resultados de diferentes critérios, podem-se evitar inferências errôneas (ARRIEL, 2004). De acordo com SNEATH & SOKAL (1973), a formação de um único grupo pode ser enganosa, se

os indivíduos em extremidades opostas da cadeia forem, de fato, completamente dissimilares.

Adotou-se um percentual de similaridade genética em torno de 65%, de acordo com as altas mudanças de níveis observadas nos dendrogramas, além da observação de dados na literatura. Dessa forma, foi constatada a formação de diferentes agrupamentos.

No método do Vizinho Mais Próximo (Figura 4), foram formados 27 grupos, destacando-se vários grupos formados por apenas um genótipo. Formou-se dois grupos envolvendo 5 genótipos (11, 24, 31, 32, 33) e (2, 5, 8, 28, 34), dois grupos que envolveram 3 genótipos (10, 16, 17) e (9, 12, 18), e um grupo englobando os genótipos 23 e 26. Os demais grupos apresentaram-se constituídos por apenas um genótipo. Tal método não se apresentou muito eficaz no agrupamento dos genótipos, pois não distinguiu grupos, o que permitiu a formação separação da maioria dos genótipos em grupos distintos.

O método do Vizinho Mais Distante (Figura 5) formou 8 grupos, com os genótipos bem distribuídos, sendo dois grupos maiores: I (32, 33, 11, 31, 3, 29, 24, 30, 21) e II (7, 20, 23, 26, 8, 34, 35) e os outros grupos menores, com 3 a 5 genótipos. Os grupos apresentaram-se compactos e discretos. A maioria dos genótipos de geração F8 foram agrupados juntos, confirmando o fato de serem oriundos de um mesmo programa de melhoramento e possuírem fontes de resistência comuns. Os genótipos 31 (JAB 99102/4.5), 32 (JAB 99102/8.1), 33 (JAB 99103/8.2) e 11 (Nobreza) permaneceram juntos no agrupamento, além dos genótipos 16 (M-SOY 8400) e 17 (M-SOY 8329), que também foram os genótipos menos dissimilares de acordo com a matriz de Coincidência Simples.

Por sua vez, o Método de Ward (Figura 6) formou apenas cinco grupos, a saber: I (32, 33, 11, 31, 24, 29, 30, 3, 21), II (16, 17, 10, 15, 37, 14, 36, 39), III (6, 27, 1, 38, 40), IV (23, 26, 7, 20, 8, 34, 35, 4, 28, 25) e V (19, 22, 13, 2, 5, 9, 18, 12) e se diferenciou apenas pela formação de um menor número de grupos, aglomerando genótipos que foram considerados menos divergentes e separando os genótipos mais dissimilares em agrupamentos diferentes.

O método de Ligação Média Dentro de Grupo (Figura 7.) possibilitou a formação de 12 grupos englobando de 1 a 9 genótipos cada. Novamente houve coincidência dos agrupamentos com as distâncias genéticas entre os genótipos, porém esse método separou alguns genótipos que apareceram juntos em outros métodos e formou um maior número de agrupamentos.

O agrupamento WPGMA (Figura 9), ou Método da Mediana, apresentou dez agrupamentos, também muito semelhantes ao UPGMA, Ward, Vizinho mais distante e Ligação Média Dentro de Grupo. Apresentou apenas o genótipo 25 (98N82) separado dos demais e manteve os genótipos mais similares em conjunto.

Pode-se observar ainda que, não houve muitas divergências entre os agrupamentos derivados da matriz de dissimilaridade, oriunda do complemento aritmético do Coeficiente de Coincidência Simples, o que pode sugerir uma boa precisão dos dados obtidos. Com exceção dos métodos de Tocher e Vizinho mais Próximo, ocorreu o compartilhamento de grupos pela maioria dos métodos. Observa-se que alguns grupos aparecem iguais nos demais métodos, são eles: um primeiro grupo formado pelos genótipos 3 (BRS Piraíba), 11 (Nobreza), 21 (Hartwig), 24 (98N71), 29 (Leflore), 30 (Jeff), 31 (JAB 99102/4.5), 32 (JAB 99102/8.1), 33 (JAB 99103/8.2) e 34 (JAB 99174/9.1), com fonte de resistência de Peking e da PI 437654; e outro grupo constituído pelos genótipos 16 (M-SOY 8400), 17 (M-SOY 8329) e 10 (Liderança), todos derivados de Peking.

No geral, são muitas as coincidências de agrupamentos hierárquicos e também do de otimização de Tocher. Alguns genótipos aparecem sempre no mesmo grupo, independente do método aglomerativo, tais como: 32 (JAB 99102/8.1), 33 (JAB 99103/8.2), 11 (Nobreza), 31 (JAB 99102/4.5) e 24 (98N71), com o parental resistente em comum; 16 (M-SOY 8400), 17 (M-SOY 8329) e 10 (Liderança), todos compartilhando a mesma fonte de resistência; os genótipos 2 (BRS Jiripoca) e 5 (FMT Tucunaré), com Peking e PI 437654 na genealogia; 9 (FMT Cachara), 12 (Robusta) e 18 (M-SOY 8757), com fonte de resistência ao NCS derivada de Peking.

Além disso, foram realizadas as projeções das distâncias no plano, que convertem as medidas de dissimilaridade em escores relativos às variáveis X e Y, que,

quando representadas em gráficos de dispersão, refletem no espaço bidimensional, as distâncias originalmente obtidas a partir do espaço n-dimensional (n = número de caracteres utilizados para obtenção das distâncias) (Cruz & Carneiro, 2006). Os dados originais foram calculados e projetados no plano através de diferentes tipos de análise gráfica. A projeção 2D (bidimensional) revelou uma distorção de 35,8% e estresse de 44,2%, com correlações entre as distâncias originais e as estimadas de 0,58 (Figura 10). Já as projeções 3D (tridimensionais) apresentaram uma menor distorção, de 20,9% e estresse de 29,7%; a correlação entre as distâncias originais e as estimadas foi de 0,73 (Figura 11 e 12).

O processo de dispersão das medidas de dissimilaridade no plano pode ser considerado satisfatório quando os coeficientes que expressam o grau de distorção e o estresse são inferiores a 20% (Cruz & Carneiro, 2006). Dessa forma, as projeções mais adequadas são as tridimensionais, que apresentam valores de distorção e estresse mais próximos do adequado, além de alta correlação entre as distâncias originais e as distâncias estimadas.

A dispersão bidimensional evidencia alguns genótipos mais distantes, tais como 12 (Robusta), 16 (MSOY 8400) e 21 (Hartwig) e alguns genótipos bem próximos como 4 (Tambaqui), 23 (98N41) e 27 (Fayette), além de 31 (JAB 99102/4.5), 33 (JAB 99103/8.2) e 11 (Nobreza); concordando em parte com os resultados dos dendrogramas.

As projeções tridimensionais ilustram um grupo bem adensado, com alguns genótipos mais distantes, como o 12 (Robusta), 16 (M-SOY 8400), 17 (M-SOY 8329), 21 (Hartwig), 32 (JAB 99102/8.1) e 35 (JAB 99174/9.15). As projeções 3D com rotação apresentaram-se mais adequadas para visualização e análise, com 3 eixos (x: x1, y: x2 e z: x3). Esse método não obteve muito sucesso na distinção dos genótipos em grupos, além de não haver adequada distinção entre as coincidências com os resultados dos agrupamentos hierárquicos.

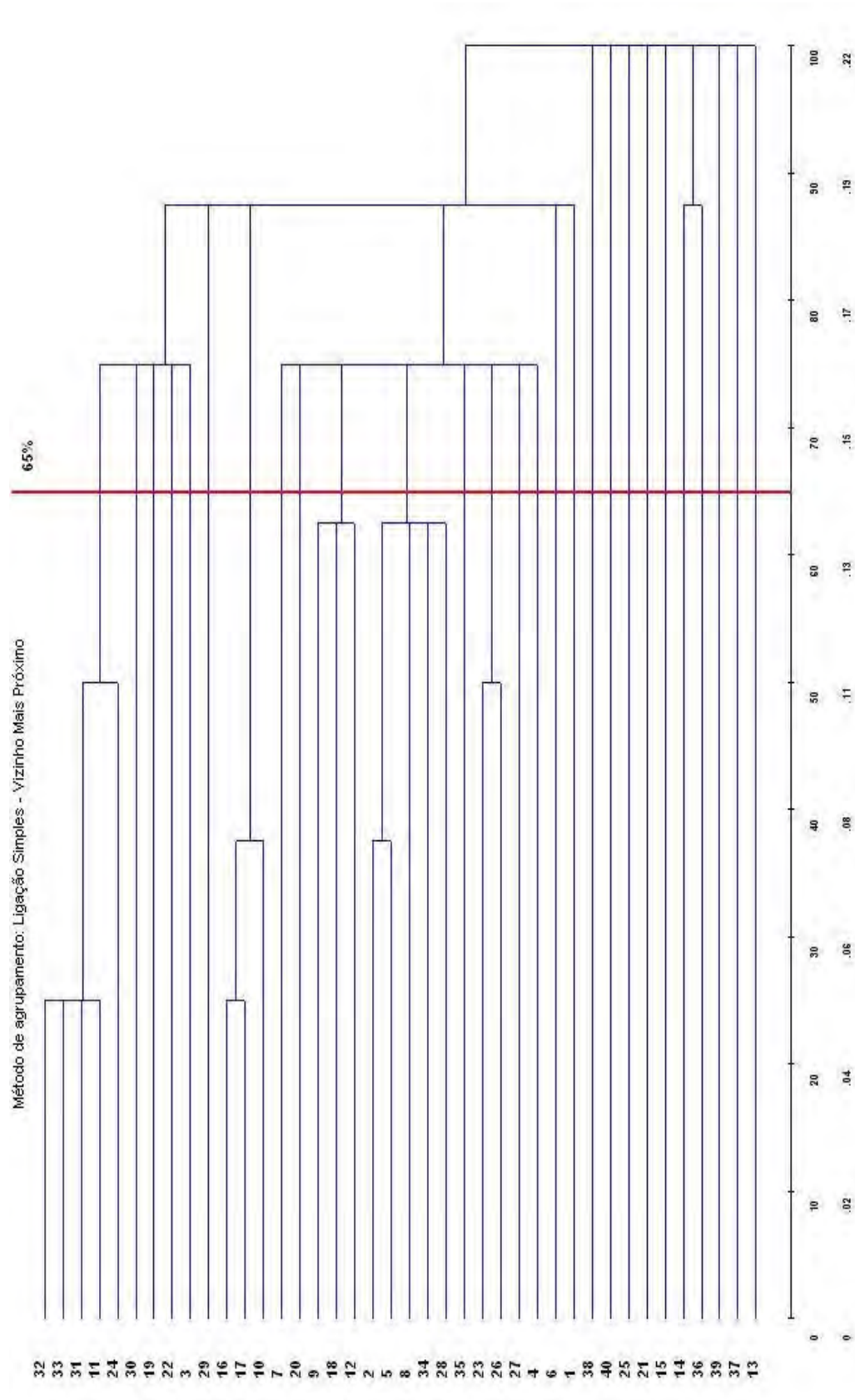


Figura 4. Representação do dendrograma obtido através do método de Ligação Simples – Vizinho Mais Próximo, a partir do complemento aritmético do Coeficiente de Coincência Simples.

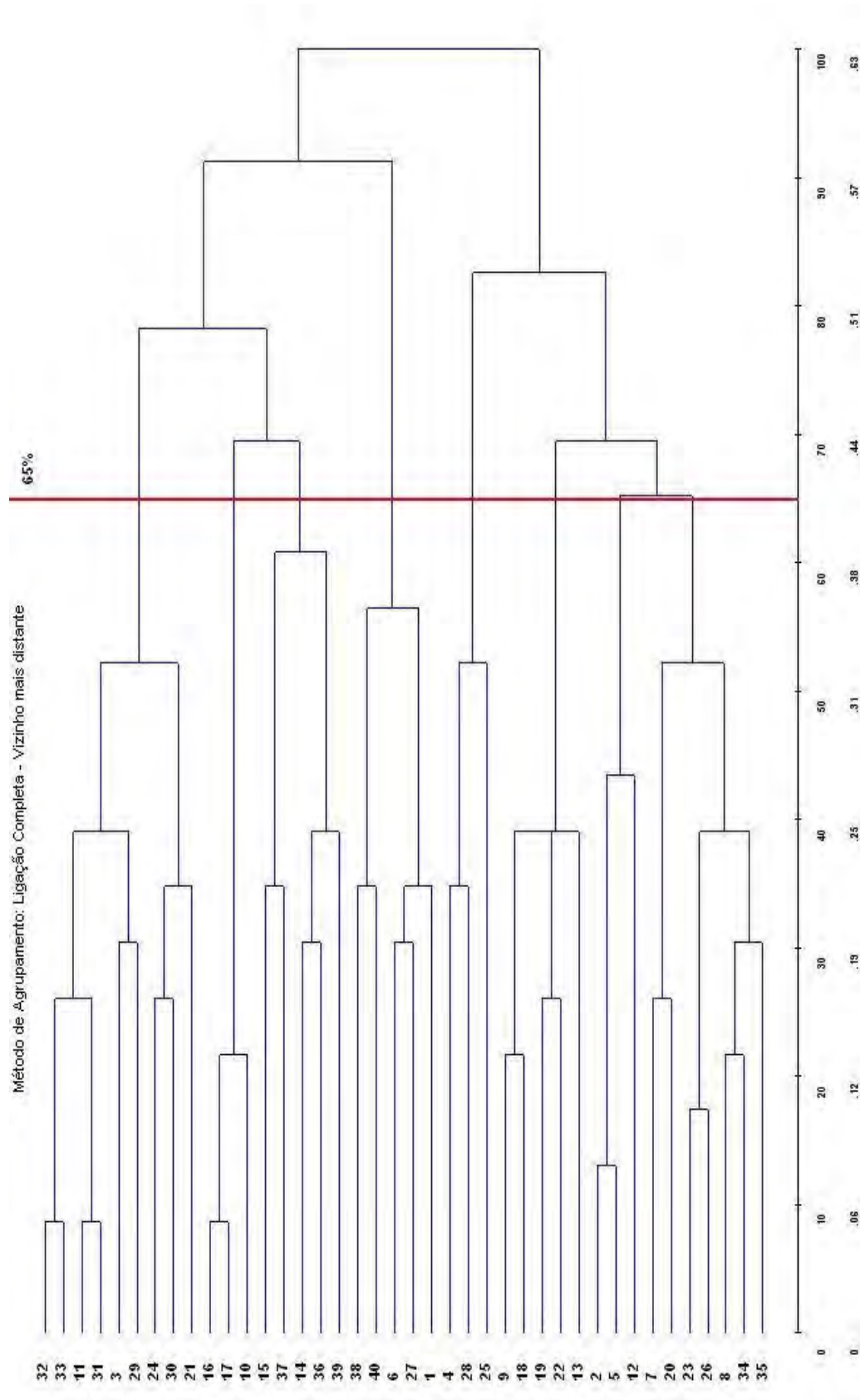


Figura 5. Representação do dendrograma obtido através do método de Ligação Completa – Vizinho Mais Distante, a partir do complemento aritmético do Coeficiente de Coincidência Simples.

