

UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE AGRONOMIA E ENGENHARIA DE ALIMENTOS

**CARACTERIZAÇÃO DE GERMOPLASMA DE SOJA
QUANTO AOS TEORES DE PROTEÍNA,
ÓLEO E ÁCIDOS GRAXOS**

KEYLA DE OLIVEIRA RIBEIRO

Orientador:
Prof. João Batista Duarte

Abril - 2011

KEYLA DE OLIVEIRA RIBEIRO

**CARACTERIZAÇÃO DE GERMOPLASMA DE SOJA
QUANTO AOS TEORES DE PROTEÍNA,
ÓLEO E ÁCIDOS GRAXOS**

Tese apresentada ao Programa de Pós-graduação em Agronomia, da Universidade Federal de Goiás, como requisito parcial à obtenção do título de Doutor em Agronomia, área de concentração: Genética e Melhoramento de Plantas.

Orientador:

Prof. Dr. João Batista Duarte

Co-orientador:

Dr. Jaison Pereira de Oliveira

Goiânia, GO – Brasil
2011

Dados Internacionais de Catalogação na Publicação (CIP)
GPT/BC/UFG

R484c **Ribeiro, Keyla de Oliveira.**
Caracterização de germoplasma de soja quanto aos teores de proteína, óleo e ácidos graxos [manuscrito] / Keyla de Oliveira Ribeiro. - 2011.
151 f. : il.

Orientador: Prof. Dr. João Batista Duarte.
Tese (Doutorado) – Universidade Federal de Goiás,
Escola de Agronomia e Engenharia de Alimentos, 2011.

Bibliografia: f. 121-123

1. Genética vegetal – Melhoramento. 2. Germoplasma – Divergência Genética. 3. Soja – Espécies cultivadas. I. Título.

CDU:631.528.1:633.34

KEYLA DE OLIVEIRA RIBEIRO

CARACTERIZAÇÃO DE GERMOPLASMA DE SOJA QUANTO AOS
TEORES DE PROTEÍNA, ÓLEO E ÁCIDOS GRAXOS

Tese DEFENDIDA e APROVADA em 26 de abril, pela Banca Examinadora
constituída pelos membros:

Prof. Dr. Natal Antônio Vello
Membro – ESALQ/USP

Dr. Cláudio Takao Karia
Membro – Embrapa Cerrados

Prof. Dr. Alexandre Siqueira Guedes Coelho
Membro – EA/UFG

Dr. Leandro Oliveira e Silva
Membro – Emater/Seagro-GO

Prof. Dr. João Batista Duarte
Orientador – EA/UFG

Goiânia, Goiás
Brasil

DEDICATÓRIA

Aos meus pais, Oribaldo e Celina, que sempre me incentivaram e apoiaram nesta jornada.

Obrigada por todo esforço dedicado à minha felicidade.

Eternamente grata.

Ao meu esposo, Luciano, pelo grande companheirismo, incentivo, compreensão, e muito mais...

Dedico minhas conquistas a você.

AGRADECIMENTOS

À Universidade Federal de Goiás, através da Escola de Agronomia e Engenharia de Alimentos (EA-UFG), à Capes, ao CNPq e à Funape, pela oportunidade de desenvolvimento e divulgação do trabalho e por minha qualificação.

À Secretaria de Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em especial ao Dr. Pedro Manuel Figueira de Oliveira Monteiro, pela oportunidade da parceria e disponibilização do material básico para o desenvolvimento deste trabalho. Graças à sua dedicação à pesquisa agropecuária em Goiás, tornaram-se viáveis os meus projetos de mestrado e de doutorado.

Ao professor Dr. João Batista Duarte (EA-UFG), pelos valiosos ensinamentos, investimento na minha qualificação e oportunidade de crescimento. Agradeço, ainda, pela fundamental ajuda na melhoria deste trabalho. Suas atitudes me motivam a permanecer na “vida” acadêmica.

Aos colegas da Estação Experimental de Senador Canedo / Seagro-GO, Regina, Luciano, Carol, Nerivaldo, Mansuêmia, Adriano, pelo companheirismo e amizade. Meus agradecimentos especiais ao Macarrão e Taco, pelo suporte essencial para realização dos ensaios em campo.

Ao Centro de Pesquisa em Alimentos (CPA), em nome do professor Dr. Albenones José de Mesquita, pela oportunidade da realização das análises de ácidos graxos. Em especial, ao Rodrigo Almeida de Oliveira, químico responsável pelo CPA, pela dedicação e orientação na realização das análises, e pela amizade que levarei eternamente.

Ao Dr. Jaison Pereira de Oliveira (Embrapa Arroz e Feijão), pelo incentivo constante e pelas orientações na execução deste trabalho.

Aos professores do Setor de Melhoramento de Plantas / EA-UFG, pelos ensinamentos durante esses anos. Ao professor Dr. Lázaro José Chaves, pela grande amizade e atenção desde o dia que ingressei no “mundo” do melhoramento genético (2002). Ao professor Dr. Alexandre Siqueira, pelo apoio na realização das análises laboratoriais.

Aos estagiários que muito me ajudaram, especialmente à Elitânia, Luíz Antônio e Ana Letycia, pela persistência e dedicação nas atividades e análises laboratoriais. Obrigada pelos diversos finais de semana dedicados à realização dessas análises.

Aos colegas de laboratório e de pós-graduação, Angel, Nara, Carol, Jarênio, Arthur, Ludmila, Cláudio (Bonfin), Cláudio Takao, Aracelle, Luíce, pelos momentos agradáveis e de crescimento.

Ao grande amigo Welinton, pelo companheirismo e “ombro amigo” de todas as horas. Ao Angel, pela amizade e carinho.

À minha amiga Camila, pelo apoio nos momentos difíceis e pela presença nos momentos de alegria.

À minha sogra Maria Helena, por tantas vezes me motivar e valorizar o meu trabalho.

Às minhas irmãs, Kelly, Késia e Keysi, pela motivação e apoio e, ainda, por cuidarem de tudo enquanto necessitei me ausentar.

Agradeço a Deus pela oportunidade de viver esta experiência, e pela capacidade de enfrentar os desafios lançados. Foi a certeza no amor de Deus que me deu força e manteve a minha fé.

SUMÁRIO

1	INTRODUÇÃO	11
2	REVISÃO DE LITERATURA	14
2.1	IMPORTÂNCIA ECONÔMICA DA SOJA	14
2.2	TEOR DE PROTEÍNA NO GRÃO EM SOJA	16
2.3	TEOR E COMPOSIÇÃO DO ÓLEO EM SOJA	23
2.4	BIODIESEL E SOJA.....	30
2.5	IMPORTÂNCIA E CARACTERIZAÇÃO DE BANCOS DE GERMOPLASMA.....	34
2.6	ANÁLISES DE COLEÇÕES DE GERMOPLASMA	38
2.6.1	Divergência genética	38
2.6.2	Técnicas de agrupamento	41
2.6.3	Análises de componentes principais e de variáveis canônicas	47
3	MATERIAL E MÉTODOS	51
3.1	DELINEAMENTO EXPERIMENTAL	51
3.2	ANÁLISES FÍSICO-QUÍMICAS	52
3.2.1	Análise do teor de proteína	52
3.2.2	Análise do teor de óleo	53
3.2.3	Análise dos teores de ácidos graxos	54
3.3	ANÁLISES ESTATÍSTICAS	55
3.3.1	Estimação das médias genótípicas dos acessos	55
3.3.2	Análises de componentes principais e de variáveis canônicas	56
3.3.3	Análise de agrupamento	57
3.3.4	Análise discriminante canônica	59
4	RESULTADOS E DISCUSSÃO	60
4.1	VARIABILIDADE DO GERMOPLASMA: ENFOQUE UNIVARIADO	60
4.2	HERDABILIDADE E CORRELAÇÃO DE CARACTERES.....	67
4.3	DIVERGÊNCIA GENÉTICA ENTRE ACESSOS: ENFOQUE MULTIVARIADO	71
4.3.1	Estrutura da diversidade entre acessos	71
4.3.2	Caracterização dos grupos de similaridade	77
4.3.3	Divergência entre acessos com óleo de melhor perfil em estabilidade oxidativa	86
4.3.4	Importância relativa dos caracteres na discriminação dos acessos	91
5	CONCLUSÕES	99
6	REFERÊNCIAS	101
	APÊNDICES	124

RESUMO

RIBEIRO, K. O. **Caracterização de germoplasma de soja quanto aos teores de proteína, óleo e ácidos graxos.** 2011. 151 f. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas) – Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2011.¹