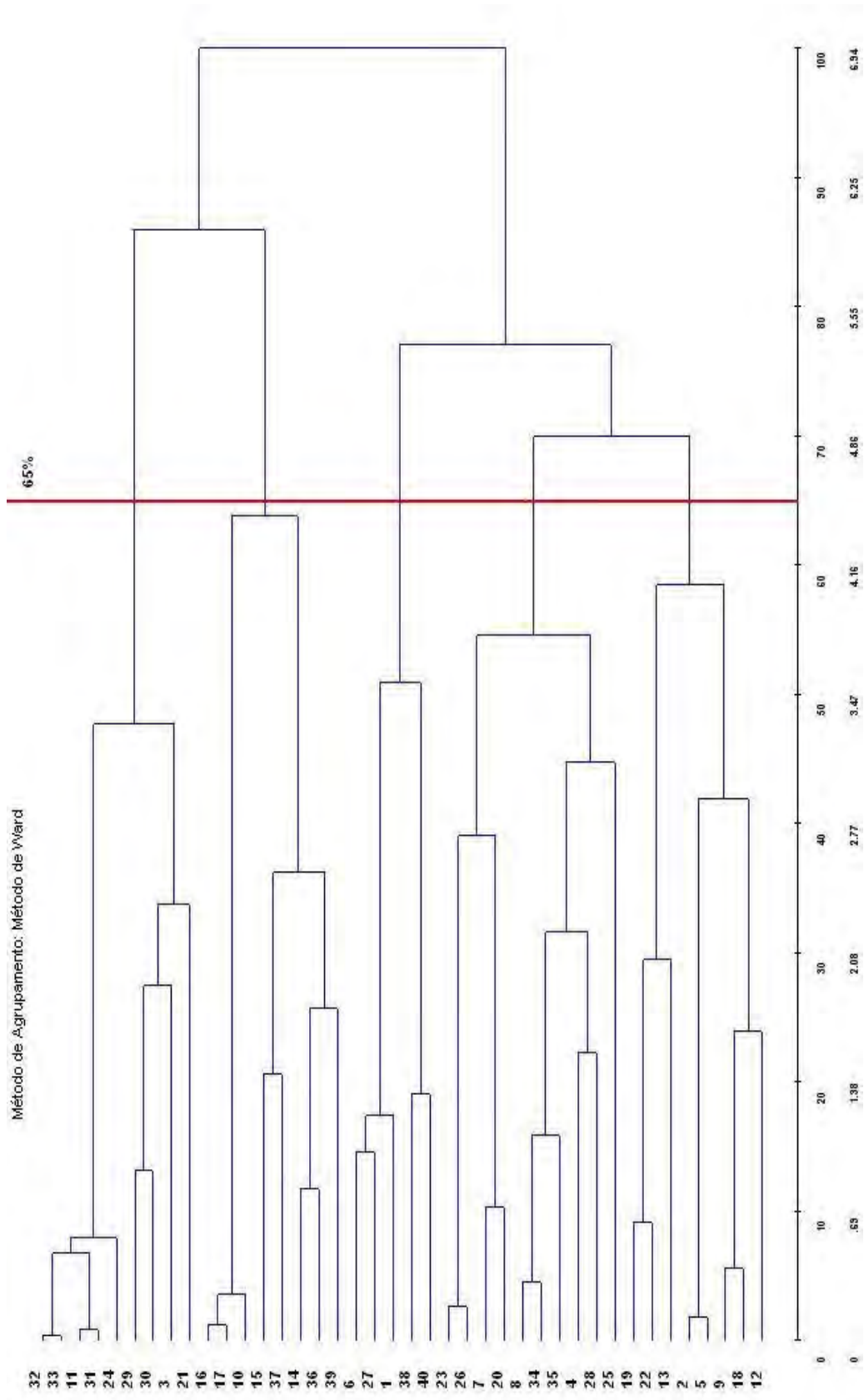


Figura 6. Representação do dendrograma obtido através do método de Ward, a partir do complemento aritmético do Coeficiente de Coincência Simples.

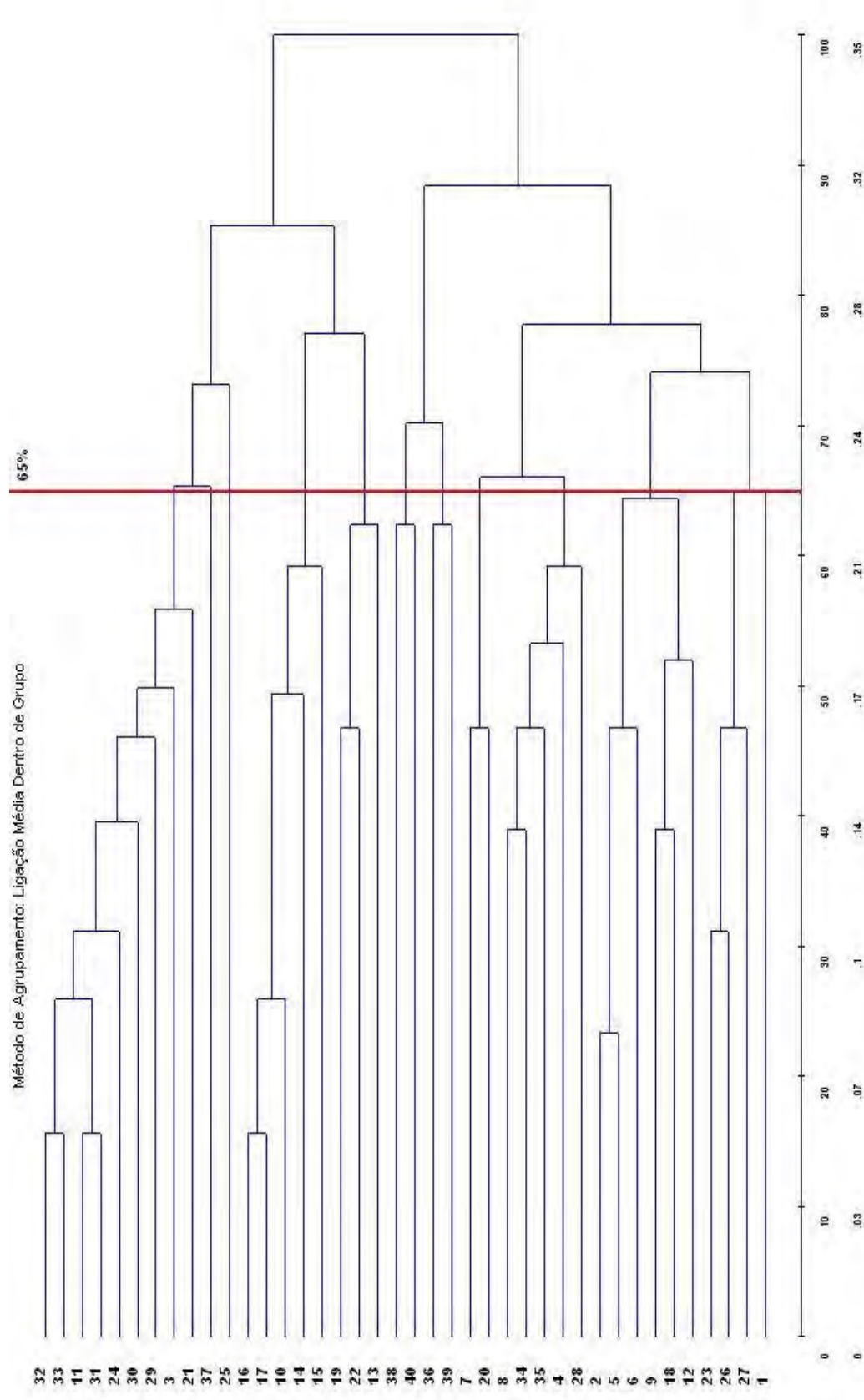


Figura 7. Representação do dendrograma obtido através do método de Ligação Média dentro do grupo, a partir do complemento aritmético do Coeficiente de Coincidência Simples.

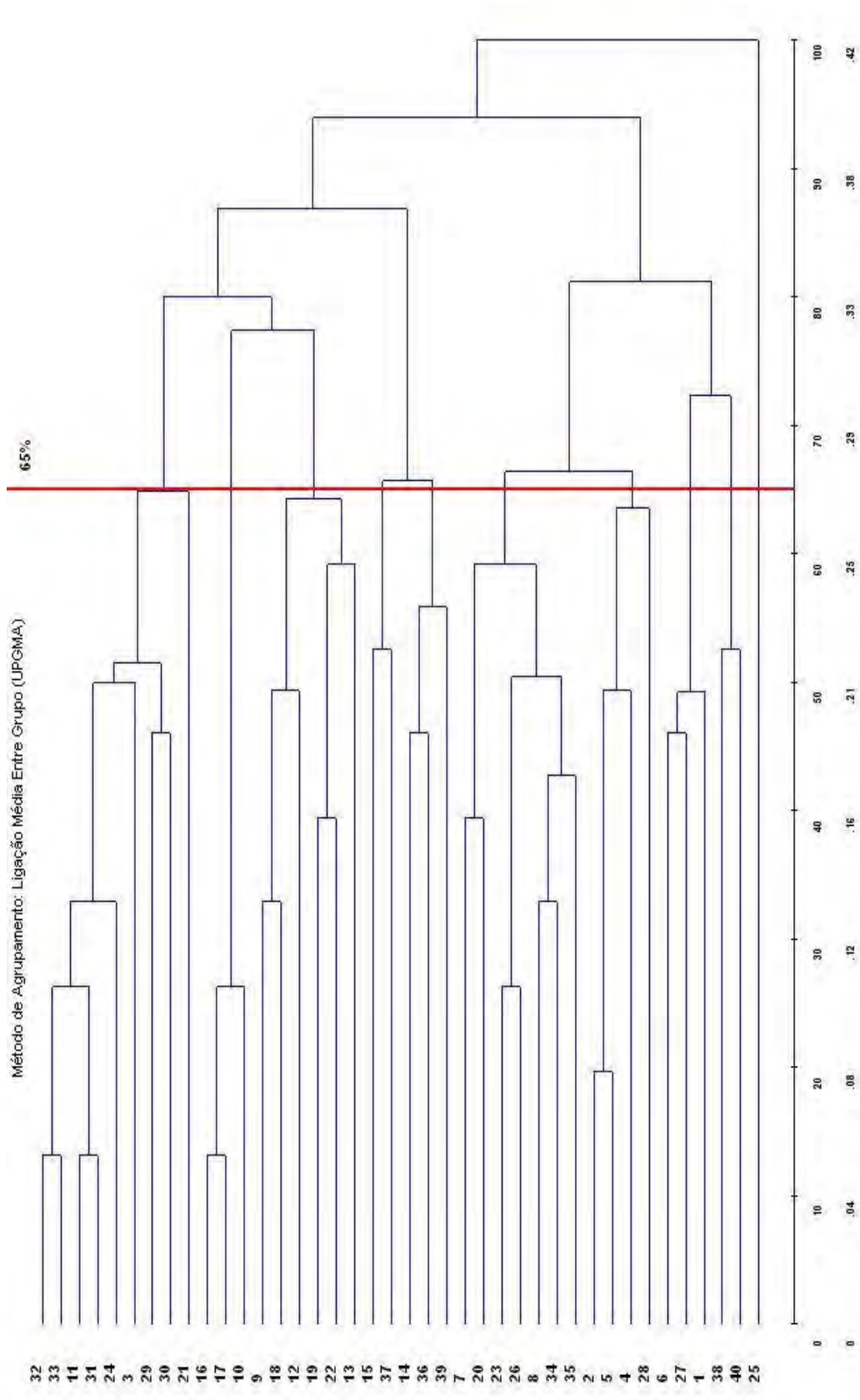


Figura 8. Representação do dendrograma obtido através do método de UPGMA, a partir do complemento aritmético do Coeficiente de Coincidência Simples.

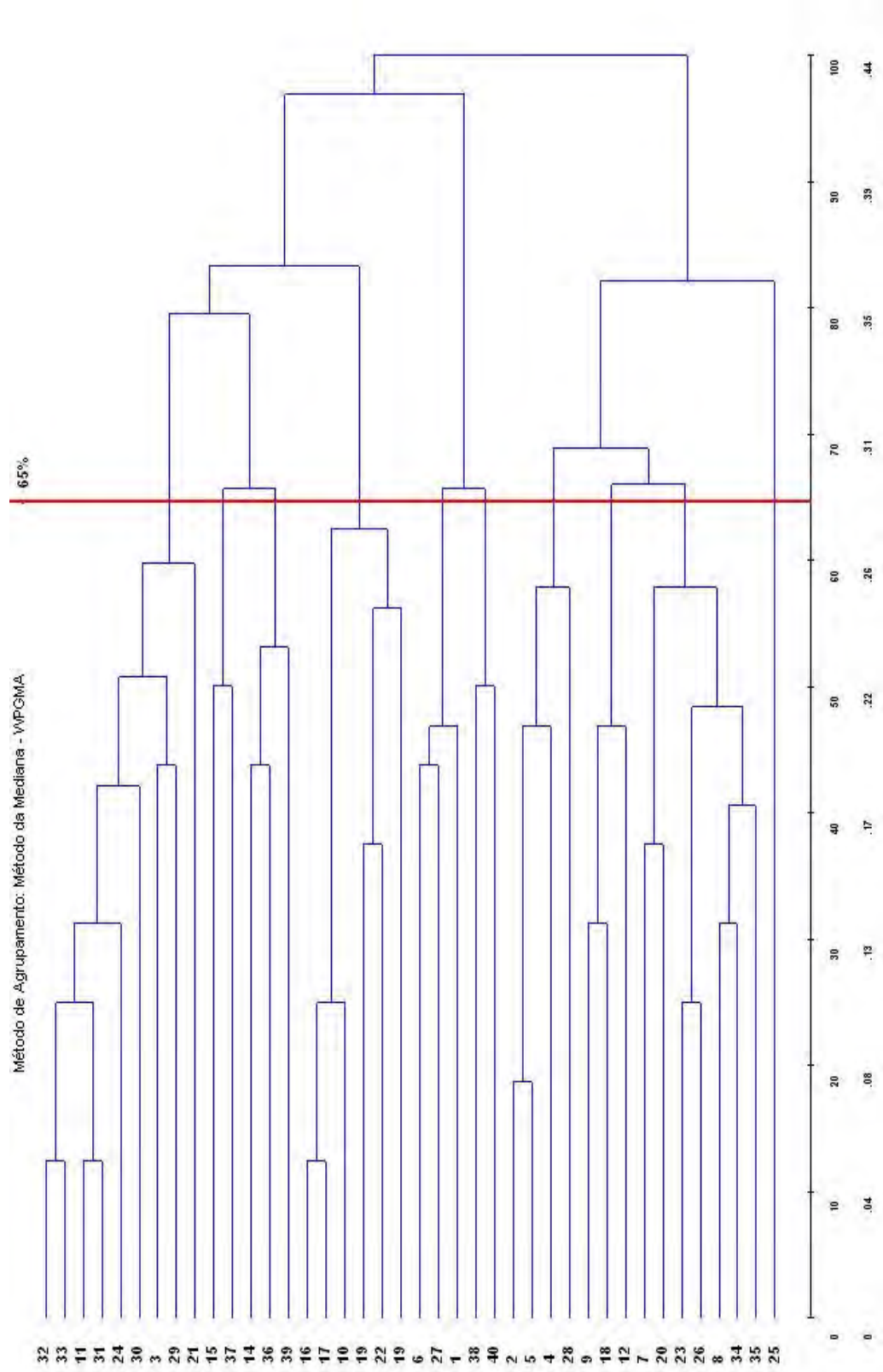


Figura 9. Representação do dendrograma obtido através do método WPGMA, a partir do complemento aritmético do Coeficiente de Coincência Simples.

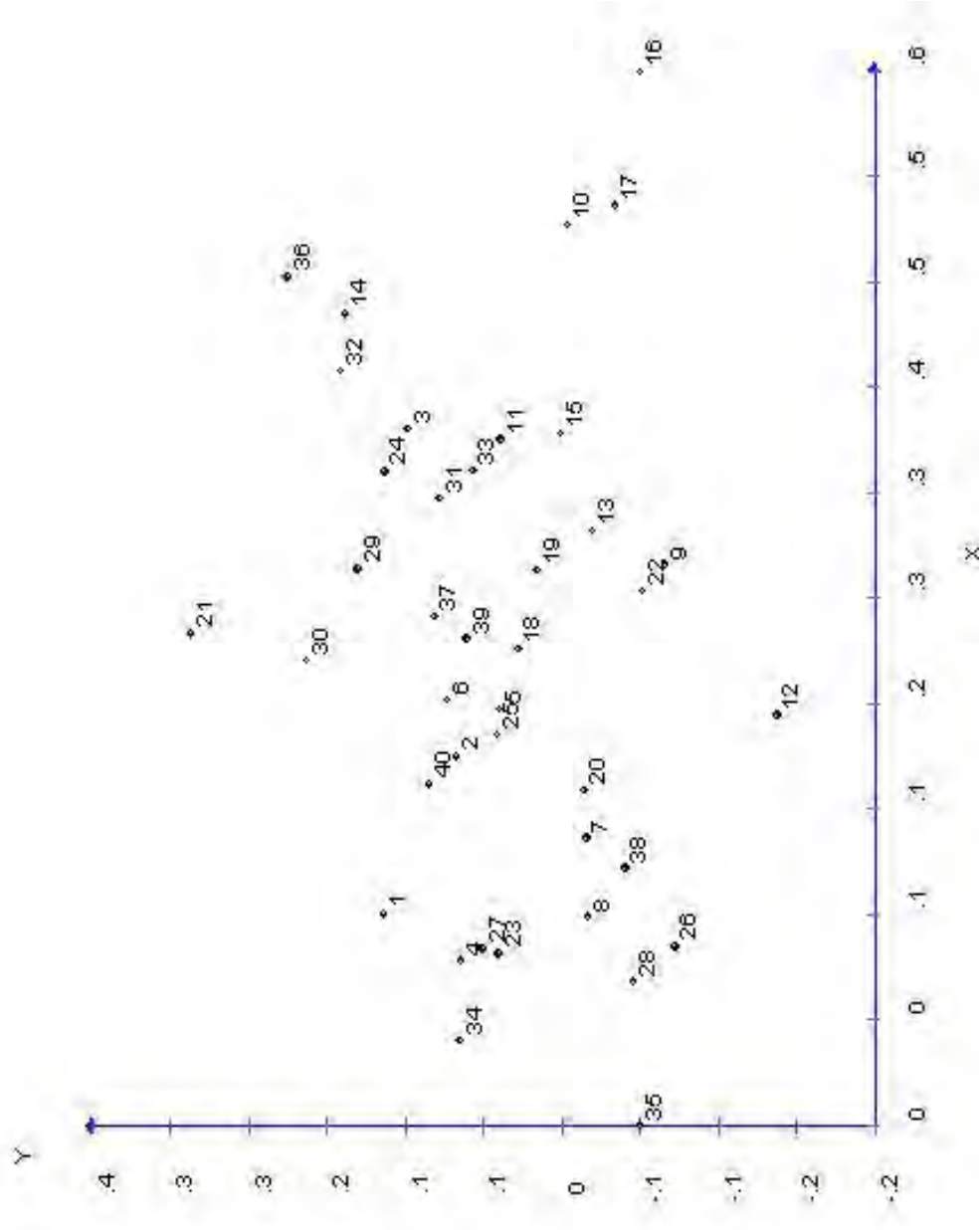


Figura 10.Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Coincidência Simples, no espaço bidimensional.

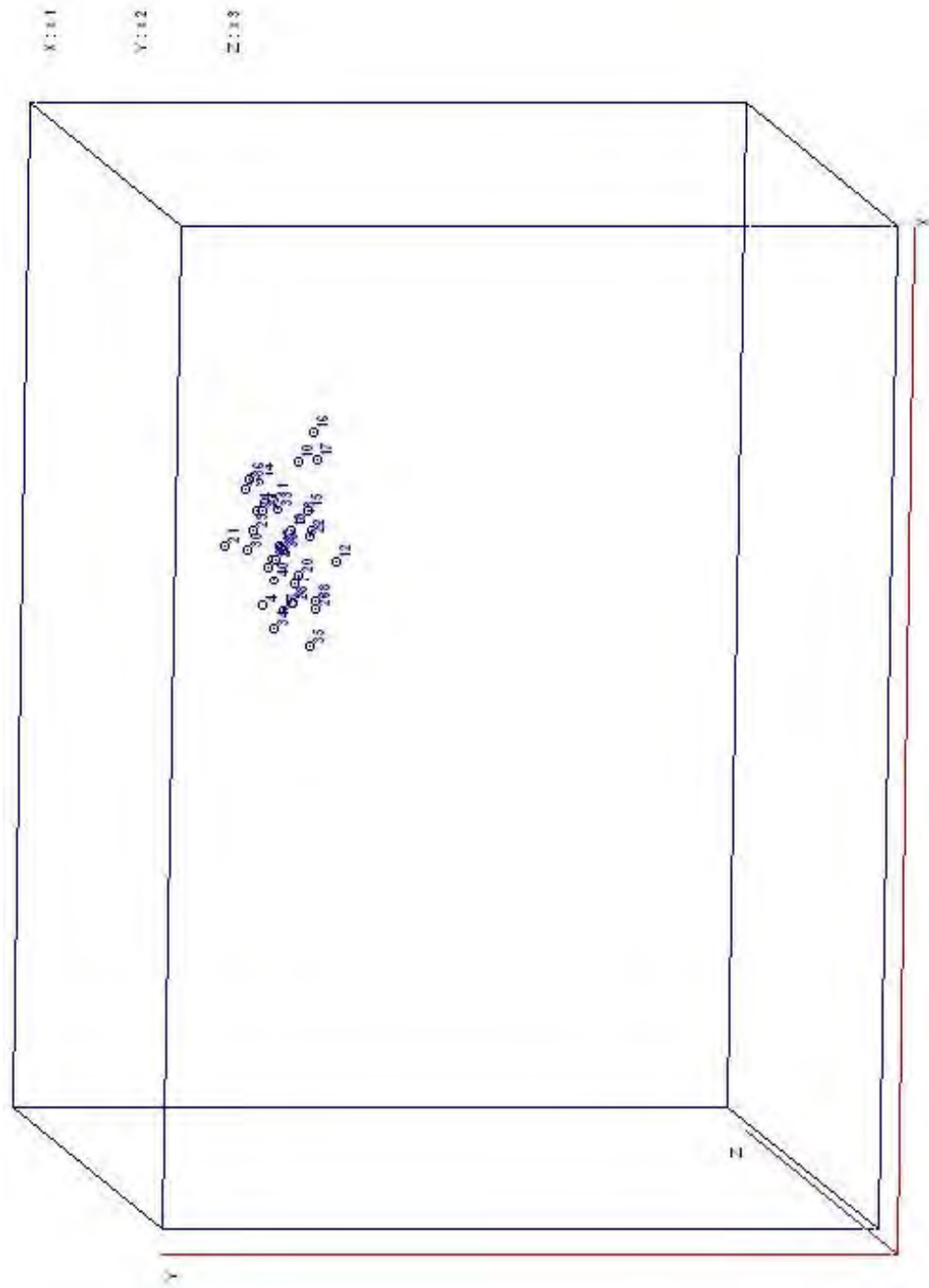


Figura 11. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Coincidência Simples, no espaço tridimensional (3D com rotação).

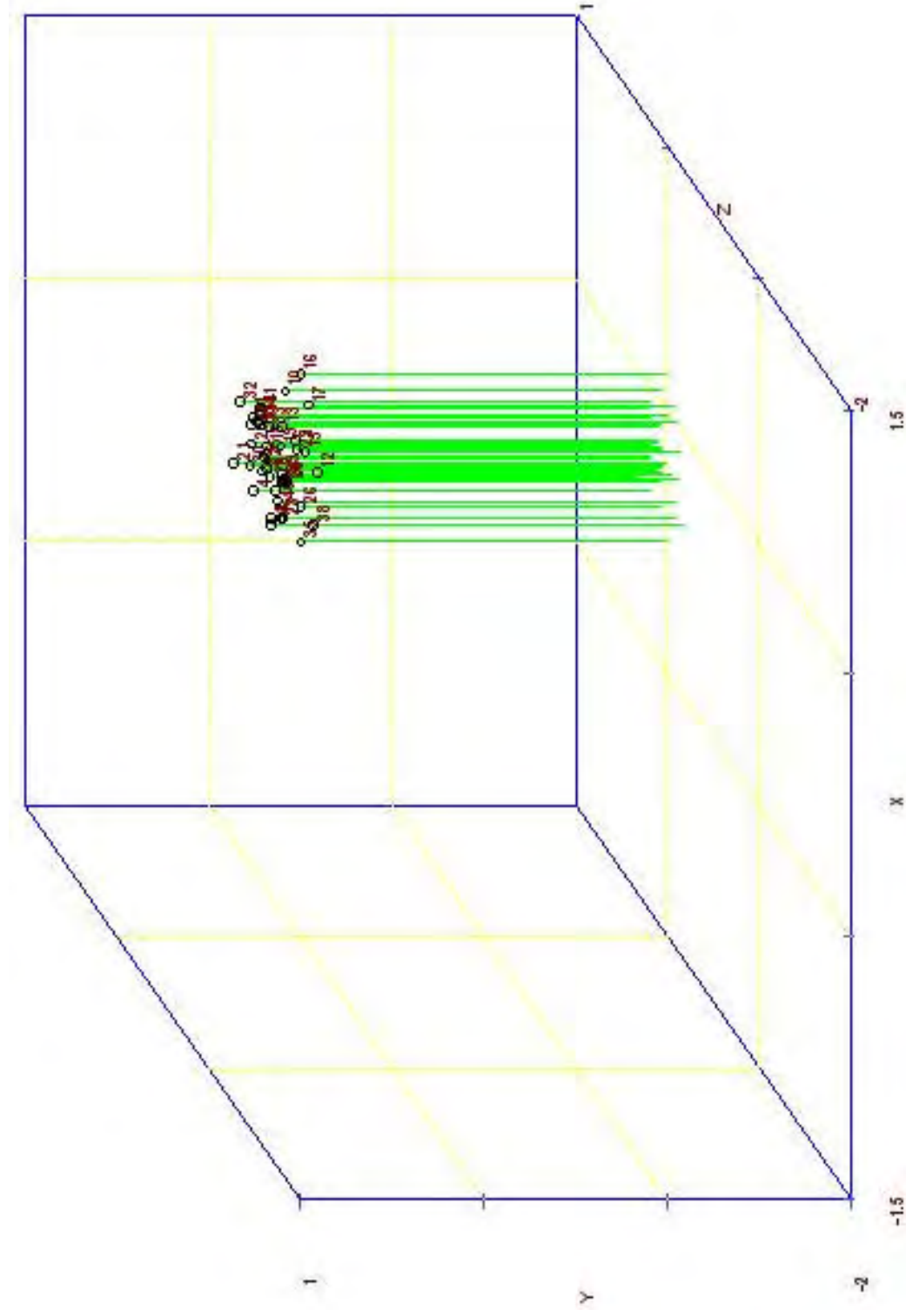


Figura 12. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Coincidência Simples, no espaço tridimensional (3D sem rotação).

Após as análises de agrupamentos foi calculada a matriz dos mais e menos similares, com a finalidade de Identificar, para cada acesso, aqueles outros que lhes são mais ou menos similares. São fornecidas as médias gerais e do grupo de acesso considerado. Foram constatadas algumas coincidências, ainda, entre os agrupamentos hierárquicos e a matriz dos genótipos mais similares (Tabela 5), havendo confirmação de alguns agrupamentos observados nos dendrogramas com os dados obtidos pela matriz que revela os genótipos mais similares.

4.2. Coeficientes de Similaridade pelo Método de Rogers e Tanimoto

A média aritmética do complemento do coeficiente de similaridade obtido pelo coeficiente de Rogers e Tanimoto entre todos os genótipos foi de 0,5033 (Tabela 6), sendo pouco superior ao valor encontrado pelo coeficiente de Coincidência Simples. Os pares de genótipos mais divergentes foram: 4 (Tambaqui) e 16 (M-SOY 8400), 34 (JAB99174/9.1) e 16 (M-SOY 8400), 35 (JAB99174/9.15) e 16 (M-SOY 8400) e o genótipo 28 (Egyptiam) com os genótipos 15 (M-SOY 8200), 16 (M-SOY 8400) e 17 (M-SOY 8329), com coeficiente de dissimilaridade igual a 0,7797. Os pares de genótipos mais similares foram: 16 (M-SOY 8400) e 17 (M-SOY 8329), 11 (Nobreza) e 31 (JAB99102/4.5), 31 (JAB99102/4.5) e 32 (JAB99102/8.1), 32 (JAB99102/8.1) e 33 (JAB99103/8.2), com coeficientes de dissimilaridade igual a 0,11. Os resultados foram coincidentes com o método anterior (Coincidência Simples).

Após análise de dissimilaridade, foram realizados os agrupamentos dos 40 genótipos resistentes ao NCS, raça 3, estudados, adotando-se um percentual de similaridade genética em torno de 65%, também utilizada no Coeficiente de Coincidência Simples.

Tabela 5. Dez genótipos mais similares, em ordem decrescente, em relação a cada um dos 40 genótipos estudados com fonte de resistência ao nematóide de cisto da soja, através do complemento aritmético do Coeficiente de Similaridade de Coincidência Simples.

Gen^{1/}	Mais similares	Gen	Mais similares
1	6, 27, 38, 30, 23, 34, 40, 21, 5, 29	21	24, 30, 14, 29, 11, 18, 19, 22, 31, 32
2	5, 6, 8, 24, 4, 23, 34, 11, 12, 18	22	19, 24, 9, 10, 11, 18, 33, 13, 20, 23
3	32, 29, 11, 24, 31, 33, 14, 15, 30, 36	23	26, 27, 34, 5, 20, 35, 2, 8, 12, 22
4	5, 34, 8, 7, 25, 28, 2, 26, 35, 6	24	11, 32, 22, 30, 31, 33, 5, 29, 2, 3
5	2, 8, 4, 6, 11, 12, 18, 24, 9, 23	25	4, 8, 11, 31, 5, 28, 18, 21, 30, 32
6	1, 5, 27, 2, 8, 18, 24, 30, 31, 32	26	23, 34, 12, 5, 7, 9, 27, 35, 8, 4
7	20, 34, 8, 4, 26, 6, 18, 5, 9, 28	27	23, 6, 1, 26, 34, 8, 18, 19, 22, 30
8	5, 28, 34, 18, 4, 7, 35, 2, 6, 12	28	8, 4, 5, 7, 34, 2, 12, 18, 9, 25
9	12, 18, 22, 5, 13, 26, 8, 19, 7, 10	29	3, 24, 30, 32, 11, 18, 19, 21, 22, 31
10	17, 16, 11, 22, 13, 14, 19, 24, 31, 32	30	24, 32, 33, 29, 6, 21, 31, 1, 23, 27
11	31, 24, 32, 33, 5, 10, 3, 22, 29, 2	31	11, 32, 33, 24, 3, 6, 18, 30, 5, 10
12	9, 5, 26, 8, 23, 2, 18, 7, 27, 28	32	31, 33, 11, 24, 3, 30, 29, 6, 18, 5
13	9, 10, 18, 19, 22, 33, 5, 11, 17, 21	33	32, 31, 11, 24, 30, 3, 19, 22, 10, 13
14	36, 10, 3, 15, 16, 21, 37, 39, 20, 29	34	8, 4, 7, 23, 26, 35, 5, 27, 2, 6
15	17, 37, 14, 3, 16, 32, 33, 36, 10, 29	35	34, 8, 23, 26, 4, 7, 20, 27, 38, 12
16	17, 10, 14, 15, 36, 11, 22, 9, 13, 3	36	14, 37, 39, 40, 3, 15, 16, 10, 29, 38
17	16, 10, 15, 11, 22, 9, 13, 14, 12, 19	37	15, 36, 14, 38, 3, 39, 34, 40, 2, 21
18	9, 8, 5, 6, 19, 22, 31, 32, 7, 13	38	1, 40, 37, 39, 34, 35, 36, 27, 2, 6
19	22, 18, 33, 9, 10, 13, 20, 27, 29, 11	39	36, 14, 20, 38, 40, 37, 23, 34, 2, 8
20	7, 23, 8, 19, 22, 39, 14, 26, 27, 29	40	38, 36, 39, 1, 2, 6, 37, 14, 23, 27

^{1/}Gen = genótipos estudados

Tabela 6. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Similaridade de Rogers e Tanimoto.

Genótipo	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0,53	0,64	0,56	0,50	0,33	0,56	0,53	0,67	0,67	0,64	0,59	0,71	0,62	0,59	0,74	0,69	0,53	0,64	0,56
2		0,50	0,40	0,15	0,36	0,53	0,36	0,47	0,59	0,43	0,43	0,53	0,59	0,71	0,62	0,67	0,43	0,62	0,59
3			0,53	0,53	0,56	0,53	0,62	0,64	0,53	0,36	0,71	0,59	0,40	0,43	0,56	0,62	0,56	0,50	0,59
4				0,29	0,47	0,36	0,33	0,56	0,67	0,53	0,53	0,56	0,67	0,64	0,78	0,74	0,47	0,64	0,50
5					0,33	0,43	0,24	0,36	0,50	0,33	0,33	0,43	0,62	0,64	0,64	0,59	0,33	0,53	0,50
6						0,40	0,36	0,53	0,47	0,43	0,50	0,59	0,53	0,56	0,62	0,56	0,36	0,56	0,47
7							0,33	0,43	0,56	0,53	0,47	0,56	0,56	0,59	0,69	0,64	0,40	0,53	0,29
8								0,40	0,59	0,43	0,36	0,53	0,64	0,67	0,71	0,67	0,29	0,50	0,40
9									0,43	0,47	0,24	0,36	0,56	0,59	0,53	0,47	0,24	0,40	0,50
10										0,33	0,59	0,36	0,36	0,47	0,24	0,15	0,47	0,40	0,50
11											0,56	0,47	0,47	0,50	0,50	0,43	0,43	0,43	0,53
12												0,53	0,69	0,56	0,56	0,50	0,43	0,56	0,53
13													0,56	0,64	0,53	0,47	0,40	0,40	0,56
14														0,40	0,40	0,47	0,53	0,47	0,43
15															0,43	0,36	0,62	0,56	0,59
16																0,11	0,62	0,56	0,64
17																	0,56	0,50	0,59
18																		0,36	0,47
19																			0,40

Continua...