A soja é atualmente a cultura de maior destaque na safra brasileira de grãos. Entre seus principais produtos estão o óleo e o farelo, os quais têm sido matéria-prima para a produção de alimentos e ração animal. O elevado teor de proteína representa o principal atrativo do uso da soja na alimentação. Além disso, o óleo de soja tem sido também utilizado na produção de biodiesel. Esses segmentos industriais buscam cultivares com perfis semelhantes de ácidos graxos, cujo óleo garanta melhor estabilidade oxidativa. Isto corresponde a óleos com baixos teores de ácidos poli-insaturados, associados a elevados teores de ácidos graxos monossaturados. Para isso, torna-se indispensável a caracterização de atributos relacionados à qualidade do grão dos cultivares comerciais; o que pode ser estendido, para os propósitos de melhoramento da soja nesses atributos, aos recursos genéticos preservados em bancos de germoplasma. Atualmente a caracterização desses recursos ainda é deficiente, sobretudo quando se tratam de caracteres especiais como aqueles anteriormente mencionados. Nesse contexto, o objetivo deste trabalho foi caracterizar uma subcoleção de 527 acessos de soja do banco de germoplasma da Secretaria de Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), com base nos teores de proteína, óleo e dos principais ácidos graxos no óleo. Os grãos foram obtidos de ensaio de campo conduzido na Unidade Experimental da Seagro-GO (16°43'S; 49°7'W; 765 m), município de Senador Canedo-GO, na safra 2007/2008. O delineamento experimental foi em blocos aumentados (blocos de Federer), com duas cultivares testemunhas (Conquista e Monsoy 6101) e parcela constituída de uma fileira de plantas (5,0 m x 0,5 m). Os teores de proteína foram quantificados pelo método Kjeldahl; de óleo, por Soxhlet; e os de ácidos graxos (palmítico, esteárico, oleico, linoleico e linolênico), pelo método FAME (*Fatty Acid Methyl Esters*). Para a avaliação da divergência genética entre os acessos, utilizaram-se as análises multivariadas de variáveis canônicas, de componentes principais e de agrupamento. Esta última foi realizada pelo método hierárquico aglomerativo, associado ao critério de ligação UPGMA. Também se realizaram análises discriminantes canônicas entre os grupos de similaridade identificados e entre os ciclos de maturação dos acessos (precoce, médio e tardio). Na análise de componentes principais foram necessárias três componentes para explicar 78% da variância total, ao passo que na análise de variáveis canônicas, apenas duas destas variáveis explicaram 81,2% dessa variação. Em relação à variação fenotípica, observou-se que a subcoleção apresenta variabilidade relativamente baixa entre os acessos, em seus teores de proteína (de 35,7% a 44,0%) e óleo (de 14,8% a 18,5%). Para os ácidos graxos, o germoplasma apresenta variação relativamente ampla, embora características de interesse industrial como elevados teores de ácido oleico (maiores que 40%) e baixos teores de ácidos poli-insaturados (menores que 40%) sejam restritas a poucos acessos. Quanto aos ácidos graxos poli-insaturados, o germoplasma apresenta variabilidade satisfatória para o ácido linoleico (ômega-6), porém, baixa variabilidade para o ácido linolênico (ômega-3), sobretudo em seu limite superior. Pela análise de agrupamento foram estabelecidos cinco

¹ Orientador: Prof. Dr. João Batista Duarte. EA-UFG.

Co-orientador: Dr. Jaison Pereira de Oliveira. Embrapa Arroz e Feijão.

grupos de similaridade multivariada (A, B, C, D e E), com formação de subgrupos em alguns deles. Os acessos dos grupos A e B apresentam os maiores teores de óleo e de proteína, respectivamente. Os acessos do grupo C destacam-se com os melhores perfis em qualidade do óleo (estabilidade oxidativa). Mesmo selecionando apenas o germoplasma com melhor perfil em estabilidade oxidativa do óleo, ainda foi possível identificar nessa coleção cinco grupos de acessos significativamente divergentes nos teores de ácidos graxos. As variáveis que se mostraram mais relevantes na estruturação da divergência genética desse germoplasma foram os teores de ácidos graxos poli-insaturados e saturados.

Palavras-chave: divergência genética, variável canônica, estabilidade oxidativa.

ABSTRACT

RIBEIRO, K. O. **Characterization of soybean germplasm for protein, oil and fatty acids contents.** 2011. 151 f. Thesis (Doctorate in Agronomy: Genetics and Plant Breeding) – Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2011.²

Soybean is currently the most prominent crop in the Brazilian production of grains. Its main products include oil and meal, which have been used for the production of food and animal feed. The high protein contents represent the main attraction of soybean use in diet. Furthermore, soybean oil has also been used in the production of biodiesel. These industrial segments demands cultivars with similar profiles of fatty acids, in which case the oil guarantees a better oxidative stability. This profile corresponds to oils with low polyunsaturated acids contents, associated to high contents of monounsaturated fatty acids. For that, it becomes essential the chemical characterization of grains in cultivars, which could be extended to the genetic resources preserved in germplasm collections, with the purpose of breeding soybean in these traits. Nowadays the characterization of these resources is still deficient, mainly regarding special traits such as those previously mentioned. In this context, the objective of this study was to characterize a sub-collection of 527 accessions of soybean from the germplasm bank of Secretaria de Agricultura, Pecuária e Abastecimento do Estado Goiás (SEAGRO-GO), based on the physicochemical traits of the grain, including protein and oil contents in the meal, and the main fatty acids contents of the oil. The grains were obtained through a field trial carried out at the experimental unit of SEAGRO-GO (16°43'S; 49°7'W; 765 m), county of Senador Canedo - GO, in the 2007-2008 growing season. The experimental design was augmented blocks (blocks of Federer), with two check cultivars (Conquista and Monsoy 6101). The plots were constituted by one row of plants (5.0 m x 0.5 m). The protein contents were quantified by the Kjeldahl method, the oil by the Soxhlet method, and the fatty acids (palmitic, stearic, oleic, linoleic and linolenic) by FAME method (Fatty Acid Methyl Esters). To assess the genetic divergence among accessions the multivariate analyses used were: canonical variables, principal components and cluster analysis. The last one was performed by the hierarchical clustering method using UPGMA criteria. Canonical discriminant analysis was also performed on established similarity groups and on maturity groups of the accessions (early, intermediate and late). In the principal components analysis it was needed three components to retain 78% of the total variance, whereas in the canonical variables analysis only two of these variables captured 81.2% of this variation. About the phenotypic variation, it was observed that the studied sub-collection presents a relatively low variability among its accessions, in protein content (from 35.7% to 44.0%) and oil content (from 14.8% to 18.5%). For fatty acids, the germplasm showed a relatively wide variation, although characteristics of industrial interest such as high contents of oleic acid (greater than 40%) and low contents of polyunsaturated acids (less than 40%) are restricted to few accessions. Regarding the polyunsaturated acids, the germplasm presents satisfactory variability to the linoleic oil (omega-6), however, low variability to the linolenic acid (omega-3), especially in its upper limit. By cluster analysis they were established five similarity groups (A, B, C, D, and E) with the formation of subgroups in some them. The accessions of groups A and B present the largest oil and protein contents,

² Adviser: Prof. Dr. João Batista Duarte. EA-UFG.

Co-adviser: Dr. Jaison Pereira de Oliveira. Embrapa Arroz e Feijão.

respectively. The accessions of group C outstood with the best profile of oil quality (oxidative stability), and they were followed by the accessions of E and B groups. Among the germplasm with the best profiles of oil oxidative stability, it was still possible to identify in this collection five divergent groups of accessions with significant differences in fatty acids contents. The polyunsaturated and saturated fatty acids are the most relevant variables in the structuring of genetic diversity of this germplasm.

Key words: divergence, canonical variable, oxidative stability.

1 INTRODUÇÃO

A soja é atualmente a cultura de maior destaque na safra brasileira de grãos, com produção de 64,6 milhões de toneladas em 2009/2010. O crescimento da produção foi 20,2% superior em relação à safra anterior. A região Centro-Oeste do país contribuiu com quase 50% dessa produção, sendo o Estado do Mato Grosso o maior produtor. A produtividade média do Brasil passou de 2.629 kg/ha, em 2008, para 2.927 kg/ha, em 2009. Este aumento foi observado em todas as regiões brasileiras, sendo mais expressivo na região Sul. Neste mesmo ano, o Brasil exportou quase 30 milhões de toneladas de soja em grão, 14 milhões de toneladas de óleo e 13 milhões de toneladas de farelo de soja. O aumento da produtividade de variedades comerciais colabora também para que o Brasil se destaque como o segundo maior produtor mundial dessa oleaginosa e, conseqüentemente, atinja o melhor desempenho na sua exportação. Ademais, a comercialização de produtos da soja, principalmente o farelo, tem grande importância na economia brasileira (Conab, 2010a).

A soja é destinada principalmente para obtenção de óleo e farelo. Estes seus derivados são matérias-primas para vários segmentos industriais. Nos últimos anos, tem crescido a procura pelos produtos da soja para uso na alimentação, pois esses contribuem ricamente na dieta. As cultivares de soja destinadas ao consumo humano apresentam excelente balanço de aminoácidos essenciais e podem ser consideradas fontes de proteína de alta qualidade (Vieira et al., 1999). Devido às ricas propriedades nutricionais do farelo, a soja tem-se tornado um dos insumos de maior importância para a avicultura, principalmente no Brasil, ocupando lugar de crescente destaque também na criação de suínos (Wilson, 2004). Acrescem-se os relatos sobre efeitos positivos da adição do óleo de soja em rações para variados fins (Barreto et al., 2010).

Outra exploração recente da soja tem sido através da utilização de seu óleo na geração de energia, com a produção de biodiesel. Este biocombustível, além de ser uma fonte renovável de energia, permite reduções nas emissões de compostos poluentes como os óxidos de enxofre e gases que contribuem para o efeito estufa (Costa Neto et al., 2000). Segundo Ramos et al. (2003), é necessário reconhecer que a soja, embora seja uma espécie

com exploração tradicional no setor alimentício, apresenta um dos maiores potenciais para servir de modelo para o desenvolvimento de biodiesel. De acordo com Pinto et al. (2005), a soja é a cultura mais citada em artigos, como fonte para a produção de biodiesel.

Em resumo, dois enfoques principais para esse tipo de exploração se apresentam aos melhoristas de soja. O primeiro está relacionado aos caracteres ligados à qualidade nutricional, e o segundo, à qualidade tecnológica da soja. O consumidor tem buscado dietas saudáveis e, conseqüentemente, a indústria manifesta uma demanda crescente por produtos melhorados em relação ao conteúdo de proteína, perfil de ácidos graxos, concentração de vitaminas, entre outros. Nesse sentido, a busca por aumento dos teores de proteína no grão de soja pode contribuir para a geração de subprodutos com altos níveis protéicos, aumentando-se, então, a disponibilidade energética. Em outras espécies como arroz, batata doce e feijão, o melhoramento genético já contribuiu mais efetivamente para o desenvolvimento de cultivares com melhor qualidade nutricional (Duarte, 2003; Guimarães et al., 2004; Nestel et al., 2006; Pfeiffer et al., 2007; Mayer et al., 2008).