Tabela 6. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Similaridade de Rogers e Tanimoto (Continuação).

Genótipo	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	0,47	0,64	0,43	0,53	0,62	0,56	0,36	0,67	0,50	0,40	0,59	0,53	0,59	0,43	0,50	0,53	0,59	0,36	0,53	0,43
2	0,50	0,50	0,40	0,36	0,59	0,47	0,47	0,47	0,59	0,50	0,50	0,50	0,56	0,40	0,53	0,56	0,50	0,53	0,50	0,47
3	0,50	0,56	0,64	0,36	0,59	0,59	0,64	0,69	0,33	0,43	0,36	0,29	0,36	0,53	0,69	0,43	0,43	0,59	0,56	0,59
4	0,53	0,64	0,50	0,53	0,36	0,43	0,56	0,36	0,56	0,53	0,53	0,53	0,59	0,29	0,43	0,69	0,59	0,56	0,53	0,62
5	0,47	0,47	0,36	0,33	0,50	0,36	0,43	0,43	0,50	0,47	0,40	0,40	0,47	0,36	0,50	0,59	0,59	0,62	0,53	0,56
6	0,50	0,43	0,47	0,36	0,64	0,53	0,33	0,53	0,47	0,36	0,36	0,36	0,43	0,40	0,53	0,62	0,56	0,53	0,56	0,47
7	0,53	0,47	0,50	0,53	0,62	0,36	0,50	0,43	0,43	0,53	0,53	0,53	0,59	0,29	0,43	0,69	0,64	0,56	0,53	0,62
8	0,56	0,50	0,40	0,50	0,47	0,40	0,40	0,24	0,47	0,56	0,43	0,50	0,56	0,24	0,33	0,67	0,62	0,59	0,50	0,53
9	0,53	0,33	0,43	0,47	0,62	0,36	0,50	0,50	0,50	0,53	0,53	0,53	0,53	0,50	0,56	0,53	0,53	0,67	0,53	0,62
10	0,53	0,33	0,62	0,40	0,67	0,62	0,62	0,71	0,50	0,53	0,40	0,40	0,40	0,67	0,71	0,47	0,59	0,71	0,59	0,67
11	0,43	0,36	0,59	0,20	0,47	0,59	0,59	0,59	0,40	0,43	0,11	0,20	0,29	0,59	0,64	0,50	0,62	0,74	0,67	0,64
12	0,67	0,50	0,40	0,56	0,59	0,33	0,47	0,47	0,59	0,62	0,62	0,62	0,62	0,47	0,47	0,62	0,62	0,59	0,56	0,59
13	0,47	0,40	0,56	0,47	0,62	0,50	0,56	0,62	0,56	0,53	0,47	0,47	0,40	0,62	0,62	0,64	0,59	0,76	0,64	0,71
14	0,40	0,47	0,62	0,53	0,62	0,67	0,67	0,71	0,43	0,47	0,47	0,47	0,47	0,67	0,71	0,33	0,40	0,62	0,40	0,50
15	0,56	0,56	0,64	0,50	0,59	0,59	0,64	0,78	0,47	0,50	0,50	0,43	0,43	0,59	0,64	0,43	0,36	0,53	0,56	0,64
16	0,67	0,50	0,74	0,56	0,69	0,74	0,74	0,78	0,64	0,67	0,56	0,56	0,56	0,78	0,78	0,43	0,56	0,64	0,56	0,64

Continua...

Tabela 6. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Similaridade de Rogers e Tanimoto (Continuação).

Genótipo	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
17	0,62	0,43	0,69	0,50	0,64	0,69	0,69	0,78	0,59	0,62	0,50	0,50	0,50	0,74	0,74	0,50	0,62	0,69	0,62	0,69
18	0,43	0,36	0,47	0,43	0,53	0,47	0,40	0,47	0,40	0,43	0,36	0,36	0,43	0,40	0,53	0,62	0,56	0,69	0,56	0,64
19	0,43	0,29	0,53	0,50	0,64	0,53	0,40	0,64	0,40	0,56	0,43	0,43	0,36	0,59	0,59	0,50	0,62	0,69	0,62	0,59
20	0,53	0,40	0,36	0,59	0,56	0,43	0,43	0,56	0,43	0,47	0,53	0,53	0,47	0,43	0,43	0,59	0,69	0,62	0,40	0,56
21		0,43	0,53	0,36	0,53	0,64	0,53	0,69	0,40	0,36	0,43	0,43	0,50	0,53	0,64	0,50	0,50	0,64	0,62	0,53
22			0,40	0,29	0,69	0,47	0,40	0,64	0,40	0,43	0,43	0,43	0,36	0,53	0,53	0,56	0,56	0,69	0,67	0,64
23				0,47	0,62	0,20	0,29	0,56	0,43	0,40	0,64	0,59	0,53	0,29	0,36	0,53	0,53	0,56	0,47	0,50
24					0,59	0,53	0,47	0,64	0,33	0,29	0,29	0,20	0,29	0,47	0,64	0,50	0,50	0,64	0,67	0,64
25						0,67	0,71	0,50	0,56	0,53	0,47	0,53	0,59	0,56	0,56	0,64	0,69	0,71	0,59	0,67
26							0,36	0,50	0,50	0,53	0,64	0,59	0,53	0,29	0,36	0,59	0,53	0,56	0,53	0,62
27								0,56	0,50	0,40	0,53	0,47	0,40	0,36	0,43	0,59	0,59	0,50	0,59	0,50
28									0,62	0,69	0,59	0,64	0,69	0,43	0,50	0,74	0,69	0,67	0,59	0,62
29										0,33	0,40	0,33	0,40	0,43	0,62	0,47	0,53	0,67	0,59	0,56
30											0,36	0,29	0,29	0,47	0,59	0,50	0,50	0,59	0,56	0,53
31												0,11	0,20	0,59	0,64	0,56	0,62	0,74	0,67	0,64
32														0,11	0,53	0,69	0,50	0,69	0,62	0,69

Continua...

Tabela 6. Medidas de dissimilaridade entre 40 genótipos de soja resistentes ao nematóide de cisto, calculadas através do complemento aritmético do Coeficiente de Similaridade de Rogers e Tanimoto (Continuação).

[illegible]

O Método de Otimização de Tocher (Tabela 7) identifica o par de indivíduos mais similares para formar o grupo inicial e a partir daí avalia a possibilidade de inclusão de novos indivíduos com o critério que a distância média intragrupo deve ser menor que a distância média intergrupo (CRUZ & CARNEIRO, 2006). Apresentou a formação de 17 grupos, idênticos aos grupos apresentados pelo Coeficiente de Coincidência Simples. Os grupos mais divergentes, foram 2 e 15 englobando os genótipos: 16 (M-SOY 8400), 17 (M-SOY 8329) e 10 (BRSMG Liderança); e 15 (M-SOY 8200) e 37 (JAB 99102/4.2), respectivamente, com coeficiente de dissimilaridade igual a 0,76. Os grupos mais semelhantes foram 9 e 12, com 0,38 de dissimilaridade. O grupo 9 foi formado pelos genótipos 14 (M-SOY 8001) e 36 (JAB 99101/2.5) e o grupo 12 composto por 39 (JAB 994311/15.20). Nota-se que os agrupamentos realizados a partir das dissimilaridades de Rogers e Tanimoto foram exatamente iguais aos resultados obtidos pelo método de Coincidência Simples, com alteração apenas nos valores médios das distâncias intra e intergrupos, o que acarretou em alterações nos grupos mais e menos similares.

No método hierárquico de Ligação Simples – Vizinho mais Próximo (Figura 13) as conexões entre grupos são feitas por ligações simples entre pares de indivíduos, ou seja, as distâncias entre grupos são definidas como sendo aquela entre os indivíduos mais parecidos entre esses grupos (CRUZ & CARNEIRO, 2006). Foram formados 32 grupos, sendo 28 destes compostos por um só genótipo (1, 3, 4, 6, 7, 8, 9, 12, 13, 14, 15, 18, 19, 20, 21, 22, 25, 27, 28, 29, 30, 34, 35, 36, 37, 38, 39, 40). Um grupo maior englobou os genótipos 11 (Nobreza), 24 (98N71), 31 (JAB 99102/4.2), 32 (JAB 99102/8.1) e 33 (JAB 99103/8.2). Houve ainda a formação de mais dois grupos: um formado pelos genótipos 10 (BRSMG Liderança), 16 (M-SOY 8400) e 17 (M-SOY 8329) e outro pelos genótipos 23 (98N41) e 26 (Foster). Tais resultados apresentaram-se bem diferentes daqueles obtidos pelo coeficiente de Coincidência Simples.

Tabela 7. Formação de grupos através do método de otimização de Tocher, utilizando o complemento aritmético do Coeficiente de Rogers e Tanimoto dos 40 genótipos de soja resistentes ao NCS estudados.

Grupo	Indivíduos
I	11, 31, 32, 33, 24, 30, 3, 29
II	16, 17, 10
III	2, 5, 8, 34, 4
IV	23, 26, 27
V	9, 12, 18
VI	7, 20
VII	19, 22
VIII	1, 6
IX	14, 36
X	15, 37
XI	38, 40
XII	39
XIII	35
XIV	21
XV	28
XVI	25
XVII	13

O método Ligação Completa – Vizinho Mais Distante (Figura 14) define a similaridade entre dois grupos como aquela apresentada pelos indivíduos de cada grupo que menos se identificam (MEYER, 2002). A partir desse método, 10 agrupamentos foram formados, abrangendo de 2 a 9 genótipos, revelando alguns grupos bem semelhantes ao mesmo agrupamento utilizando o complemento aritmético do Coeficiente de Coincidência Simples. Alguns grupos apresentaram-se iguais ao

método aplicado anteriormente, tais como: I (32, 33, 11, 31, 3, 29, 24, 30, 21); II (16, 17, 10); III (4, 28, 25); IV (9, 18, 19, 22, 13); V (2, 5, 12); VI (7, 20, 23, 26, 8, 34, 35).

O método de Ward (Figura 15) considera para formação inicial do grupo, aqueles indivíduos que proporcionam a menor soma de quadrados dos desvios (CRUZ & CARNEIRO, 2006). Tal método revelou 6 grandes grupos: I (32, 33, 11, 31, 24, 29, 30, 3, 21); II (16, 17, 10); III (15, 37, 14, 36, 39); IV (9, 18, 12, 19, 22, 13); V (6, 27, 1, 38, 40) e VI (23, 26, 7, 20, 2, 5, 8, 34, 35, 4, 28, 25).

A Ligação Média Dentro de Grupo (Figura 16) calcula a similaridade média dos indivíduos do grupo e depois os coeficientes de similaridade do indivíduo que pretende se juntar ao grupo. Este método apresentou 18 grupos com muita distribuição dos genótipos e muitos grupos formados por apenas um genótipo. Alguns genótipos mantiveram-se agrupados de forma semelhante aos outros métodos, mas com alterações na composição geral dos grupos.

O agrupamento por UPGMA (Figura 17) é, recentemente, o mais utilizado nos estudos de divergência genética (DIAS, 1998; CRUZ & CARNEIRO, 2006), este algoritmo não considera a estrutura de subdivisão do grupo, dando pesos iguais a cada indivíduo do grupo e calcula a similaridade média de um indivíduo que pretende se juntar ao grupo já existente (MEYER, 2002). Tal método propôs a formação de 15 grupos, assim como o WPGMA (Figura 18), que difere do método anterior por ser não ponderado, ou seja, não aferir pesos diferentes aos indivíduos que já se uniram o grupo e aos que pretendem se unir. Ambos, UPGMA e WPGMA apresentaram um padrão de agrupamento completamente semelhante, com formação de grupos iguais, fato também observado em partes em relação ao Coeficiente de Coincidência Simples.

Na maioria das vezes, verifica-se que, apesar da estrutura final mostrar-se parecida, ocorreram algumas alterações nos níveis em que os genótipos são agrupados. Entretanto, apesar de tal fato, existem genótipos que se mantiveram no mesmo grupo na maioria dos métodos aglomerativos, como os genótipos 10 (BRSMG Liderança), 16 (M-SOY 8400), 17 (M-SOY 8329), todos com fonte de resistência de Peking; os genótipos 3 (BRS Piraíba), 11 (Nobreza), 24 (98N71), 29 (Leflore), 30 (Jeff), 31 (JAB 99102/4.5), 32 (JAB 99102/8.1), 33 (JAB 99103/8.2), com a cultivar Peking

envolvida na genealogia; os genótipos 9 (FMT Cachara), 12 (Robusta), 18 (M-SOY 8757), também oriundos de Peking; os genótipos 2 (BRS Jiripoca) e 5 (FMT Tucunaré), compartilhando Peking e PI 437654 em sua genealogia; os genótipos 19 (BRSMT Pintado) e 22 (MGBR 9520937); entre outros.

As projeções das distâncias das dissimilaridades calculadas pelo Coeficiente de Rogers e Tanimoto no espaço bidimensional e tridimensional apresentaram-se menos satisfatórias do que as obtidas pelo coeficiente de Coincidência Simples. A projeção 2D (bidimensional) revelou uma distorção de 47% e estresse de 51,8%, com correlações entre as distâncias originais e as estimadas de 0,55 (Figura 19). Já as projeções 3D (tridimensionais) apresentaram uma menor distorção, de 34,95% e estresse de 39,85%; onde a correlação entre as distâncias originais e as estimadas foi de 0,66 (Figuras 20 e 21).

Enquanto que as projeções 2D apresentaram uma certa dispersão dos genótipos, com os indivíduos 12 (Robusta), 16 (M-SOY 8400), 21 (Hartwig) e 35 (JAB99174/9.15) mais afastados dos demais; as projeções tridimensionais no plano revelaram uma grande proximidade entre todos os genótipos. Podem-se verificar claramente os genótipos 10 (Liderança), 16 (M-SOY 8400) e 17 (M-SOY 8329) mais unidos entre si e um pouco mais afastados dos outros, concordando com métodos de agrupamento hierárquicos e de otimização. Novamente o genótipo 21 (Hartwig) aparece mais distanciado dos grupos, assim como o genótipo 12 (Robusta) e o 35 (JAB99174/9.15).

A matriz dos genótipos mais similares revelou muitas coincidências de agrupamento, sendo útil como complemento para as análises (Tabela 8).

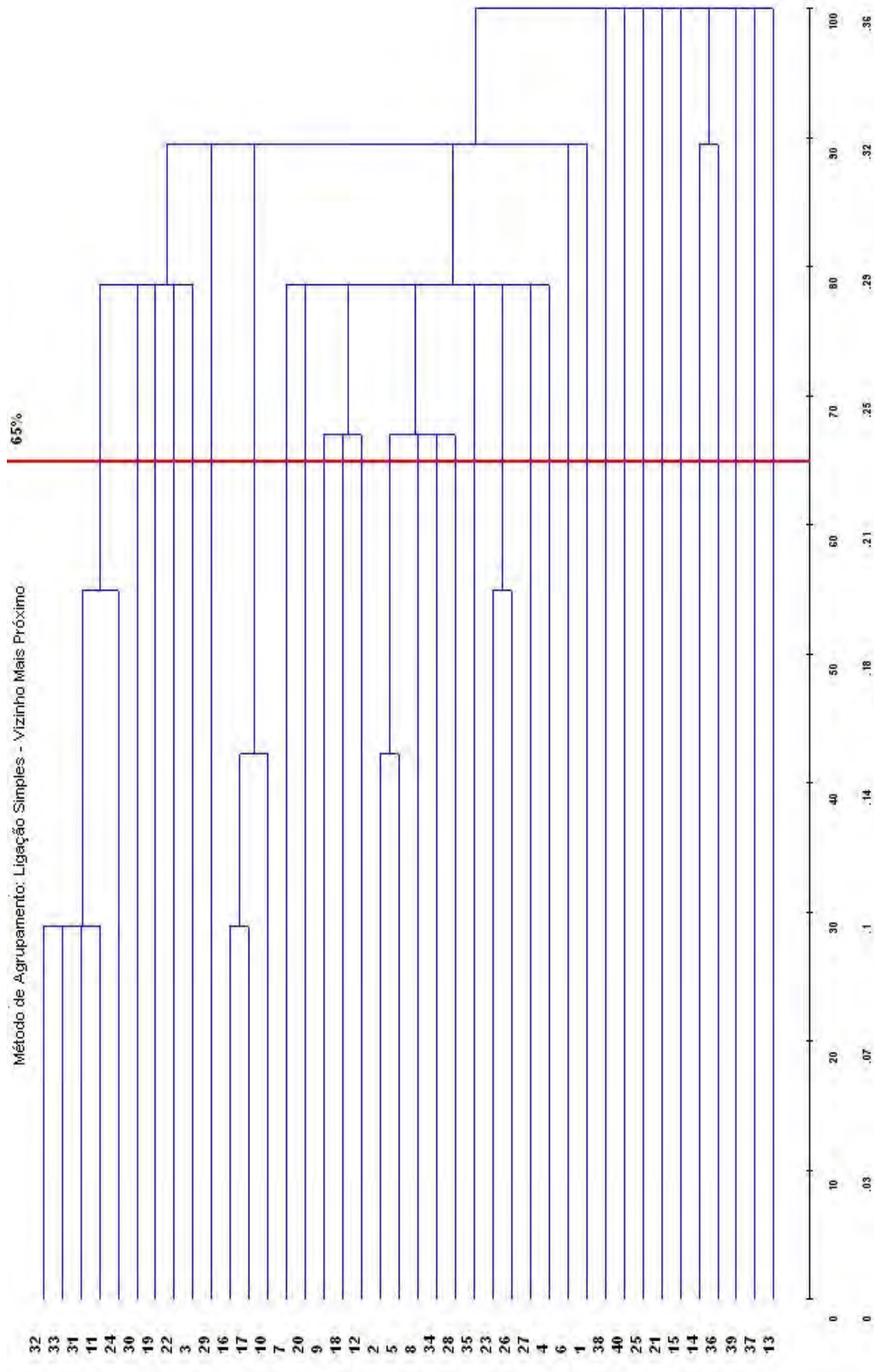


Figura 13.Representação do dendrograma obtido através do Método de Ligação Simples – Vizininho Mais Próximo, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.

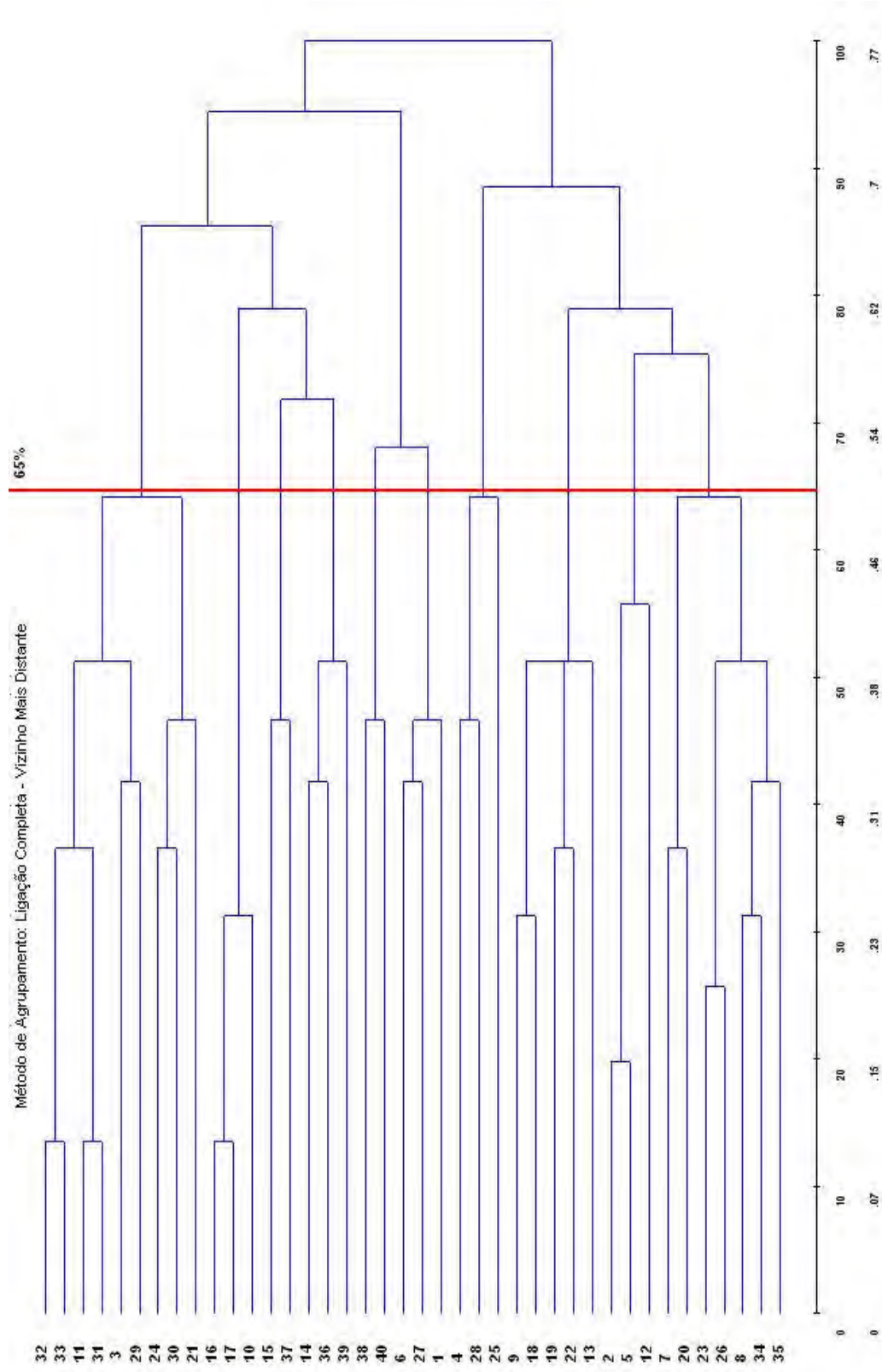


Figura 14. Representação do dendrograma obtido através do Método de Ligação Completa - Vizinheiro mais Distante, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.

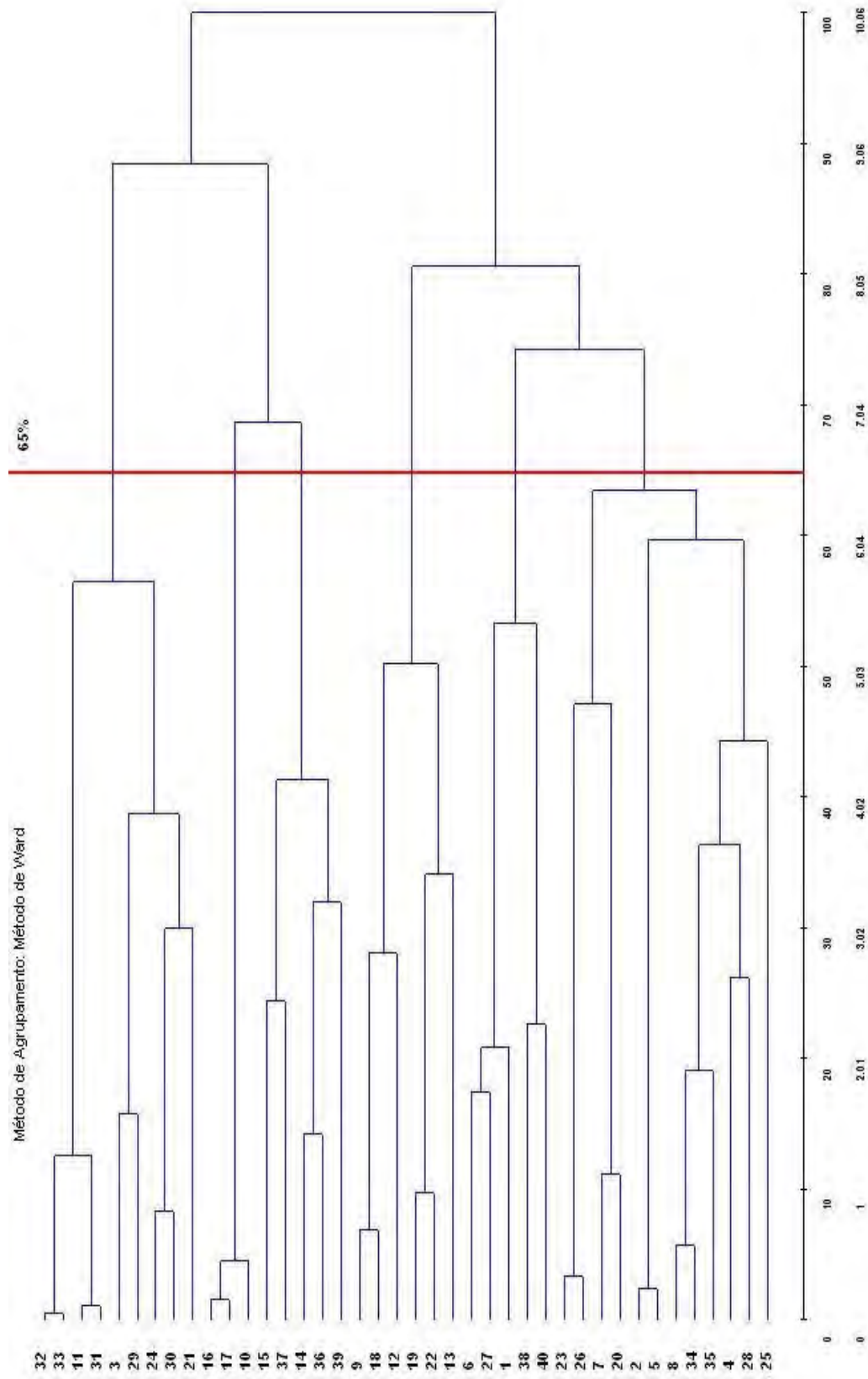


Figura 15. Representação do dendrograma obtido pelo Método de Ward, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.

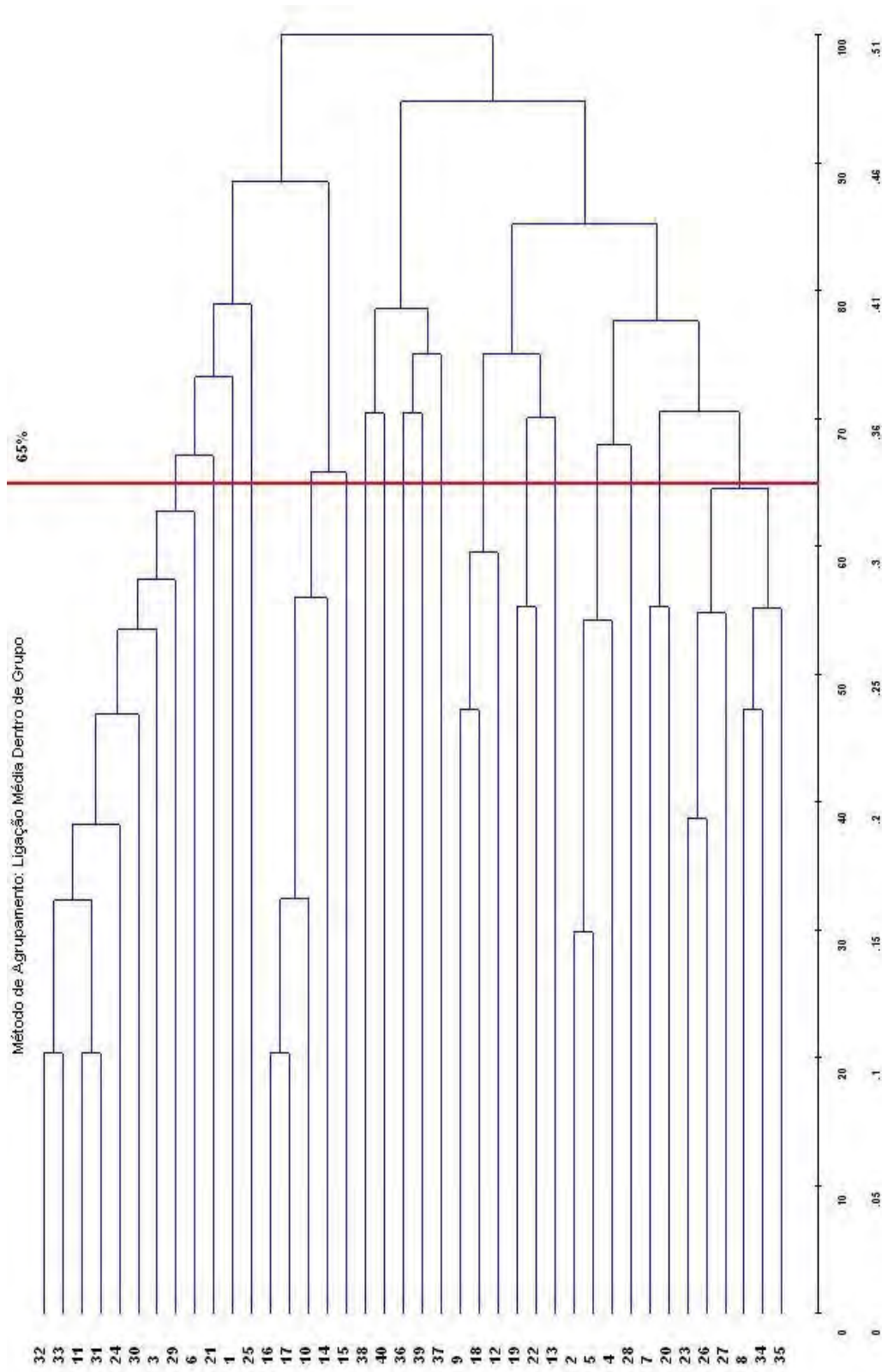


Figura 16. Representação do dendrograma obtido através do Método de Ligação Média Dentro de Grupo, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.

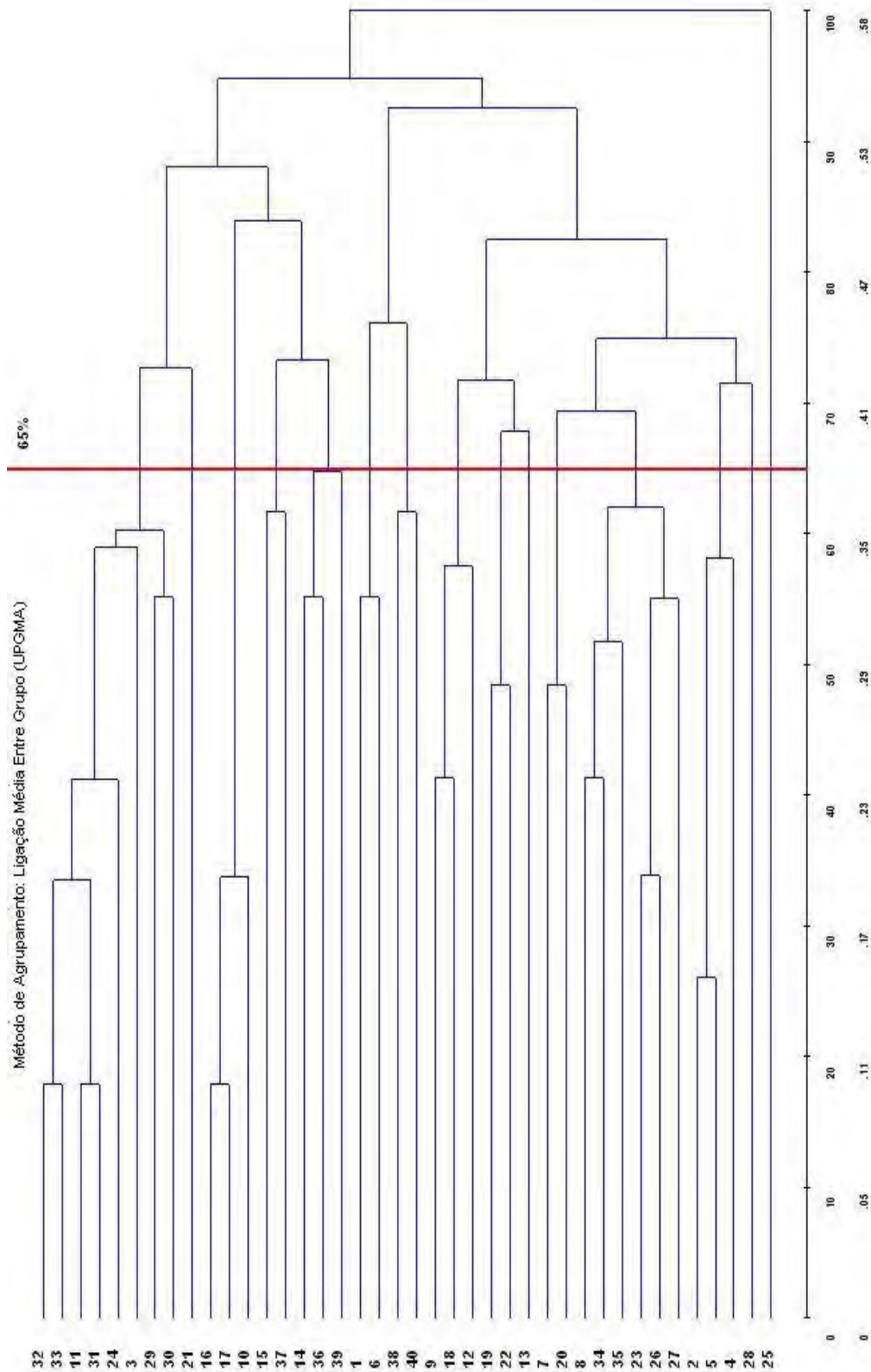


Figura 17. Representação do dendrograma obtido através do Método de Agrupamento UPGMA, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.