Outro enfoque potencial a ser perseguido pelo melhoramento genético da soja é o dos caracteres ligados à qualidade tecnológica e nutricional do grão. Exemplos na soja são aqueles que se relacionam à estabilidade do óleo, seja para a produção de alimentos e rações ou para a produção de biodiesel. Estes segmentos industriais buscam cultivares com perfis semelhantes de ácidos graxos, cujo óleo garanta melhor estabilidade; isto é, baixos teores de ácidos graxos saturados e de ácidos poli-insaturados, associados a teores elevados de ácidos graxos monossaturados. Enfim, o crescimento e a diversificação do mercado de alimentos à base de soja, aliado à sua exploração energética para biodiesel, possivelmente deverão ser acompanhados pelo desenvolvimento de cultivares próprias para cada tipo de uso.

Nesse sentido, estudos que venham caracterizar os acessos dos bancos de germoplasma de soja, no que tange a características químicas ligadas à qualidade do grão (ex. teores de óleo e de proteína, aliados aos perfis de ácidos graxos e de aminoácidos), permitiriam acelerar o processo de melhoramento da espécie, para melhor atender essas demandas. Assim, o uso desse material genético tanto poderia se dar de maneira direta pelo cultivo de genótipos que reúnam propriedades de interesse, ou utilizando-os como genitores em cruzamentos com cultivares comerciais, para obter progênie adaptadas e melhoradas nesses atributos. Vários trabalhos relatam a existência de variabilidade entre linhagens de soja para os teores de óleo e de proteína, e, também, com melhor qualidade

nestes componentes (Tango et al., 1983; Vieira et al., 1999; Wilson, 2004; Ribeiro, 2006; Oliveira, 2007; Oliveira et al., 2010).

Em síntese, a caracterização dos genótipos nos bancos de germoplasma é passo fundamental, pois permite ao melhorista o uso programado dos recursos genéticos neles preservados. Em várias espécies, a caracterização morfológica desses acessos tem sido realizada (Barros, 1991; Bered et al., 2002; Ohara & Shimamoto, 2002; Silva & Costa, 2003; Karia, 2008). Com isso, informações adicionais mais precisas sobre a variabilidade genética do banco podem ser obtidas, incluindo a identificação de grupos heteróticos e, conseqüentemente, o planejamento de cruzamentos que permitam segregações com maior potencial de exploração. Isso permite também ao melhorista acelerar o processo de melhoramento para determinado caráter de interesse, minimizando-se o tempo para se obter novas variedades melhoradas. Além disso, esse tipo de trabalho possibilita a identificação de duplicidades (duplicatas) entre os acessos ou mesmo a de grupos de similaridade entre eles, que subsidiam decisões acerca da regeneração e preservação dos acessos, bem como a formação de coleções nucleares.

O Banco de Ativo de Germoplasma (BAG) de soja mantido pela Secretaria de Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), objeto do presente estudo, vem sendo constituído ao longo de trinta anos; concomitantemente ao melhoramento da cultura para a região central do Brasil, desde a década de 1970, quando ainda em atividade a extinta Empresa Goiana de Pesquisa Agropecuária (EMGOPA). A potencialidade desses recursos, relativamente a demandas específicas atuais do setor agroindustrial, entretanto, ainda não foi avaliada. Além disso, há possibilidade de se identificar diretamente nesse banco, material promissor, de boa adaptação às condições de cultivo da soja na região e com características de interesse específico. Isto porque boa parte desses acessos deriva de cruzamentos que foram planejados para a obtenção de cultivares adaptadas às condições do Cerrado da região central do Brasil.

Nesse contexto, o objetivo da presente pesquisa foi a caracterização de subcoleção representativa do BAG de soja mantido pela Seagro-GO, no que tange a caracteres relacionados à qualidade tecnológica e nutricional do grão, especificamente quanto aos teores de proteína, óleo e ácidos graxos. Isto poderá subsidiar, tanto o processo de recomendação de cultivares para fins mercadológicos específicos, bem como os programas de melhoramento que visem utilizar tais recursos para a obtenção de cultivares que melhor atendam a esse tipo de demanda.

2 REVISÃO DE LITERATURA

2.1 IMPORTÂNCIA ECONÔMICA DA SOJA

O crescimento agrícola no Brasil se caracterizou por uma forte expansão da área total plantada a partir de 1990, rompendo com um padrão de crescimento agrícola. A soja foi a cultura que mais contribuiu com esta expansão, registrando aumento na taxa média anual de crescimento de 3,6%, entre 1991 e 2001, para 13,8% em 2001 a 2004. No Centro-Oeste, a cultura apresentou aumento de 118% na área plantada, entre os anos de 1990 e 2004 (Brandão et al., 2006). Após esse período, até 2010 não se observou aumentos anuais superiores ao da safra 2004/2005, que foi quase 23 milhões de hectares (Conab, 2010b).

Segundo a Conab (2010b), espera-se crescimento entre 1,3% e 3,1% na área a ser cultivada com soja em 2010/2011, o que corresponde aos totais de 23,8 milhões e 24,2 milhões de hectares, respectivamente. De 2008/2009 para 2009/2010, a área plantada cresceu de 21 milhões de hectares para um pouco mais de 23 milhões de hectares. Isso se deve principalmente ao crescimento nas regiões Centro-Oeste, com acréscimo de 199,2 mil hectares, e na região Sul, onde se registrou crescimento de 197,9 mil hectares. Os Estados mais representativos neste índice, em cada região, são Mato Grosso e Paraná, respectivamente. A produção nacional de soja foi de 68,7 milhões de toneladas na safra 2009/2010, representando incremento de 11,5 milhões de toneladas em relação à safra anterior (Conab, 2010b).

A influência da soja no avanço da fronteira agrícola e na política econômica é assunto amplamente discutido, principalmente a respeito das implicações da expansão da área plantada sobre o meio ambiente. Fearnside (2001) relata que a soja representa uma ameaça para a biodiversidade tropical no Brasil. O autor defende o desenvolvimento de estratégias efetivas para conter e minimizar o impacto ambiental do cultivo da soja. Ainda, alerta que algumas medidas devem ser tomadas, como a proteção efetiva de áreas antes da delimitação da fronteira da soja, eliminação de alguns subsídios e realização de estudos para avaliar os custos dos impactos sociais e ambientais associados à expansão da cultura.

Brandão et al. (2006), entretanto, sugerem não haver razões para se temer que a expansão de área com a soja ameace a política ambiental brasileira. O assunto é polêmico no que diz respeito ao binômio “impacto econômico *versus* ambiental”, sobretudo no que se refere às discussões de ordem política, social e ambiental (Tilman, 1999; Nepstad et al., 2006; Coronel et al., 2008; Queiroz, 2009; Souza et al., 2010).

O Brasil é o segundo maior produtor de soja em grão do mundo, atrás apenas dos Estados Unidos (FAO, 2010). As exportações brasileiras de soja grão, em 2010, atingiram cerca de 29,0 milhões de toneladas, resultando numa receita cambial equivalente a US\$ 11,0 bilhões. O farelo de soja foi responsável por quase 43% deste faturamento (Conab, 2011). O principal importador de grãos de soja do Brasil, em 2009, foi a China, responsável por 56% do total de grãos exportados. Já para o farelo de soja, os principais países foram França e Países Baixos, responsáveis por quase 41% do total de 12 milhões de toneladas exportadas pelo Brasil (Conab, 2010a).

A soja começou a ser explorada devido à crescente demanda de óleo comestível. O seu cultivo comercial para produção de grãos no Brasil se deu a partir dos anos 1940, em algumas regiões do Rio Grande do Sul e São Paulo. A produção de soja começou a ser incentivada pelo governo e foi introduzida no sistema de rotação de culturas junto com o trigo. Nos anos de 1970, a produção nacional alcançou crescimento de 32% ao ano (Bonato & Dallagnol, 1985). A partir daí, a soja passou a assumir, definitivamente, papel de destaque na economia brasileira. Com a crise do café o governo introduziu programa visando substituir o café por culturas alimentares como o trigo e a soja. Inicialmente o agronegócio da soja, representado principalmente pelo farelo e óleo, era voltado para o mercado interno. Com o desenvolvimento de políticas de crédito subsidiado pelo governo, o aumento da produção do grão e, conseqüentemente, das exportações, foi estimulado. Após esse período, as perspectivas da cadeia produtiva da soja estiveram voltadas para o mercado europeu, onde se concentram os principais importadores do farelo de soja (Kaimowitz & Smith, 2001; Cassuce & Silva, 2006).

Destacam-se na evolução da cadeia produtiva da soja, as indústrias esmagadoras, que tiveram papel fundamental no desenvolvimento do agronegócio e no aumento da produção e do cultivo da soja. Outro fator que também contribuiu para o desenvolvimento desse complexo foi a articulação da indústria de esmagamento com a atividade integrada de pecuária. A sua utilização na alimentação animal, principalmente para suínos e aves, tem despertado o promissor mercado de rações. Com o aumento deste

complexo e a expansão da produção de suínos e aves, a necessidade de se aumentar também a produção de soja foi imediata (Cassuce & Silva, 2006).

Diversos produtos da soja são utilizados historicamente como alimentos, principalmente pela população da Ásia oriental. Os povos dessa região têm, tradicionalmente, na sua dieta vários processados da dessa cultura, como o nimame (grão integral cozido), o edaname (grão verde e fresco), o extrato hidrossolúvel de soja (leite de soja), o tofu (queijo), o molho de soja ou shoyu, entre outros (Rosa et al., 2009).