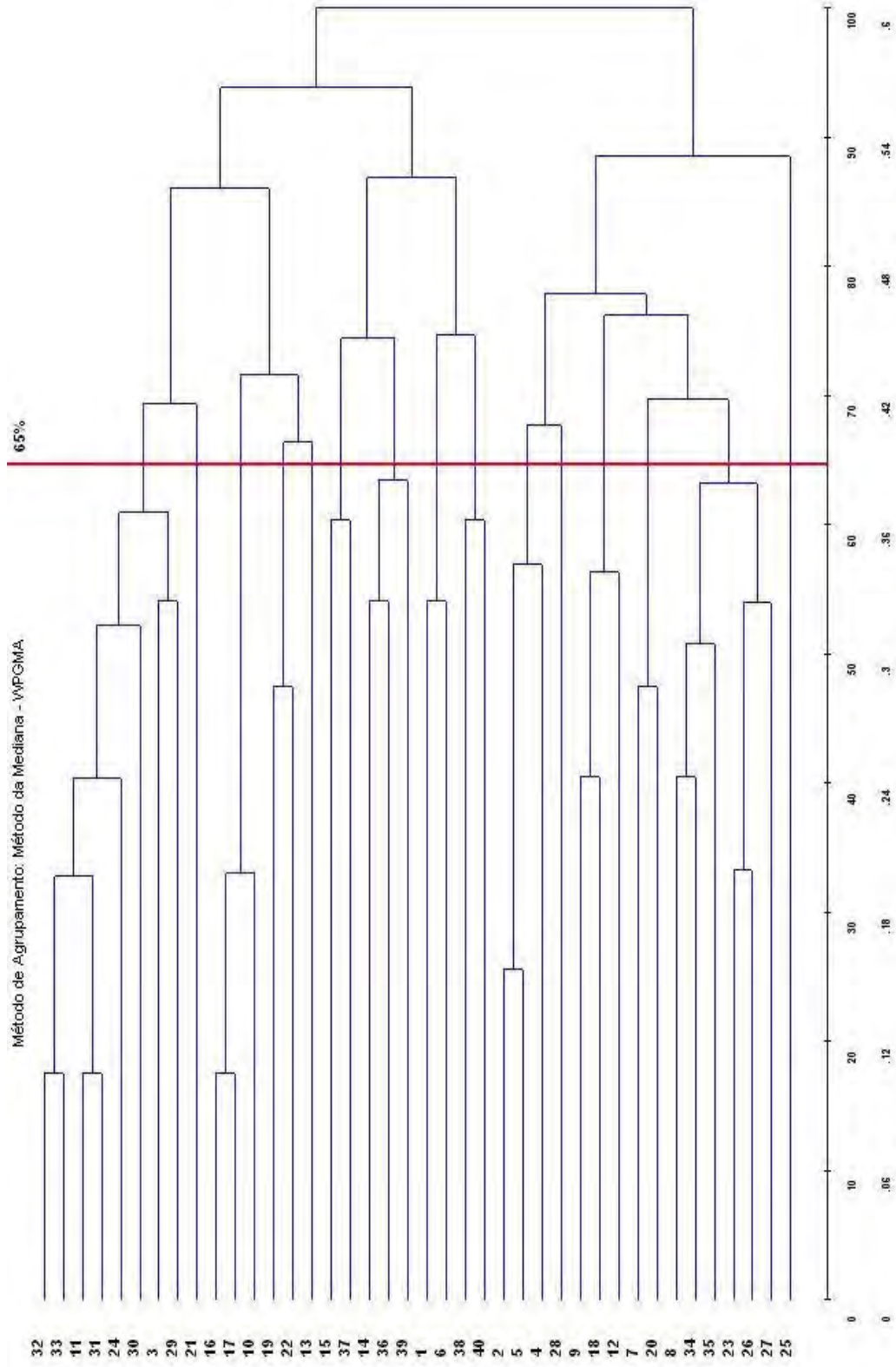


Figura 18. Representação do dendrograma obtido pelo Método da Mediana - WPGMA, a partir do complemento aritmético do Coeficiente de Rogers e Tanimoto.

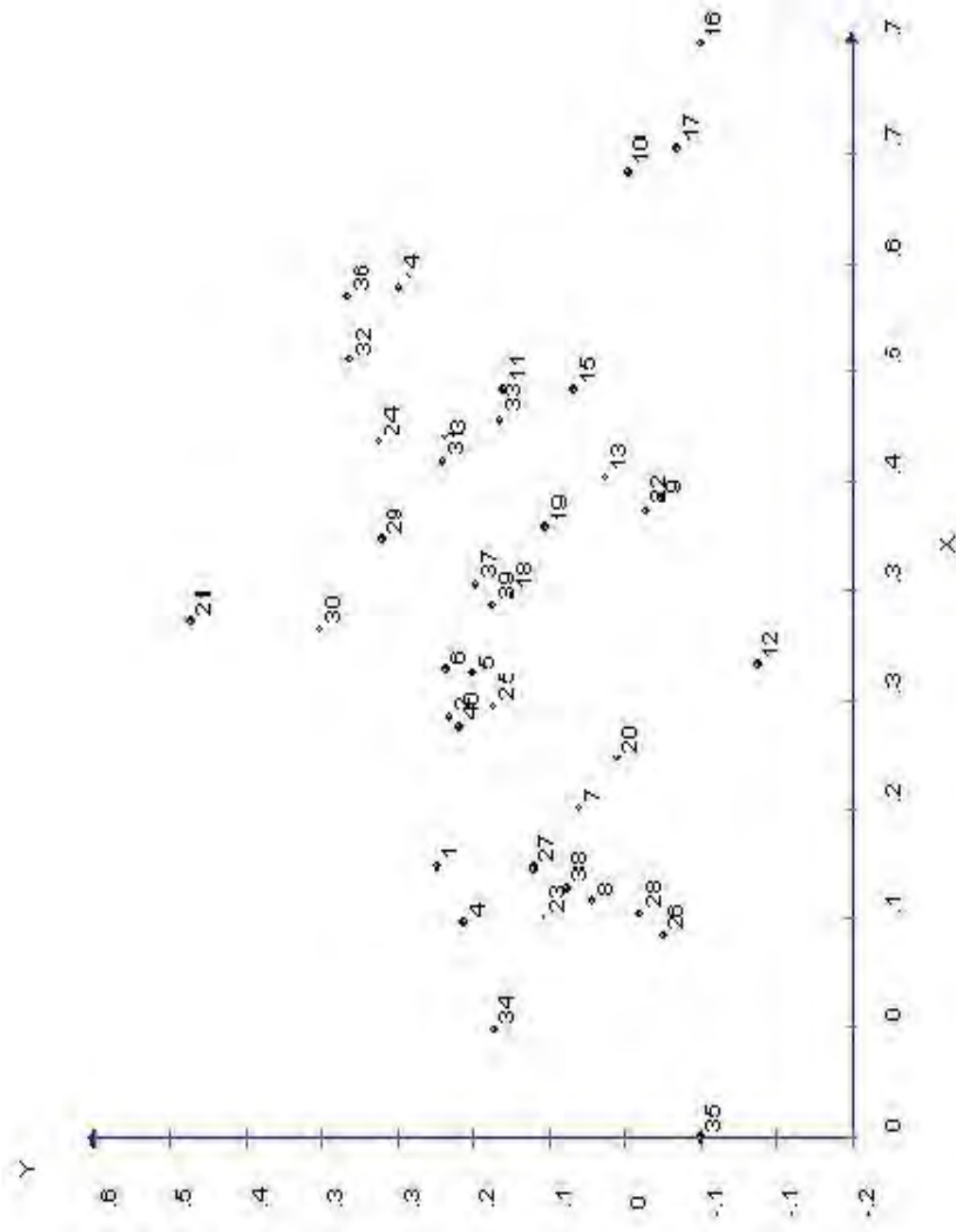


Figura 19. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Rogers e Tanimoto, no espaço bidimensional.

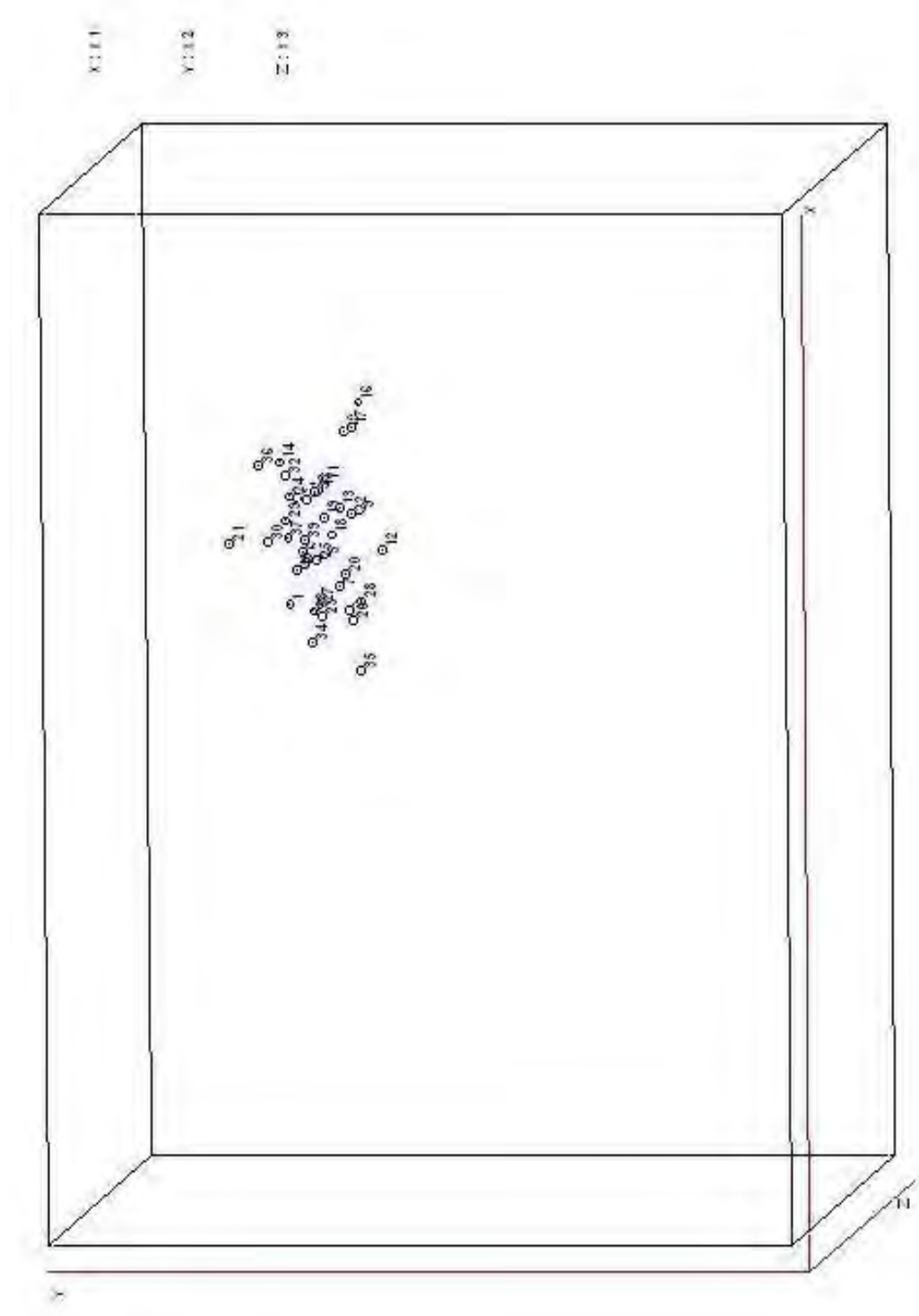


Figura 20. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Rogers e Tanimoto, no espaço tridimensional (3D com rotação).

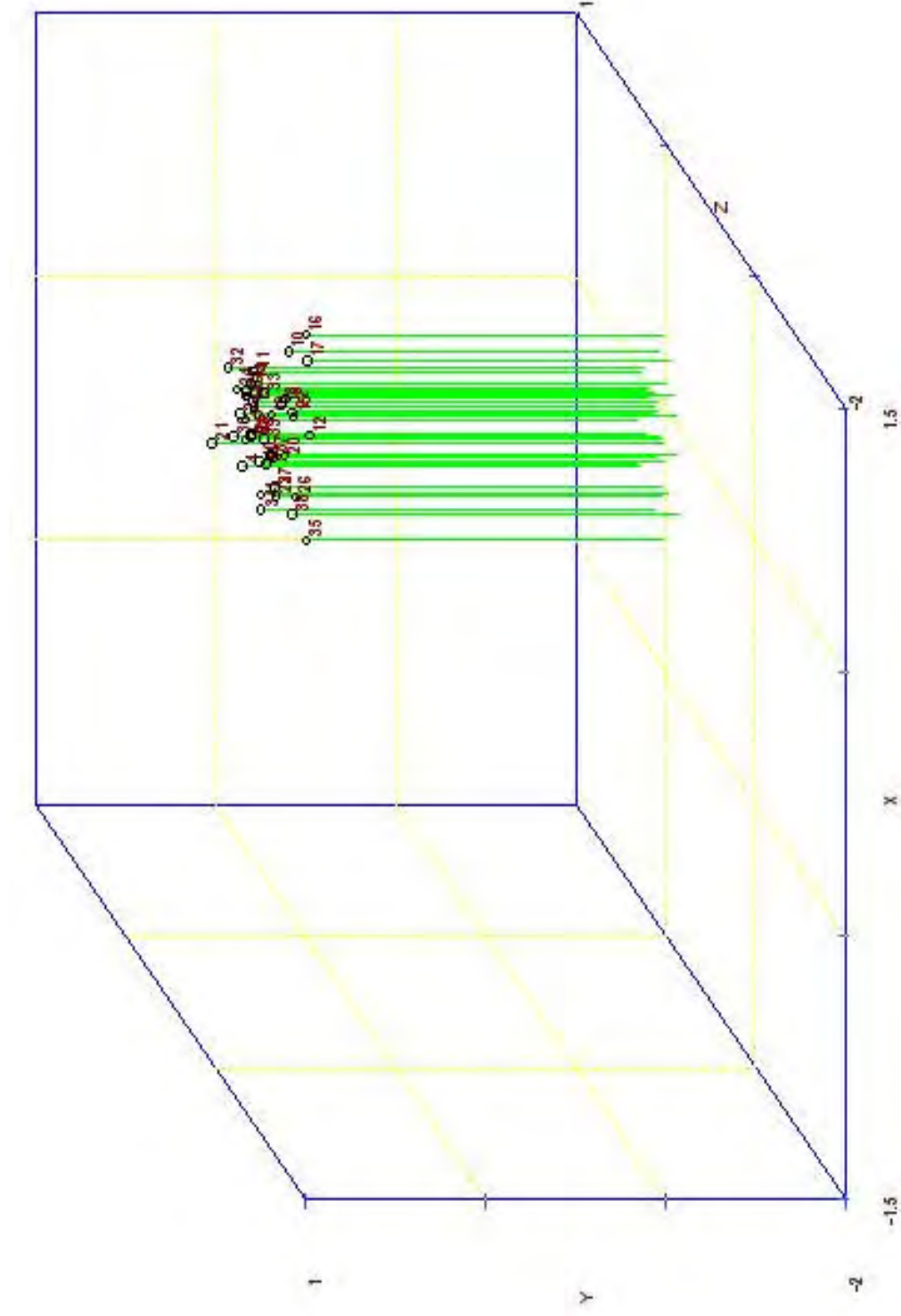


Figura 21. Projeção da dissimilaridade entre 40 genótipos de soja resistentes ao NCS, expressa pelo complemento aritmético do Coeficiente de Rogers e Tanimoto, no espaço tridimensional (3D sem rotação).

Tabela 8. Dez genótipos mais similares, em ordem decrescente, em relação a cada um dos 40 genótipos estudados com fonte de resistência ao nematóide de cisto da soja, através do Coeficiente de Rogers e Tanimoto.

Gen^{1/}	Mais similares	Gen	Mais similares
1	6, 27, 38, 30, 23, 34, 40, 21, 5, 29	21	24, 30, 14, 29, 11, 18, 19, 22, 31, 32
2	5, 6, 8, 24, 4, 23, 34, 11, 12, 18	22	19, 24, 9, 10, 11, 18, 33, 13, 20, 23
3	32, 29, 11, 24, 31, 33, 14, 15, 30, 36	23	26, 27, 34, 5, 20, 35, 2, 8, 12, 22
4	5, 34, 8, 7, 25, 28, 2, 26, 35, 6	24	11, 32, 22, 30, 31, 33, 5, 29, 2, 3
5	2, 8, 4, 6, 11, 12, 18, 24, 9, 23	25	4, 8, 11, 31, 5, 28, 18, 21, 30, 32
6	1, 5, 27, 2, 8, 18, 24, 30, 31, 32	26	23, 34, 12, 5, 7, 9, 27, 35, 8, 4
7	20, 34, 8, 4, 26, 6, 18, 5, 9, 28	27	23, 6, 1, 26, 34, 8, 18, 19, 22, 30
8	5, 28, 34, 18, 4, 7, 35, 2, 6, 12	28	8, 4, 5, 7, 34, 2, 12, 18, 9, 25
9	12, 18, 22, 5, 13, 26, 8, 19, 7, 10	29	3, 24, 30, 32, 11, 18, 19, 21, 22, 31
10	17, 16, 11, 22, 13, 14, 19, 24, 31, 32	30	24, 32, 33, 29, 6, 21, 31, 1, 23, 27
11	31, 24, 32, 33, 5, 10, 3, 22, 29, 2	31	11, 32, 33, 24, 3, 6, 18, 30, 5, 10
12	9, 5, 26, 8, 23, 2, 18, 7, 27, 28	32	31, 33, 11, 24, 3, 30, 29, 6, 18, 5
13	9, 10, 18, 19, 22, 33, 5, 11, 17, 21	33	32, 31, 11, 24, 30, 3, 19, 22, 10, 13
14	36, 10, 3, 15, 16, 21, 37, 39, 20, 29	34	8, 4, 7, 23, 26, 35, 5, 27, 2, 6
15	17, 37, 14, 3, 16, 32, 33, 36, 10, 29	35	34, 8, 23, 26, 4, 7, 20, 27, 38, 12
16	17, 10, 14, 15, 36, 11, 22, 9, 13, 3	36	14, 37, 39, 40, 3, 15, 16, 10, 29, 38
17	16, 10, 15, 11, 22, 9, 13, 14, 12, 19	37	15, 36, 14, 38, 3, 39, 34, 40, 2, 21
18	9, 8, 5, 6, 19, 22, 31, 32, 7, 13	38	1, 40, 37, 39, 34, 35, 36, 27, 2, 6
19	22, 18, 33, 9, 10, 13, 20, 27, 29, 11	39	36, 14, 20, 38, 40, 37, 23, 34, 2, 8
20	7, 23, 8, 19, 22, 39, 14, 26, 27, 29	40	38, 36, 39, 1, 2, 6, 37, 14, 23, 27

^{1/}Gen = genótipos estudados

4.3. Correlação entre Matrizes

A correlação entre as matrizes de dissimilaridade obtidas pelos Coeficientes de Coincidência Simples e de Rogers e Tanimoto revelou-se significativa a 1% de probabilidade pelo Teste t e pelo Teste de Mantel baseado em 1000 simulações, demonstrando-se altamente correlacionadas, com valor de correlação igual a 0,99409 (Figura 22).

O emprego dos dois coeficientes causou poucas alterações no agrupamento das cultivares, o que também foi observado por DUARTE et al. (1999), na comparação de oito coeficientes de similaridade, baseados em marcadores RAPD, em feijão; e ROCHA (2000), que comparou oito coeficientes de similaridade por meio de marcadores RAPD em batata e não observou alteração no agrupamento dos genótipos estudados.

4.4. Considerações Finais

A média aritmética dos coeficientes de dissimilaridade obtidos através do complemento aritmético dos coeficientes de Coincidência Simples e Rogers e Tanimoto apresentaram valores de reduzidos a médios, 0,36 e 0,50, respectivamente. PRIOLLI et al. (2002), encontraram em seus estudos coeficientes de dissimilaridade entre 0,4 e 0,9, quando realizam-se comparações entre cultivares brasileiras de soja, através de marcadores SSR. Os valores médios de similaridade também variaram em outros estudos, como em SMITH et al. (2000), que evidenciou um valor médio de similaridade de 0,38, utilizando marcadores RFLP e 0,36 utilizando marcadores SSR, em sorgo. Por sua vez, ABDELNOOR et al. (1995) obtiveram média de similaridade de 0,82 em estudos com marcadores RAPD em soja.

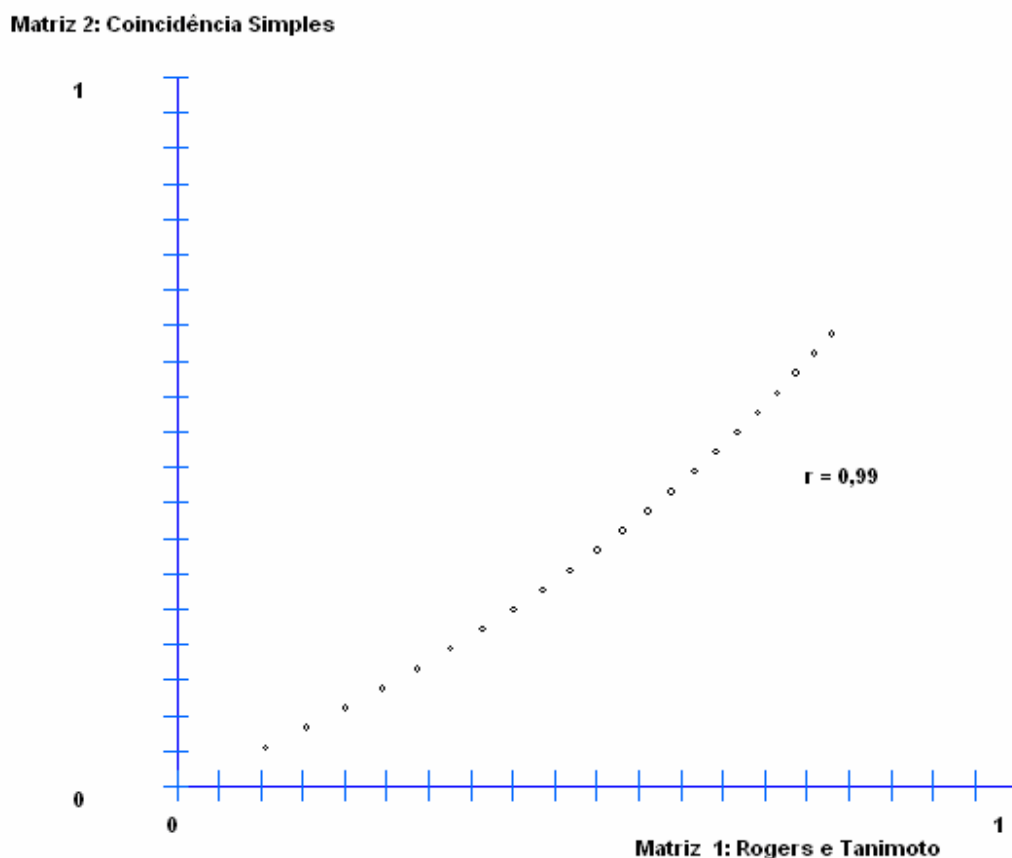


Figura 22. Correlação entre as matrizes de dissimilaridade de Rogers e Tanimoto e de Coincidência Simples.

Os resultados dos diferentes métodos aglomerativos confirmaram a existência de uma estrutura natural de agrupamento em função da similaridade genética entre os genótipos. Ressalta-se, entretanto, a importância do emprego de mais de um método de agrupamento, em função das diferenças na hierarquização, otimização e ordenação dos mesmos. Isto permite que as classificações dos grupos se complementem em função dos critérios que cada técnica utiliza, impedindo que inferências errôneas sejam adotadas na alocação de genótipos dentro de um determinado subgrupo (Arriel, 2004).

Alguns métodos mais comumente utilizados como o Vizinho Mais Distante, Ward, WPGMA e UPGMA foram considerados adequados ao agrupamento de genótipos no

presente estudo, apresentando coincidência de resultados, inclusive com relação às fontes de resistências encontradas nas genealogias dos genótipos. Conclusões semelhantes foram obtidas por ARRIEL et al. (2006) para gergelim e MEYER et al. (2004) para genótipos de milho. Além disso, muitos estudos atualmente adotam apenas o método UPGMA para análises de agrupamentos, como SMITH et al. (2000) com SSR em sorgo; PEREIRA & KERR (2001) com RAPD em abacaxizeiro; OLIVEIRA et al. (2002) com RAPD em citrus e AREIAS et al. (2006) com RAPD em arroz.

As projeções bi e tridimensionais não apresentaram índices satisfatórios, sugerindo que tal método não é o mais adequado para esses tipos de dados e as projeções não são eficientes para representar as matrizes de similaridade. De acordo com CRUZ & CARNEIRO (2006) o processo de dispersão das medidas de dissimilaridade no plano pode ser considerado satisfatório quando os coeficientes que expressam o grau de distorção e o estresse são inferiores a 20%, fato que não ocorreu no presente estudo para quaisquer dos métodos utilizados. Tais resultados também foram obtidos por MEYER et al (2004), que obtiveram valores de distorção entre 44 e 67% e estresse chegando a 71,35%, para marcadores AFLP e RAPD e para vários coeficientes de similaridade ou dissimilaridade. EMYGDIO et al. (2003) em estudos com genótipos de feijão obtiveram, para o coeficiente de Coincidência Simples, distorção de 29,7%, correlação de 0,924 e estresse de 32,9% na projeção das distâncias genéticas em um espaço bidimensional. Para o coeficiente de Rogers e Tanimoto, encontraram maiores problemas, assim como no presente estudo (51,8%), com distorção de 33%, correlação de 0,008 e estresse de 61,4%, sendo o maior valor de estresse observado pelos coeficientes de Russel e Rao e Rogers e Tanimoto, com níveis considerados insatisfatórios.

Apesar de algumas diferenças na estrutura dos agrupamentos, e alterações nas relações entre os genótipos nos diferentes coeficientes, alguns genótipos mantiveram-se agrupados, com baixa dissimilaridade genética e outros se apresentaram mais divergentes, em todos os agrupamentos.

5. CONCLUSÕES

De acordo com os dados obtidos, pode-se concluir que:

1. Os genótipos de soja resistentes ao nematóide de cisto apresentaram baixa divergência genética entre si;
2. Os marcadores microssatélites permitem identificar grupos similares e genótipos divergentes, apresentando-se eficientes no estudo de divergência genética entre genótipos de soja;
3. As projeções no plano não apresentaram resultados satisfatórios;
4. A combinação das informações provenientes dos diferentes métodos de agrupamentos mostrou-se eficiente na separação dos genótipos e cultivares de soja em grupos distintos.

6. 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, **Revista Brasileira de Genética**, v.18, p. 265-273, 1995.

ALMEIDA, A.M.R.; FERREIRA, L.P.; YORINORI, J.T.; SILVA, J.F.V.; HENNING, A.A. Doenças da Soja (*Glycine max* L.). In: KIMATI, H.; AMORIM, L.; FILHO, A.B.; CAMARGO, L.E.A.; REZENDE, J.A.M. (Ed.). **Manual de fitopatologia: doenças das plantas cultivadas**. v 2, 1997. p. 642–664.

ANAND, S.C.; GALLO, K.M. Identification of additional soybean germplasm with resistance to race 3 of soybean cyst nematode, **Plant Disease**, St. Paul, v.68, n.7, p.593-595, 1984.

ANAND, S.C. Sources of resistance to the soybean cyst nematode. In: LAMBERTI, F.; TAYLOR, C.E. (Ed) **Cyst nematodes**. New York: Plenum Press, 1985. p.269-276.

ANAND, S.C. Registration of ‘Hartwig’ soybean, **Crop Science**, Madison, v.32, n.4, p.1069-1070, 1992.

AREIAS, R.G.de B.M.; PAIVA, D.M. de; SOUZA, S.R.; FERNANDES, M.S. Similaridade genética de variedades crioulas de arroz, em função da morfologia, marcadores RAPD e acúmulo de proteína nos grãos, **Bragantia**, v.65, n.1, p.19-28, 2006.

ARRIEL, N.H.C. **Diversidade genética em gergelim (*Sesamum indicum* L.) a partir de marcadores moleculares (RAPD) e caracteres morfológicos e agrônômicos**. 2004. 114f. Tese (Doutorado em Agronomia-Genética e Melhoramento de Plantas) -

Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, 2004.

ARRIEL, N.H.C.; MAURO, A.O.Di; MAURO, S.M.Z.Di; BAKKE, O.A.; UNÊDA-TREVISOLI, S.H.; COSTA, M.M.; CAPELOTO, A.; CORRADO, A.R. Técnicas multivariadas na determinação da diversidade genética em gergelim usando marca2dores RAPD, **Pesquisa Agropecuária Brasileira**, v.41, n.5, p. 801-809, 2006.

BECELAERE, G.V.; LUBBERS, E.L.; PATERSON, A.H.; CHEE, P.W. Pedigree-vs. DNA Marker-Based genetic similarity estimates in cotton, **Crop Science**, v. 45, p. 2281-2287, 2005.

BLACK, R.J. Complexo soja: fundamentos, situação atual e perspectivas. In: CÂMARA, G.M.S. (Ed.). **Soja: tecnologia da produção II**. Piracicaba: Ed. ESALQ, 2000. p. 1-17.

BONETTI, L.P. Cultivares e seu melhoramento genético. In: VERNETTI, F. de J. (Coord.). **Soja, genética e melhoramento**. Campinas: Fundação Cargill, 1983. p.741-800.

BORÉM, A. **Melhoramento de plantas**. Viçosa: Ed. UFV, 1997. 547p.

BOTSTEIN, D.; WHITE, R.L.; SKOLNICK, M.; DAVIS, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, v. 32, p.314-331, 1980.

BRIM, C.A.; ROSS, J.P. Registration of Pickett soybeans, **Crop Science**, Madison, v.6, n.3, p.305, 1966.

BUSSAB, W. de O.; MIAZAKI, E.S.; ANDRADE, D.F. **Introdução à análise de agrupamentos**, São Paulo: Associação Brasileira de Estatística, 1990. 105p.

CAIXETA, E.T.; OLIVEIRA, A.C.B.de; BRITO, G.G.de; SAKIYAMA, N.S. Tipos de marcadores moleculares. In: BORÉM, A.; CAIXETA, E.T. (Ed.). **Marcadores Moleculares**. Viçosa: 2006. 374p.

CALDWELL, B.E. Inheritance of resistance of soybeans to cyst nematode, *Heterodera glycines*, **Agronomy Journal**, v.52, p.635-636, 1960.