Os principais produtos da soja, entretanto, ainda são óleo e farelo. O farelo é matéria-prima de vários derivados proteicos, como a proteína texturizada, o isolado proteico e o concentrado proteico. O óleo é destinado à alimentação e uso em indústrias de alimentos e, mais recentemente, para a produção de biodiesel. Além do seu alto valor energético, várias outras características da soja têm despertado o interesse da indústria por essa cultura. Atualmente essa cultura tem sido matéria-prima de vários produtos, tanto para alimentação humana, quanto animal (Wilson, 2004).

A demanda mundial de soja pelos consumidores tem aumentado devido ao apelo da qualidade nutricional de produtos que a utiliza como matéria-prima. Na alimentação humana, a soja entra na composição de produtos como embutidos, chocolates, temperos para saladas, misturas para bolo, entre outros produtos (Paetau et al., 1994; Golbitz, 1995; Lusas & Riaz, 1995; Mujoo et al., 2002; Pérez et al., 2008; Amadou et al., 2009; Cheng et al., 2009; Weng & Chen, 2010). Essas pesquisas têm buscado melhorar o processamento da soja e o seu uso tecnológico, pela elaboração de diversos subprodutos alimentícios com maior valor agregado. Um bom exemplo nesse sentido é o corrente uso da proteína texturizada de soja na elaboração de produtos cárneos, com destaque para os chamados “embutidos”. Isso acarreta redução de aproximadamente 13% no custo, além de aumentar o teor proteico e diminuir o teor de gordura final do produto processado (Singh et al., 2008; Rosa et al., 2009).

2.2 TEOR DE PROTEÍNA NO GRÃO EM SOJA

A soja é caracterizada principalmente pelo seu alto nível de proteína. Assim, representa um alimento de alto valor energético e, ainda, sua proteína é de alta qualidade (Santos et al., 2005; Zarkadas et al., 2007). Além do valor nutricional, são também bastantes exploradas as propriedades tecnológicas da proteína de soja. A partir da farinha

desengordurada do grão podem-se obter diversos derivados proteicos como a proteína texturizada, os concentrados e os isolados proteicos na indústria de alimentos. O uso desses produtos, com altos teores proteicos, pela indústria alimentícia tem aumentado devido às suas propriedades tecnológicas como emulsificante e estabilizante, o que ainda agrega valor ao produto por melhorar o rendimento e a qualidade nutricional do produto final (Rosa et al., 2009). Devido ao aumento da utilização da proteína de soja nas diversas áreas da alimentação, espera-se para o próximo ano um crescimento de 24% no consumo de farelo, em relação a 2009/2010 (Souza et al., 2010).

O esforço da indústria em elaborar produtos com altos níveis de proteína e reduzidos teores de gordura tem sido estimulado pelo interesse do consumidor em dietas mais saudáveis. Com base nisso, vários produtos combinados com a proteína de soja têm sido elaborados (Pérez et al., 2008; Rosa et al., 2009; Amadou et al., 2009; Cheng et al., 2009; Weng & Chen, 2010). Assim, a indústria consegue aumentar ainda mais a qualidade nutricional dos produtos, sem que haja alteração significativa nos seus custos.

O farelo de soja é o produto resultante da extração do óleo, apresentando, portanto, maiores teores de proteína que o grão, atingindo em torno de 47%. Entre os produtos provenientes da soja e com alto teor proteico, podem-se destacar a proteína de soja texturizada e o isolado proteico (proteína isolada de soja), que contém até 90% de proteína. Esses produtos se destacam por representarem alternativas para substituição das proteínas animais. Bonacina (2006) informa que a proteína de soja texturizada tem sido largamente utilizada na fabricação de embutidos, devido, principalmente, à sua boa capacidade de retenção de água e à estabilidade em emulsões. Outros autores confirmam a capacidade do produto para melhorar propriedades tecnológicas tais como formação e estabilidade de emulsão (Silva, 2007; Albuquerque et al., 2009). O mercado de produção de barras de cereais também tem-se mostrado em grande ascensão, com demanda bastante promissora por proteína de soja texturizada como constituinte desses produtos (Freitas & Moretti, 2006).

Outro importante subproduto proteico da soja é o extrato solúvel, também chamado de “leite de soja”. O extrato de soja não contém lactose e colesterol, sendo indicado para portadores de intolerância à lactose ou para indivíduos com dieta reduzida em colesterol. Um quilo de soja fornece cerca de seis a nove litros de extrato de soja, restando, aproximadamente, 700 g de resíduo com alto teor proteico. A composição química do extrato de soja varia em função da matéria-prima utilizada e do processamento

empregado (Oliveira, 1981; Tashima & Cardello, 2003). O principal produto do extrato solúvel de soja é o tofu, o queijo de soja. O tofu é o produto obtido da precipitação das proteínas, pela adição de sais de cálcio e de magnésio ou ácidos, resultando num gel que advém da formação de uma rede proteica com retenção de água, lipídeos e outros constituintes. Sua textura é lisa, macia e elástica, sendo importante fonte de proteína, minerais e vitaminas, ao mesmo tempo em que apresenta baixa proporção de gorduras saturadas e ausência (total) de colesterol. O tofu representa um alimento saudável, com alto valor nutritivo e de custo reduzido. Devido a tais propriedades é bastante utilizado em preparações alimentícias como substituto de ovos, queijos, carnes e outros alimentos de origem animal (Ciabotti et al., 2007). Nos Estados Unidos, 25% da produção de fórmulas (leite para crianças) são derivadas de soja (Bhatia & Greer, 2008).

Deve-se destacar também que a proteína de soja tem um bom balanço de aminoácidos essenciais (isoleucina, leucina, lisina, metionina, fenilalanina, treonina e triptofano). Os aminoácidos metionina e cistina encontram-se em baixas concentrações, porém, os teores de lisina e triptofano são elevados no grão de soja. Assim, devido ao alto valor nutricional na dieta, a soja destaca-se também na indústria de produção de rações. Para a avicultura do Brasil e de outros países, é um dos insumos de maior importância na fabricação de rações (Santos et al., 2005).

A sua importante contribuição nutricional na dieta, tanto na quantidade como na qualidade protéica, tem despertado o uso combinado com outros grãos na elaboração de rações para diversos fins, principalmente na alimentação de aves e suínos (Santos et al., 2005; Kemp, 2006; Bandegan et al., 2010; Terrazas et al., 2010; Baker et al., 2011). Lima (1999) reportou grandes perspectivas de utilização dos derivados da soja na alimentação de suínos e aves, sendo o farelo o mais utilizado como fonte proteica, e em associação com o milho, como fonte de energia. Hernández et al. (2007) concluíram que a inclusão do farelo de soja na dieta reduz os custos com alimentação de peixes.

As propriedades da soja não estão restritas à tecnologia industrial. Tem-se observado também que a soja desempenha papel preventivo e terapêutico em doenças cardiovasculares, osteoporose e em alívio dos sintomas da menopausa (Potter, 1998; Albertazzi et al., 1998), sendo classificada pela Anvisa (Agência Nacional de Vigilância Sanitária) como um alimento funcional. Têm sido observadas, ainda, reduções na concentração de colesterol LDL (*low density lipoprotein*) no sangue, com a ingestão de alimentos que contêm proteína de soja, tanto em seres humanos quanto em animais

(Anderson et al., 1995; Anthony, 2000). Somado a estas propriedades funcionais, a soja apresenta também efeito protetor contra o desenvolvimento de câncer de mama em mulheres (Lee et al., 2009).

Hoje, diversas cultivares de soja estão disponíveis no mercado. No entanto, a escassez de informações intrínsecas sobre as propriedades nutricionais e tecnológicas dessas cultivares não permitem aproveitar o potencial específico de cada uma. Alguns trabalhos buscam avaliar a contribuição diferenciada das cultivares na qualidade do extrato de soja. Khatib et al. (2002) avaliaram a qualidade de tofu obtido a partir de quatro variedades de soja (Macon, Saturno, OHIOFG1 e OHIOFG2). As variedades mostraram diferenças significativas para os efeitos de rendimento, retenção de água, viscosidade e textura do tofu, sendo que a variedade Macon destacou-se positivamente. Assim, os autores sugeriram que se deve considerar os efeitos intrínsecos na seleção de cultivares destinadas à produção de tofu.

Evans et al. (1997) encontraram correlação significativa entre o teor de proteína no grão, o teor de proteína no extrato de soja e a qualidade do queijo tofu. E, ainda, as cultivares de soja com teores de proteínas mais altos e grãos maiores resultaram em maiores rendimentos e melhor qualidade do tofu. Aziadekey et al. (2002) avaliaram a influência de doze genótipos em diferentes locais e anos, sobre a qualidade do extrato de soja. Observaram que o efeito de genótipo foi a principal fonte de variação para a maioria das características de qualidade, tanto do extrato, quanto do subproduto tofu. A interação de genótipo com ambientes foi importante para as características de consistência e viscosidade do tofu. Assim, os autores recomendaram avaliações considerando vários ambientes no desenvolvimento de cultivares de soja para qualidade do extrato de soja e do tofu.

O Ministério da Agricultura, Pecuária e Abastecimento (Brasil, 1993) define, pela Portaria nº 795, de 15 de dezembro de 1993, as normas de identidade, qualidade, embalagem, marcação e apresentação do óleo e do farelo de soja. Essa regulamentação classifica o farelo de soja em três tipos, e um dos critérios de classificação é o teor de proteína. Farelos com teor mínimo de 48% de proteína são classificados como tipo 1; mínimo de 46%, como tipo 2; e mínimo de 44%, como tipo 3. Desta forma, grãos de soja com maiores teores de proteína produzirão farelos com melhores chances de comercialização e, talvez, com melhores preços.

O teor de proteína de cultivares de soja no Brasil é em torno de 40%, porém, este valor pode variar bastante dependendo da região. A média dos teores de proteína das variedades lançadas no Rio Grande do Sul, entre 1970 e 1980, era de 40,7%, e entre 1991 e 1996, de 39,0% (Bonato et al., 2000). Nos Estados Unidos, cultivares plantadas na região Norte, entre 1949 e 1998, também apresentavam teor médio de proteína de 40,7%, e na região Sul, de 41,1% (Yaklich et al., 2002). Albrecht et al. (2009) verificaram teores de proteína variando entre 35,5% a 43,2% em cultivares do Paraná. Minuzzi et al. (2009) verificaram que os teores de proteínas variaram de 31% a 35%, entre quatro cultivares plantadas em dois locais de Mato Grosso Sul, em dois anos agrícolas (2004/2005 e 2005/2006).

O teor de proteína é influenciado por fatores ambientais, principalmente durante o período de enchimento dos grãos (Pípolo, 2002; Rangel et al., 2004). Segundo Gonçalves et al. (2007), quando a temperatura aumenta e se aproxima de 30°C, o acúmulo de proteína total na semente de soja é afetado, possivelmente favorecendo o metabolismo no sentido de biossíntese. Albrecht et al. (2009) observaram que a antecipação da semeadura influenciou os teores de proteína e de óleo no grão. Os autores associaram períodos de maiores restrições hídricas com a elevação nos teores de proteína. Aumentos nos teores de proteína e reduções nos teores de óleo como resultado da elevação da temperatura, sob restrição hídrica durante o enchimento de grãos, também já haviam sido relatados por Dornbos & Mullen (1992). Por outro lado, Wilson (2004) relata que o teor de proteína em soja é inversamente proporcional à temperatura de crescimento da planta. O autor acrescenta que plantas sob estresse hídrico concentram mais óleo e menos proteína no grão pelo fato de as temperaturas mais altas reduzirem a disponibilidade de água à planta. Enfim, a causa da variação nos teores dos constituintes do grão em soja, sob diferentes temperaturas, ainda não é totalmente conhecida, podendo também ser afetado pelo estágio de desenvolvimento da planta que o estresse ocorre. A situação relatada por Wilson (2004) pode ser explicada pelo fato de que fixação biológica do nitrogênio ser muito sensível à deficiência hídrica (Serraj et al., 1999), comprometendo o conteúdo de proteína no grão. A evidência de que o estresse hídrico e altas temperaturas diminuem o acúmulo de proteína no grão de soja faz sentido, uma vez que, esses são os principais fatores ambientais limitantes à fixação biológica do nitrogênio, afetando a simbiose em todos os estádios (Hungria et al., 2001).

Em geral, variedades de soja produtivas e com altos teores de óleo satisfazem melhor aos produtores e às indústrias de óleo, enquanto o alto teor de proteína vem agregar valor aos produtos derivados do farelo. Todavia, o aumento do teor de proteína no grão tem sido dificultado por este caráter apresentar correlações negativas com teor de óleo e com produtividade de grãos. Vários estudos demonstram tais correlações, com teor de óleo (Johnson et al., 1955; Kwon & Torrie, 1964; Thorne & Fehr, 1970; Hymowitz et al., 1972; Simpson Junior & Wilcox, 1983; Bonato et al., 2000; Cober & Voldeng, 2000; Feng et al., 2004; Silva, 2008) e com rendimento de grãos (Voldeng et al., 1997; Wilcox & Guodong, 1997; Cober & Voldeng, 2000; Marega Filho et al., 2001; Silva, 2008). Vale ressaltar que no caso deste último caráter, Wilcox & Guodong (1997), embora evidenciaram a correlação negativa em populações de crescimento indeterminado, não a encontraram em populações de hábito determinado. Enfim, o conhecimento da associação entre caracteres é de grande importância nos trabalhos de melhoramento, pois a seleção de um caráter pode provocar mudanças indesejáveis em outros, quando a correlação genética é existente (Cruz & Regazzi, 2001).

Outros estudos revelam os efeitos da ênfase do melhoramento genético da soja dirigida, principalmente, para aumento na produção de grãos e teor de óleo, em detrimento do conteúdo de proteína no grão. Guodong & Jinling (1989) avaliaram, nos Estados Unidos, os teores de óleo e de proteína em variedades soja, em função da década e da posição geográfica da avaliação. As variedades liberadas na década de 1950 superaram, em relação ao teor de proteína, as variedades liberadas nos anos de 1960, 1970 e 1980. Os autores argumentaram que, nos anos seguintes a 1950, iniciou-se o crescimento da demanda por óleos comestíveis. Isso pode explicar a queda no teor de proteína, já que o óleo, naquela época era considerado produto principal e de maior valor, é negativamente correlacionado com o teor de proteína.

Mahmoud et al. (2006) também apresentaram evidências de redução no teor de proteína das cultivares de soja liberadas nos Estados Unidos, entre 1940 e 2005. Segundo os autores, essa redução pode ser atribuída à atenção especial dada ao aumento da produtividade de grãos, que é negativamente correlacionada com o teor de proteína. Eles dividiram as cultivares em três grupos, baseados no período de liberação comercial. O primeiro, formado por cultivares que apresentavam teores entre 38% e 43% de proteína, foi constituído principalmente por genótipos que formam a base do *pool* gênico do Norte dos Estados Unidos. O segundo grupo foi formado por cultivares com teores variando de 36%

a 41%, e o terceiro, entre 36% a 39%. Este último grupo é composto por cultivares liberadas mais recentemente, mais próximo ao ano de 2005.

No Brasil, Bonato et al. (2000) também observaram queda no teor de proteína no grão, em cultivares comercializadas no Rio Grande do Sul após 1980, especialmente em 1996 e 1997. Os autores sugerem que tal redução esteja ligada ao uso de determinados genitores nos programas de melhoramento.

A aparente correlação negativa entre teor de proteína e produção de grãos em soja, entretanto, não parece oferecer limitações definitivas ao desenvolvimento de variedades que sejam, simultaneamente, produtivas e com altos teores de proteínas (Wehrmann et al., 1987; Holbrook et al., 1989). Wilcox & Cavins (1995), utilizando uma série crescente de retrocruzamentos, obtiveram, ao final do terceiro deles, progênies com alto teor de proteína e alta produtividade. Mello Filho et al. (2004), após avaliarem oito populações de soja, também verificaram que é possível obter linhagens com teor proteico mais elevado, mantendo-se as médias de produção de grãos de seus respectivos genitores recorrentes.

A correlação negativa entre os teores de proteína e de óleo parece impor limitações maiores à seleção de genótipos que combinem características favoráveis para os dois atributos. Miller & Fehr (1979), após praticar um ciclo de seleção recorrente para teor de proteína no grão, obtiveram aumento de 1,5 g/100g na média do teor de proteína, embora com redução de 0,7 g/100g no teor de óleo. Da mesma forma, Brim & Burton (1979), após seis ciclos de seleção recorrente, verificaram que os ganhos no teor de proteína foram obtidos em detrimento do teor de óleo. Entretanto, também neste caso, alguns trabalhos têm mostrado que através de seleção recorrente para maiores teores de proteína é possível manter o teor de óleo, sem reduções significativas (Xu & Wilcox, 1992; Wilcox, 1998; Chung et al., 2003; Mello Filho et al., 2004).

Essa possibilidade encontra sustentação na ampla variabilidade detectada no germoplasma de soja, tanto para teor de proteína, quanto para teor de óleo (Wilson, 2004; Iqbal et al., 2010). A concentração de proteína em 11.726 acessos de soja, da coleção de germoplasma do Departamento de Agricultura dos Estados Unidos (USDA), variou de 34,1% a 56,8%, com média de 42,1% (Wilson, 2004). Também verificaram-se que cultivares adaptadas ao sul dos Estados Unidos apresentam maiores teores médios de proteína no grão do que cultivares adaptadas ao norte e, ainda, maior amplitude nos teores

de proteína e de óleo. É raro, porém, encontrar um genótipo que combine concentrações relativamente altas de ambos os teores. Apesar disso, Wilson (2004) sustenta que é possível selecionar combinações gênicas com teores de proteína maiores que o normal, mantendo-se aceitáveis níveis de óleo e de produtividade.

Oliveira (2007) reuniu informações sobre 15.558 acessos de soja do germoplasma mantido pelo USDA tendo encontrado variação entre 31,7% e 57,9%, nos teores de proteína, e entre 8,1% e 25,4%, nos teores de óleo. Isso demonstra a existência de ampla variabilidade para estes caracteres.

Em estudo objetivando aumentar os teores de proteína e de óleo no grão, Scott & Kephart (1997) avaliaram o potencial de cruzamentos entre linhagens de soja introduzidas e genótipos adaptados. Encontraram variabilidade genética significativa, em ambos os caracteres, entre as progênies de dois dos cruzamentos realizados. Thorne & Fehr (1970) também já haviam evidenciado variabilidade significativa entre populações provenientes do cruzamento de genótipos exóticos com adaptados, para teores de proteína e de óleo, bem como para o conteúdo de ambas as características somadas (proteína + óleo). Os autores concluíram que o germoplasma exótico é capaz de transferir as características de elevados teores de proteína para suas progênies. Concluíram, então, que a seleção para alta produtividade e altos teores de proteína é possível a partir do cruzamento entre germoplasma exótico e adaptado.

2.3 TEOR E COMPOSIÇÃO DO ÓLEO EM SOJA

O Brasil é o quarto produtor de óleo de soja no mundo, tendo produzido 6,9 milhões de toneladas em 2010, o que representou quase 17% da produção mundial. O país é ainda o segundo maior exportador de óleo de soja, atrás somente da Argentina (Souza et al., 2010; USDA, 2010). O crescimento médio anual de óleo exportado, entre 1995 e 2004, foi de 6,5% ao ano (Coronel et al., 2008).