CARLINI-GARCIA, L.A. **Estudo da estrutura genética populacional através de marcadores moleculares**. 1998. 118f. Monografia, (Graduação em Engenharia Agrônômica). Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1998.

CARUSO, R. **Uma caminhada sem fim**. Ed. Fundação Cargill, 1997, 95p.

CATELLI, L.L. **Caracterização de cultivares de soja utilizando marcadores moleculares microssatélites**. 2005. 175f. Dissertação (Mestrado em Genética e Biologia Molecular) - Universidade Estadual de Londrina, Londrina, 2005.

COLLINS, F.S.; BROOKS, L.D.; CHAKRAVARTI, A. A DNA polymorphism discovery resource for research on human genetic variation. **Genome**, v.8, p.1229-1231, 1998.

COMPANHIA NACIONAL DE ABASTECIMENTO. Disponível em: <<http://www.conab.gov.br>>. Acesso em: 10 jun 2007.

CONCIBIDO, V.C.; DENNY, R.L.; LANGE, D.A.; ORF, J.H.; YOUNG, N.D. RFLP mapping and marker-assisted selection of soybean cyst nematode resistance in PI 209332, **Crop Science**, v.36, p. 1643-1650, 1996.

CONCIBIDO, V.C.; LANGE, D.A.; DENNY, R.L.; ORF, J.H.; YOUNG, N.D. Genome mapping of soybean cyst nematode resistance genes in 'Peking', PI 90763, and PI 88788 using DNA markers, **Crop Science**, v.37, p. 258-264, 1997.

CORRÊA, R.X.; ABDELNOOR, R.V.; FALEIRO, F.G.; CRUZ, C.D.; MOREIRA, M.A.; BARROS, E.G.DE. Genetic distances in soybean based on RAPD markers, **Bragantia**, Campinas, v.58, n.1, p. 15-22, 1999.

COSTA, J.A.; MANICA, I. **Cultura da soja**. Porto Alegre: Ed. Evangraf, 1996. 233p.
CRUZ, C.D. **Programa Genes**: versão Windows: aplicativo computacional em genética e estatística. Viçosa: Editora UFV, 2001. 648p.

CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: Ed. UFV, 2006. v.2, 585p.

CRUZ, C.D.; REGAZZI, A.J.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: Ed. UFV, 2004. v.1, 480p.

CRUZ, C.D.; VIANA, J.M.S. A method of genetic divergence analysis based on sample uni projection on two-dimensional space, **Revista Brasileira de Genética**, Ribeirão Preto, v.17, n.1, p. 69-73, 1994.

DIAS, L.A., dos S. Análises multidimensionais. In: ALFENAS, A.C. **Eletroforese de isoenzimas e proteínas afins**: fundamentos e aplicações em plantas e microorganismos. Viçosa: UFV, 1998. 574p.

DIAS, W.P. **Genética da resistência da soja à raça 4+ do nematóide de cisto, *Heterodera glycines***. 2003. 83f. Dissertação (Doutorado) - Universidade Federal de Lavras, Lavras, 2003.

DIAS, W.P.; SILVA, J.F.V.; WAIN, A.L.; PEREIRA, J.E. Distribuição de raças de *Heterodera glycines* no Brasil. In: SOCIEDADE BRASILEIRA DE NEMATOLOGIA. **O nematóide de cisto da soja: a experiência brasileira**. Jaboticabal: Artsigner Editores, 1999. p.95-103.

DUARTE, J.M.; SANTOS, J.B.dos; MELO, L.C. Comparison of similarity coefficients based on RAPD markers in the common bean, **Genetics and Molecular Biology**, Ribeirão Preto, v.22, n.3, p.427-432, 1999.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Recomendações técnicas para a cultura da soja na região central do Brasil 1999/2000**. Londrina, 1999. 226 p.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Tecnologias de Produção de Soja - Região Central do Brasil - 2003**. Londrina, 2003. p.11-24.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Tecnologias de Produção de Soja – Região Central do Brasil – 2005**. Londrina, 2004. 239p.

EMYGDIO B.M.; ANTUNES, I.F.; CHOER, E.; NEDEL, J.L. Eficiência de coeficientes de similaridade em genótipos de feijão mediante marcadores RAPD, **Pesquisa Agropecuária Brasileira**, v.38, n.2, p.243-250, 2003.

FALCONER, D.S. **Introdução à genética quantitativa**. Viçosa: UFV, Impr. Univ., 1981. 279p.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3ª ed. Brasília, DF: EMBRAPA-CENARGEN, 1998. 220p.

GOWER, J.C.; LEGENDRE, P. Metric and Euclidean properties of dissimilarity coefficients, **Journal of Classification**, v.3, p.5-48, 1986.

HARTWIG, E.E. Breeding productive soybeans with resistance to the soybean cyst nematode. In: **World Soybean Conference**, 3., 1985, Boulder, Proceedings... Boulder: editora, 1985. p. 394-399.

HARTWIG, E.E; YOUNG, L.D. Registration of soybean germplasm line J81-116. **Crop Science**, Madison, v.26, n.1, p.209, 1986.

HIROMOTO, D.M.; VELLO, N.A. The genetic base of Brazilian soybean (*Glycine max* (L.)Merrill) cultivars, **Revista Brasileira de Genética**, v.9, p. 295-306, 1986.

HYMOWITZ, T. On the domestication of the soybean, **Economic Botany**, v.24, p.408-421, 1970.

JACKSON, A.A.; SOMERS, K.M.; HARVEY, H.H. Similarity coefficients: measures for co-occurrence and association or simply measures of occurrence?, **American Naturalist**, v.133, p.436-453, 1989.

JEFFREY, A.J.; WILSON, V.; THEIN, S.L. Hypervariable minisatellite regions in human DNA, **Nature**, v. 314, p.67-73, 1985.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. New Jersey: Prentice-Hall, 1988. 607p.

KEIM, P.; DIERS, B.W.; OLSON, T.C.; SHOEMAKER, R.C. RFLP mapping in soybean: association between marker loci and variation in quantitative traits, **Genetics**, v.126, p. 735-742, 1990.

KIIHL, R.A.S.; ALMEIDA, L.A.; SOUZA, P.I.M.; HIROMOTO, D.M.; ARANTES, N.E.; FARIAS, L.C.; MIRANDA, M.A.C.; MORCELI, A.; TAKEDA, C.; BONATO, E.R.; BERTAGNOLLI, P. F.; PRADO, E.E. **Desenvolvimento de germoplasma de soja adaptados às várias regiões ecológicas e aos vários sistemas de produção: resultados de pesquisa da Embrapa Soja – 1999.** Londrina: Embrapa Soja, 2000. p.144-145. (Embrapa Soja. Documentos, 142).

KRESOVICH, S.; SZEWC-MCFADDEN, A.K; BLICK, S. Abundance and characterization of simple sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed), **Theoretic Applied Genetics**, v.91, p.206-211, 1995.

LITT, M.; LUTY, J.A. A hyper variable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. **Am J Hum Genet.**, v.44, p.398-401, 1989.

MAHALIGAN, R.; SKORUPSKA, H.T. DNA markers for to *Heterodera glycines* I. Race 3 in soybean cultivar Peking. **Breed Science**, v.45, p. 435-443, 1995.

MATSON, A.L.; WILLIAMS L.F. Evidence of a fourth gene for resistance to the soybean cyst nematode. **Crop Science**, Madison, v.5, n.5, p.477, 1965.

MAURO, A.O.; OLIVEIRA, A.L.de; MAURO, S.M.Z. Genetics of resistance to soybean cyst nematode, *Heterodera glycines* Ichinohe (race 3), in a Brazilian soybean population. **Genetics and Molecular Biology**, v.22, n.2, p.257-260, 1999.

MESSMER, M.M.; MELCHINGER, A.E.; HERRMANN, R.G.; JURGEN BOPPENMAIER. Relationships among early European maize inbreds: II. Comparison of pedigree and RFLP data, **Crop Science**, v.33, p. 944-950, 1993.

MEYER, A.S. **Comparação de coeficientes de similaridade usados em análises de agrupamentos com dados de marcadores moleculares dominantes**. 2002. 106f. Dissertação. Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba. 2002.

MEYER, A.S.; GARCIA, A.A.F.; SOUZA, A.P.DE, SOUZA JR., C.L. Comparison of similarity coefficients used for cluster analysis with dominant markers in maize (*Zea mays* L). **Genetics and Molecular Biology**, v.27, n.1, p. 83-91, 2004.

MUDGE, J.; CONCIBIDO, V.C.; DENNY, R.L.; YOUNG, N.D.; ORF, J.H. Genetic mapping of a yield depression locus near a major gene for soybean cyst nematode resistance, **Soybean Genetic Newsletters**, v.23, p. 175-178, 1996.

MUHLEN, G.S.; MARTINS, P.S.; ANDO, A. Variabilidade genética de etnovariedades de mandioca, avaliada por marcadores de DNA, **Scientia Agrícola**, v.57, n.2, p. 319-328, 2000.

MULLIS, K.; FALOONA, F. Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. **Methods in Enzymology**, v. 55, p. 335-350, 1987.

MYERS, G.O.; ANAND, S.C. Inheritance of resistance and genetic relationships among soybean plant introductions to races of soybean cyst nematode, **Euphytica**, v.55, p. 197-201, 1991.

OLIVEIRA, R.P.de; CRISTOFANI, M.; AGUILAR-VILDOSO, C.I.; MACHADO, M.A. Diversidade genética entre híbridos de tangerina 'Cravo' e laranja 'Pera'. **Pesquisa Agropecuária Brasileira**, Brasília, v.37, n.4, p. 479-484, 2002.

PARAN, I; MICHELMORE, R.W. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce, **Theoretic Applied Genetics**, v. 85, p.985-993, 1993.

PEREIRA, A.V. **Utilização de análise multivariada na caracterização de germoplasma de mandioca (*Manihot esculenta*, Crantz).** 1989. 179f. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1989.

PEREIRA, C.D.; KERR, W.E. Divergência entre doze genótipos de abacaxizeiro (*Ananás comosus* L, Merrill.) estimada por análise de marcadores RAPD. **Revista Brasileira Fruticultura**, v.23, n.2, p.335-338, 2001.

PRABHU, R.R.; NJITI, V.N.; BELL-JOHNSON, B.; JOHNSON, J.E.; SCHIMIDT, M.E.; KLEIN, J.H.; LIGHTFOOT, D.A. Selecting soybeans cultivars for dual resistance to soybean cyst nematode and sudden death syndrome using two DNA markers. **Crop Science**, Madison, v.39, n.3/4, p. 356-364, 1999.

PRIOLLI, R.H.G.; MENDES-JUNIOR, C.T.; ARANTES, N.E.; CONTEL, E.P.B. Characterization of brasilian soybean cultivars using microsatellite markers. **Genetics and Molecular Biology**, v.25, n.2, p. 185-193, 2002.

RAO-ARELLI, A.P.; ANAND, S.C. Genetic relationships amng soybean plant introductions for resistance to race 3 of the soybean cyst nematode. **Crop Science**, v.28, p.650-652, 1988.

RAO-ARELLI, A.P.; ANAND, S.C.; WEATHER, A. Soybean resistance to soybean cyst nematode race 3 is conditioned by an additional dominant gene. **Crop Science**, v.28, p.650-652, 1992.

RAO-ARELLI, A.P. Inheritance of resistance to *Heterodera glycines* race 3 in soybean accessions, **Plant Disease**, St. Paul, v.78, n.9, p. 898-900, 1994.

RAO, C.R. **Advanced statistical methods in biometric research**. New York: John Wiley and Sons, 1952. 390p.

ROCHA, B.H.G. **Caracterização de genótipos e análise de pureza varietal em sementes de batata através de marcadores moleculares**. 2000. 54 f. Tese (Doutorado em Ciências e Tecnologia de Sementes)-Universidade Federal de Pelotas, Pelotas, 2000.

ROMESBURG, H.C. **Cluster analysis for researchers**. California: Lifetime Learning Publications, 1984. 334p.

SCHIMIDT, D.P.; NOEL, G.R. Nematode parasites of soybeans. In: NICKLE, W.R. (Ed.) **Plant and insect Nematodes**. New York: Marcel Dekker, 1984. p.14-59.

SCHMITT, R.D.; BARKER, K.R. **Plant-parasitic nematodes on soybean in North Carolina**. The North Carolina Agricultural Extension Service, 1985. 8p.

SLEPER, D.A.; SHANNON, J.G. Role of public and private soybean breeding programs in the development of soybean varieties biotechnology, **The Journal of Agrobiotechnology Management e Economics**, v.6, 2003.

SMITH, J.C.S.; KRESOVICH, S.; HOPKINS, M.S.; MITCHELL, S.E.; DEAN, R.E.; WOODMAN, W.L.; LEE, M.; PORTER K. Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats, **Crop Science**, v.40, p.226-232, 2000.

SNEATH, P. H. A.; SOKAL, R. R. **Numeric taxonomy: the principles and practice of numerical classification**. San Francisco: W.H. Freeman, 1973. 573p.

TAYLOR, A.L. **Introductions to research on plant nematology**. Rome: FAO, 1971. 133p.

TODD, J.J.; VODKIN, L.O. Duplications that suppress and deletions that restore expression from a chalcone synthase multigene family, **Plant Cell**, v.8, p. 687-699, 1998.

VELLO, N. A.; HIROMOTO, D. M.; AZEVEDO-FILHO, A. J. B.V. Coefficient of parentage and breeding of Brazilian soybean germplasm, **Brazilian Journal Genetics**, v.11, p.679-697, 1988.

VOS, P. HOGERS, R.; BLEEKER, M.; REIJANS, M.; LEE, T. VAN DE; HORNER, M.; FRITJERS, A.; POT, J.; PELEMAN, J.; KUIPER, M.; ZABEAU, M. AFLP: a new technique for DNA fingerprinting. **Nucleic acids Research**, v. 23, p.4407-4414, 1995.

WEBB, D.M.; BALTAZAR, B.M.; RAO-ARELLI, A.P.; SCHUPP, J.; CLAYTON, K.; KEIM, P.; BEAVIS, W.D. Genetic mapping of soybean cyst nematode race-3 resistance loci in the soybean PI 437.654. **Theoretic Applied Genetics**, v.91, p. 574-581, 1995.

WEIR, B.S. **Genetic data analysis II**. 2 ed. Sunderland: Sinauer Associates, 1996. 445p.

WEISEMANN, J.M.; MATTHEWS, B.F.; DEVINE, T.E. Molecular marker located proximal to the soybean cyst nematode resistance gene, *Rgh4*. **Theoretic Applied Genetics**, v.85, p. 136-138, 1992.

WELLE, R.; SCHRODER, G.; SCHILTZ, E.; GRISEBACH, H.; SCHRODER, J. Induced plant responses to pathogen attack. **Eur. J. Biochem.**, v.196, p. 423-430, 1991.

WILLIAMS, K. A.; KUBELIK, A.R.; LIVAK, K.J.; RAFALSKI, J.A.; TINGEY, S.V. DNA polymorphism amplified by arbitrary primers re useful as genetic markers, **Nucleic. Acid. Res.**, v.18, p.6531-6535, 1990.

YORINORI, J.T. Controle integrado das principais doenças da soja. In: CÂMARA, G. M. S. **Soja: tecnologia da produção**. ESALQ/ USP, 1998. 293p.

YORINORI, J.T. Fungal diseases in EMBRAPA. Empresa Brasileira de Pesquisa Agropecuária. Centro Nacional de Pesquisa da Soja. CNPSo. FAO. **Plant Production and Protection**, Londrina, n 27, 1994. 254 p.

YORINORI, J.T. Riscos de surgimento de novas doenças na cultura da soja. In: CONGRESSO DE TECNOLOGIA E COMPETITIVIDADE DA SOJA NO MERCADO GLOBAL, 2000, Cuiabá. **Anais...**Cuiabá: Fundação MT, 2000. p.165-169.

YORINORI, J.T. Situação atual das doenças potenciais no cone sul. In: CONGRESSO BRASILEIRO DE SOJA E MERCOSOJA, 2002, Foz do Iguaçu, PR. **Resumos...** Londrina: Embrapa Soja, 2002. p.171-186.

YUE, P.; SLEPER, D.A.; RAO-ARELLI, A.P. Genetic analysis of resistance to soybean cyst nematode in PI 438489B. **Euphytica**, Dordrecht, v.116, n.2, p.181-186, 2000.