O teor de óleo na soja comercial varia de 18% a 20% no grão (Bonato et al., 2000). Outras culturas oleaginosas apresentam maiores teores de óleo, porém, ainda não substituem o óleo de soja. As perspectivas, entretanto, são de que a demanda por óleo irá aumentar ainda mais e, conseqüentemente, a necessidade de cultivares com maior produtividade de óleo. Wilson (2010) avaliou a previsão de consumo de óleo até 2020 e concluiu que o aumento, para alimentação, poderá chegar a 5,2 bilhões de toneladas por

ano; e para outros fins prevê aumento na mesma proporção. O autor conclui, ainda, que essa demanda deverá ser suprida principalmente pela soja.

A produção de óleo de soja tem sido estimulada também pelo aumento da demanda internacional no seu uso como combustível (Souza et al., 2010). O óleo de soja corresponde a 80% do volume produzido de biocombustível no Brasil, enquanto a gordura animal e as outras oleaginosas respondem, respectivamente por 15% e 5% (Garcia & Romeiro, 2009). Em 2004, 95% do óleo de soja nos Estados Unidos foi usado pela indústria de alimentos, no entanto, em 2008, 19% do consumo doméstico de óleo no país já foi destinado para usos não alimentares, principalmente à produção de biodiesel (Clemente & Cahoon, 2009).

Além da busca por cultivares com maiores teores de óleo no grão, o melhoramento genético vem definindo metas que visam melhoria na sua qualidade. A busca por características ligadas ao óleo na soja tem três objetivos principais: redução da atividade de lipoxigenase, redução dos teores de poli-insaturados e redução de ácidos graxos saturados. Os dois primeiros são responsáveis por melhorar, respectivamente, a qualidade sensorial e a vida de prateleira do óleo. O terceiro objetivo tem sido adicionado como resultado da preocupação com o excesso de gorduras saturadas na dieta humana (Burton, 1991).

Uma vez que a qualidade tecnológica e nutricional do óleo está em função da sua composição em ácidos graxos é necessário entender as características destes constituintes. Os óleos vegetais são constituídos por ácidos graxos que podem conter cadeias de 8 a 24 átomos de carbono, com diferentes números de insaturações (duplas ou triplas ligações). Dependendo da espécie, há variações também no tamanho das cadeias da molécula e no grau de insaturações desses ácidos. Em síntese, a qualidade do óleo se baseia na medida dos ácidos graxos (Costa Neto et al., 2000).

A análise genômica da biossíntese de lipídios em *Arabidopsis* revelou 614 genes envolvidos na síntese de ácidos graxos. Já na soja, esse número é bem maior, 1.127. A natureza multigênica das atividades envolvidas no metabolismo lipídico sugere, assim, um controle mais complexo na soja em relação ao de *Arabidopsis* (Schmutz et al., 2010).

A composição de ácidos graxos no óleo de soja comercial, em média, é: 10,8% de ácido palmítico (16:0), 3,3% de ácido esteárico (18:0), 21,5% de ácido oleico (18:1), 55,1% de ácido linoleico (18:2) e 4,8% de ácido linolênico (18:3). Nesta notação, os

valores entre parênteses indicam o número de átomos de carbono na molécula do ácido graxo e, em seguida, o número de insaturações (Sanibal & Mancini Filho, 2004).

Os ácidos graxos podem ser classificados quanto ao número de duplas ligações (insaturações) na sua cadeia de carbonos, são eles: ácidos saturados, quando não há duplas ligações; monoinsaturados, presença de uma só dupla ligação; e ácidos poli-insaturados, com duas ou mais insaturações. Os ácidos poli-insaturados são normalmente divididos em duas famílias, ômega-3 e ômega-6, sendo seus principais representantes, no óleo de soja, os ácidos linolênico e linoleico, respectivamente. Essa classificação deve-se à posição da primeira dupla ligação a partir do grupo metil (Sizer & Whitney, 2008; Insel et al., 2010). Os ácidos linoleico e linolênico são essenciais ao organismo humano, isto é, somente são adquiridos na dieta; e, o aumento de consumo destes ácidos, em relação aos ácidos graxos saturados, está associado à melhoria desempenho do cérebro, do coração e do sistema imunológico do organismo humano (Katan, 2009). Ainda que ambos exerçam efeitos supressores no surgimento de doenças cardiovasculares, câncer e inflamatórias, deve-se buscar a redução da proporção entre os ácidos ômega-6 e ômega-3 na dieta (Simopoulos, 2008; Deckelbaum, 2010; Farzaneh-Far et al., 2010). Segundo Simopoulos (2003), nos últimos cem anos a sociedade tem diminuído o consumo de ácidos ômega-3 e, conseqüentemente, a relação dos ácidos ômega-6/ômega-3 tem aumentado. Os autores associam esse aumento a maiores incidências de arterosclerose e outras doenças coronárias. Em síntese, os ácidos ômega-3 e ômega-6 são considerados benéficos à saúde humana; mas, para isto, deve-se manter também na dieta uma baixa relação ômega-6/ômega-3, com preferência de consumo de ômega-3.

O grupo dos ácidos graxos ômega-3 tem mostrado efeitos benéficos também sob condições psiquiátricas, doenças autoimunes e degenerativas. Os ácidos do tipo ômega-3 apresentam papéis coadjuvantes úteis na manutenção da função do cérebro durante o envelhecimento, particularmente em relação à doença de Alzheimer (Mazza et al., 2007; Freemantle et al., 2006; Harris et al., 2008; Kendall-Tackett, 2010). Além disso, Silva et al. (2008) apontam que a ingestão de ômega-3 pode ser usado com um efeito antidepressivo ou como terapia adjuvante com algum outro medicamento no tratamento da doença de Parkinson.

Outra vertente sobre a importância da composição de ácidos graxos é em relação à estabilidade oxidativa do óleo. Para isso, no óleo de soja são desejáveis maiores concentrações de ácidos monoinsaturados, isto é, especificamente de ácido oleico. Os

ácidos graxos poli-insaturados são mais susceptíveis à oxidação, fazendo do óleo de soja um produto relativamente instável, devido às altas concentrações de ácido linoleico (Fehr, 2007; Graef et al., 2009). A suscetibilidade do óleo à oxidação reduz sua vida útil e, ainda, produz substâncias indesejáveis, comprometendo sua qualidade sensorial. Por isso, para evitar a oxidação do óleo, a indústria utiliza o processo de hidrogenação, o qual, entretanto, além de acumular custos ao processamento industrial, introduz desvantagens nutricionais (Garcia & Romeiro, 2009).

O processo de hidrogenação, que aumenta a estabilidade do óleo, consiste em promover a redução de insaturações nos ácidos graxos, diminuindo a proporção de ácidos poli-insaturados. Durante esse processo, tem-se ainda a formação de ácidos graxos *trans*. Estas substâncias são consideradas prejudiciais à saúde do homem (Garcia & Romeiro, 2009). Nesse contexto, o melhoramento genético da soja para aumentar teores de ácido oleico e diminuir o de ácido linoleico no óleo poderá reduzir a necessidade de hidrogenação e, conseqüentemente, os custos de produção e a vida de prateleira de produtos derivados da soja (Fehr, 2007; Bachlava et al., 2008). Tudo isso vem caracterizar uma demanda clara da indústria por óleos mais estáveis, ao que se devem agregar requisitos de ordem nutricional, que representam, por conseguinte, estímulo à obtenção de genótipos com perfis de ácidos graxos mais atrativos.

Na indústria de biodiesel, a composição ideal do óleo é atribuída também à presença elevada de ácidos graxos monoinsaturados, presença reduzida de ácidos poli-insaturados e controlado conteúdo de ácidos saturados. Assim, o balanço entre estes ácidos graxos influenciará na qualidade e no rendimento do biodiesel. Especificamente, os ácidos graxos oleico (18:1) e palmitoleico (16:1) são os que apresentam, entre outras propriedades desejáveis ao biodiesel, melhor estabilidade oxidativa e boa condição de fluidez sob baixas temperaturas, o que se relaciona também com seu ponto de solidificação (Pinzi et al., 2009; Knothe, 2010). Daí a razão de a genética ser ferramenta valiosa para o desenvolvimento de óleos com perfis de ácidos graxos mais adequados à esta finalidade (Pinzi et al., 2009).

Vieira et al. (1999) observaram diferenças significativas entre cultivares de soja, quanto ao teor de ácido oleico, o que indica possibilidade de uso específico de cultivares com menores índices de insaturações para a produção de biodiesel. No germoplasma do Departamento de Agricultura dos Estados Unidos (USDA) foram encontrados valores de ácido esteárico variando entre 1,1% e 10,6%, de ácido palmítico

entre 8,0% e 18,4%, de ácido oleico entre 11,1% e 50,3%, de ácido linoleico entre 30,3% e 65,0% e de ácido linolênico entre 3,4% a 18,1% (Oliveira, 2007; Oliveira et al., 2010).

Além das técnicas tradicionais do melhoramento de plantas, a engenharia genética tem fornecido inovações em relação à qualidade do óleo de soja (Napier, 2007). Em 2010, o USDA aprovou para cultivo a soja PlenishTM High Oleic, da empresa DuPont. O genótipo apresenta teores de ácido oleico acima de 75% e, ainda, teores de ácidos graxos saturados abaixo de 20%. A Monsanto também já desenvolveu algo semelhante, a soja Vistive-GoldTM, que apresenta baixos teores de ácidos palmítico (3%) e poli-insaturados, e altos teores de ácido oleico (68% a 74%) (Wilson, 2010).

A busca de baixos índices de ácidos graxos poli-insaturados no óleo de soja, principalmente em relação aos ácidos linoleico e linolênico, parte tanto da indústria alimentícia, quanto da indústria de biodiesel. Devido à alta capacidade oxidativa destes ácidos, a vida de prateleira do óleo e do biodiesel é diminuída com o aumento de suas concentrações. Por outro lado, há demanda por óleos com baixos teores de ácidos graxos saturados, devido à relação destes ácidos com problemas de saúde de ordem cardiovascular. Reduções do índice de saturação no óleo de soja podem melhorar a sua competitividade comercial frente aos óleos de canola, girassol e milho, os quais apresentam, em ordem crescente, menores índices de ácidos graxos saturados (Tyson & McCormick, 2006).

Esforços para melhorar a composição dos ácidos graxos no óleo, além de trazer vantagens competitivas para a soja, podem expandir a sua aplicação na indústria alimentícia e de biodiesel (Wilson, 2010). Em razão disso, busca-se no óleo de soja o aumento dos teores de ácidos monoinsaturados, que tem o ácido oleico como principal representante, e a redução nos teores de ácidos saturados, cujo principal representante é o ácido palmítico. Este objetivo promoveria, conjuntamente, melhor qualidade industrial, tecnológica e nutricional do óleo de soja.

A mutagênese tem sido uma ferramenta para obtenção de genótipos com diferentes composições em ácidos graxos no óleo de culturas como soja, girassol e canola (Bhatia et al., 1999). Os autores relacionam as principais linhagens de soja obtidas por mutagênese, com baixos teores de ácidos palmítico e linolênico, e com maiores teores de ácido oleico. Com reduzidos teores de ácido palmítico desatacam-se as linhagens C1726 (8,5% de ácido palmítico) e ELLP2 (7,1%) foram desenvolvidas por aplicação de metanosulfonato de etila (EMS) às cultivares Century e Elgin87, respectivamente (Wilcox

& Cavins, 1990; Stojsin et al., 1998). A linhagem A22 (7,8% de ácido palmítico) foi desenvolvida por mutação induzida com N-nitroso-N-metil-uréia (Fehr et al., 1991); e as linhagens A1937 e J3 (5,7% de ácido palmítico), com raios-X, ambas a partir do cultivar Bay (Takagi et al., 1995; Stojsin et al., 1998). Além disso, foi descoberta na linhagem N79-2077 uma mutação natural com 6,4% de ácido palmítico. A linhagem N79-2077-12, derivada dessa mutante e com baixo teor de ácido palmítico, foi obtida a partir do quinto ciclo de seleção recorrente para aumento do teor de ácido oleico (Burton et al., 1983; Burton et al., 1994).

Com altos teores de ácido oleico foram desenvolvidas, também por mutagênese, as linhagens M11 (30%), M23 (46%), KK21 (47%) e N98-4445A (54,9%) (Rahman et al., 1994; Rahman et al., 1996; Burton et al., 2006; Rahman et al., 2001; Anai et al., 2008). Takagi & Rahman (1996) observaram ausência de efeito materno e dominância parcial no loco que controla os teores de ácido oleico, em cruzamentos, considerando os recíprocos, provenientes da linhagem M23 e sua cultivar original, Bay. Esses autores ainda verificaram que o alto teor de ácido oleico na linhagem mutante era controlado por um alelo recessivo (*ol*), e o baixo teor da cultivar Bay, por um alelo com dominância parcial (*OL*) sob o alelo recessivo (*ol*).

Outras linhagens com teores de linolênico menores que 2,5% também foram desenvolvidas por mutagênese (Hammond & Fehr, 1983; Wilcox & Cavins, 1985; Fehr et al., 1992; Fehr & Hammond, 1996). As linhagens mutantes IL-8 e M-5 apresentam teores menores que 4,5% de ácido linolênico (Rahman et al., 1996). Ainda, os autores observaram que o loco que determina menor teor de ácido linolênico na linhagem M-5 é diferente do loco da mutante KL-8, também com menor teor desse ácido graxo. Segundo Clemente & Cahoon (2009), teores de ácido linolênico abaixo 4% são considerados baixos, e valores menores que 2% muito baixos.

Uma vez que os teores dos ácidos palmítico e oleico seriam caracteres alvos de um programa de melhoramento para qualidade do óleo em soja, há interesse na correlação destes com outros caracteres agronômicos. Ndzana et al. (1994) avaliaram a influência da redução do ácido palmítico sobre características agronômicas e nas sementes. Linhagens com reduzidos teores de ácido palmítico tiveram menor produtividade, menos óleo na semente, maiores teores de proteína, menor teor de ácido oleico e maiores teores dos ácidos linoleico e linolênico. Horejsi et al. (1994), através do retrocruzamento de linhagens com reduzidos teores de ácido palmítico com cultivares de alta produtividade, obtiveram

linhagens com conteúdo reduzido deste ácido graxo, e com produtividade comparável à dos genitores recorrentes. Outra observação dos autores foi que as linhagens com baixos teores de ácido palmítico geralmente tinham reduzidos teores de óleo.

Rebetzke et al. (1998) avaliaram o efeito da seleção para reduzidos teores de ácido palmítico sobre características agrônômicas e qualidade da semente. Foram obtidas linhagens provenientes do cruzamento entre N87-2122-4, linhagem portadora do alelo *fapnc* e com reduzido teor de ácido palmítico, e duas linhagens com teores normais desse ácido graxo. Linhagens homozigotas com o gene principal para reduzido teor de ácido palmítico tiveram rendimento de grãos 10% menor do que as linhas homozigotas com o gene normal. Observaram correlação negativa entre os teores de ácido palmítico e ácido oleico, e correlação positiva entre o primeiro e o ácido linolênico. Os genes que reduzem o teor ácido palmítico não influenciaram significativamente os teores de proteína e de ácido linoleico no grão.

Cherrak et al. (2003) avaliaram a qualidade do óleo e a produtividade, numa população proveniente do cruzamento entre a cultivar Anand (normal) e a linhagem N97-3708-13 (baixos teores dos ácidos palmítico e linolênico). Encontraram correlações positivas, porém fracas, entre teor de ácido palmítico e produção de sementes ($r = 0,12$), e entre os teores de ácido palmítico e de óleo no grão ($r = 0,13$). Não foram encontradas correlações entre o teor de ácido linolênico e a produção de sementes, ou entre esse teor e conteúdo de óleo no grão. Os autores concluíram que reduções nos teores dos ácidos palmítico e linolênico não acarretam, necessariamente, queda na produtividade.

Mais recentemente, Bachlava et al. (2008) encontraram correlação negativa entre: teor de ácido oleico e produtividade; teores de ácido oleico e linoleico; teores de ácido oleico e linolênico; e teores de ácido oleico e ácido palmítico. Esses resultados sugerem que uma seleção para maiores teores de ácido oleico poderá resultar em menores teores de ácidos poli-insaturados. Os autores relatam ainda que a obtenção de cultivares com maiores teores de ácido oleico pode ser dificultada pela redução na produtividade de grãos da soja. A correlação negativa entre os teores de ácido oleico e linoleico também foi encontrada por Alt et al. (2005).

Outro aspecto importante no melhoramento visando a qualidade do óleo em soja é o efeito do genótipo nos diversos ambientes. Hou et al. (2006) evidenciaram que a temperatura é o fator ambiental predominante na constituição dos ácidos graxos, sendo o ácido linoleico o mais vulnerável às mudanças ambientais. Os autores observaram

significância no efeito da interação de genótipos com ambientes, para os ácidos graxos palmítico, esteárico, oleico, linoleico e linolênico. Outros trabalhos relatam aumento dos teores de ácidos graxos poli-insaturados e redução do teor de ácido oleico, sob baixas temperaturas (Cheesbrough, 1989; Dornbos & Mullen, 1992; Heppard et al., 1996; Wilson, 2004; Lannal et al., 2005). Esses estudos têm buscado compreender o aumento dos teores dos ácidos graxos poli-insaturados em temperaturas baixas. Essa correlação deve-se, principalmente, à ação das enzimas desnaturase durante a fase de enchimento dos grãos. A temperatura apresenta também correlação positiva com o teor de óleo, mas a correlação é negativa com o teor de proteína (Wilson, 2004).

2.4 BIODIESEL E SOJA

A utilização de biodiesel como combustível tem apresentado um potencial promissor. Além das contribuições ao meio ambiente, o biodiesel é fonte estratégica de energia renovável em substituição ao óleo diesel e a outros derivados do petróleo. Vários países vêm investindo fortemente na produção e viabilização comercial desta fonte energética (Costa Neto et al., 2000).

A expectativa nas reduções das reservas de petróleo e a preocupação com suas consequências ambientais negativas vêm estimulando a substituição do petróleo por combustíveis alternativos. Uma das grandes preocupações ambientais com o uso de diesel é com os impactos ambientais devido aos efeitos da emissão de gases. O biodiesel é, ambientalmente, menos poluente, atóxico e biodegradável do que o diesel do petróleo (Pinto et al., 2005). Outra vertente para o estímulo da produção de biodiesel é em relação ao aumento da demanda e, conseqüentemente, aumento do preço do petróleo, que ainda é a principal fonte primária de energia no mundo (Masiero & Lopes, 2008). O custo do biodiesel, no entanto, ainda é o principal obstáculo para a comercialização do produto (Demirbas, 2007).

A utilização de óleos comestíveis na fabricação de biodiesel é controversa (Ponti & Gutierrez, 2009). Algumas Organizações Não Governamentais (ONGs) e movimentos sociais, por exemplo, a *Global Forest Coalition*, sugerem que a produção de biocombustíveis a partir de matérias-primas alimentares pode ser a principal causa do aumento dos preços no mercado global de alimentos. Segundo a Organização das Nações Unidas para Alimentação e Agricultura (FAO, 2008), outros fatores que podem inflacionar

os preços dos alimentos incluem a escassez de sua oferta devido a más condições climáticas em regiões produtoras e as mudanças nos hábitos alimentares, gerando forte demanda por alimentos.

A influência da indústria do biodiesel sob os preços de óleo comestível foi observada por Bergsma et al. (2006). Segundo os autores, o preço do óleo de canola, entre 2001 e 2007, subiu quase 100%, em função de sua utilização na produção de biodiesel. Em consequência disso, este óleo foi substituído pelo óleo de girassol em aplicações alimentícias. Pinzi et al. (2009) defenderam que o foco da matéria-prima para produção de biodiesel deva ser colocado sobre óleos não comestíveis, em vez daqueles comestíveis, para se ganhar a aceitação social do biodiesel.

Outra preocupação com o uso de biocombustíveis é o esgotamento dos recursos ecológicos, devido às práticas de agricultura intensiva na produção agrícola (Pinzi et al., 2009). A idéia de que o uso de biocombustíveis pode provocar aumento de áreas cultivadas e, ainda, inflacionar o preço dos alimentos em longo prazo tem sido o patamar desta preocupação. Porém, é discutida também a possibilidade de a população, principalmente de países em desenvolvimento, se beneficiar com o aumento de oportunidades de emprego, pois a cadeia para obtenção de biodiesel é complexa e requer mais mão-de-obra que a indústria do petróleo (Pinzi et al., 2009; Rajagopal & Zilberman, 2007).

A FAO (2008) ameniza as preocupações acerca dos impactos do biodiesel sobre o meio ambiente, informando que a proporção de terras dedicadas à produção de biocombustíveis no mundo é muito pequena, sendo de 1% no Brasil e Europa, e 4% nos Estados Unidos. A instituição ainda estimula a bioenergia por esta representar boa oportunidade para impulsionar a economia rural e reduzir a pobreza, desde que a produção obedeça critérios de sustentabilidade. Entretanto, para se garantir o critério de sustentabilidade, a FAO defende o suporte da produção com incentivos e políticas públicas, tais como estabelecimento de metas obrigatórias para a utilização de biocombustíveis e incentivos fiscais que favoreçam os biocombustíveis em relação aos combustíveis fósseis. No Anuário FAO (2008) é evidenciada a necessidade de revisão urgente das políticas de biocombustíveis e dos subsídios relacionados. O documento sugere que as políticas atuais tendem a favorecer os produtores de países desenvolvidos, em relação aos produtores de países em desenvolvimento.

Embora um número considerável de publicações sobre biodiesel possa ser encontrado na literatura, vários problemas ainda são presentes no aspecto técnico econômico e social. Segundo Rajagopal & Zilberman (2007), a aplicação de uma política de incentivo ao biodiesel pode estimular a agricultura sustentável. Mas, a produção de biocombustíveis representa riscos substanciais. Entre as situações de risco, os autores descrevem duas: estoque positivo de petróleo, como assistido em 1980, ou estoques negativos de biocombustíveis, devido à queda na safra da matéria-prima. Por causa destes riscos, a adoção vem sendo e continuará a ser desencadeada por políticas que mantêm a demanda por biocombustíveis. O biodiesel pode representar uma alternativa verdadeiramente competitiva ao diesel. Para isso, a isenção de impostos e os subsídios às culturas energéticas são medidas necessárias (Demirbas, 2009).

Entre as medidas de incentivo citam-se aquelas que promovem redução nos custos de matérias-primas e de processamento. Isto porque, a matéria-prima é responsável por cerca de 70% a 95% do custo total de produção do biodiesel. As isenções fiscais serão instrumento fundamental para se inserir o biodiesel como combustível alternativo para o transporte, em futuro próximo (Rajagopal & Zilberman, 2007; Demirbas, 2009; Pinzi et al., 2009).

Pinto et al. (2005) consultaram bases de catalogação de artigos científicos e constataram crescimento no interesse em biodiesel, pelos pesquisadores. O número de artigos publicados relacionados à palavra-chave "biodiesel" cresceu de quinze para 53, e o número de patentes registradas, de nove para trinta, de 2000 para 2004. O Brasil é responsável por quase 11% dessas patentes. Dos 134 artigos analisados, 86 eram exclusivamente de universidades, 26 de universidades em colaboração com centros de pesquisa ou indústrias, e 22 de centros de pesquisa em associação com indústrias. Por outro lado, entre 84 patentes, apenas três foram apresentadas por universidades.

Entre as propriedades do biodiesel, relacionadas à sua qualidade, citam-se principalmente o índice de cetano e a estabilidade. Cadeias de ácidos graxos lineares e mais longas promovem maior índice de cetano e, conseqüentemente, melhor qualidade na ignição dos motores. O aumento no nível de insaturação dos ácidos graxos é o fator mais significativo na redução do índice de cetano. Em relação à estabilidade do biodiesel, como descrito anteriormente, ácidos graxos com maior número de insaturações promovem menor estabilidade (Knothe, 2005). Além das propriedades intrínsecas aos ácidos graxos, uma boa

fonte de óleo para biodiesel deve apresentar também baixos custos de produção e alta produtividade de óleo (Pinto et al., 2005).

Dentre as matérias-primas mais utilizadas na produção mundial de biodiesel destaca-se o óleo de canola (84%) e de girassol (13%) (Thoenes, 2006; Gui et al., 2008). No Brasil, destacam-se o óleo de soja e, em seguida, o de canola e alguns tipos de óleos derivados de frituras (Ramos, 1999; Costa Neto et al., 2000). Existem outras possíveis fontes: palma (dendê), amêndoa, andiroba (*Carapa guianensis*), babaçu (*Orbignia* sp.), cevada, camelina (*Camelina sativa*), coco, copra, cumaru (*Dipteryx odorata*), *Cynara cardunculus*, óleo de peixe, amendoim, pinhão manso (*Jatropha curcas*), karanja (*Pongamia glabra*), louro, aveia, *Lesquerella fendleri*, *Madhuca indica*, microalgas (*Chlorella vulgaris*), pequi (*Caryocar* sp.), nim (*Azadirachta indica*), kamani (*Calophyllum inophyllum*), *Pongamia pinnata*, entre outras como sementes de papoula, arroz, mamona, gergelim, sorgo e trigo (Azam et al., 2005; Pinto et al., 2005). Entre estas, o óleo de palma tem-se destacado, com participação de aproximadamente 1% na produção de biodiesel (Thoenes, 2006; Gui et al., 2008; Lam et al., 2009).

No Brasil, a produção de biodiesel foi adaptada para atender as culturas de cada região. Nas regiões Sudeste e Centro-Oeste destacam-se o algodão, a mamona, o girassol e a soja; no Sul, o óleo de canola; e no Norte e Nordeste, o dendê (Pinto et al., 2005). A mamona e o dendê têm sido alvos de incentivos do Governo Federal, para a produção de biodiesel, tendo aprovado reduções de até 100% dos impostos federais incidentes sobre combustíveis, quando estes forem derivados de matéria-prima fornecida por agricultores familiares (MDIC, 2006).

Independentemente das especificidades regionais, a soja tem-se destacado como a principal matéria-prima para essa finalidade, sendo também responsável por 80% da produção total de biodiesel no país (Pinto et al., 2005). Devido às diferenças nas propriedades do biodiesel de óleo de soja, tem sido sugerido o uso de aditivos para melhorar o seu desempenho (Dunn et al., 2002). Para isso, tem sido sugerida também a mistura de matérias-primas, o que melhora a qualidade do balanço entre os ácidos graxos (Knothe, 2010).

Segundo Pinto et al. (2005), Estados Unidos, Japão e países europeus têm dedicado esforços para se alcançar um bom nível tecnológico na produção do biodiesel. Nestes países a soja também é a fonte mais usada para essa produção, embora isto possa variar dependendo de sua disponibilidade em cada local.

2.5 IMPORTÂNCIA E CARACTERIZAÇÃO DE BANCOS DE GERMOPLASMA

Os recursos genéticos constituem-se na parte essencial da biodiversidade, que é usada pelo homem para a promoção do desenvolvimento sustentável da agricultura e para a produção de alimentos. Esses recursos tratam da variabilidade genética de determinada espécie, a qual pode ser preservada “in situ” ou “ex situ”, incluindo-se a possibilidade de sua manutenção em bancos de germoplasma. O conceito de germoplasma pode ser definido como a estrutura física vegetal, animal ou de micro-organismos, dotada de caracteres hereditários, capaz de gerar um novo indivíduo, transmitindo suas características de geração em geração (Ribeiro, 1995). Dessa forma, o banco de germoplasma consiste em um elemento dos recursos genéticos, que permite a manipulação da variabilidade genética dentro de cada espécie, com fins de utilização para pesquisa, especialmente no melhoramento genético (Goedert, 2007).

Os bancos de germoplasma suprem os programas de melhoramento das espécies de grande importância econômica, como soja, milho, arroz, entre outras. O uso desses recursos como fonte de criação de novas variedades é fator de fundamental importância para a obtenção de produtividades mais elevadas e para as transformações tecnológicas necessárias ao processo de modernização do agronegócio (Goedert, 2007). Com o intuito de preservar a diversidade genética e disponibilizar material para os programas de melhoramento, tem sido constituídas, então, coleções de germoplasma de diversas espécies, as quais são, posteriormente, avaliadas (Anti, 2000; Silva et al., 2001; Silva & Salomão, 2006; Borba, 2007).

Em decorrência do uso desorganizado dos recursos genéticos e da expansão das cidades e da agricultura, a preservação destes recursos tem sido foco de vários países. Em 1972, na Conferência sobre o Homem e o Meio Ambiente, em Estocolmo, na Suécia, Sir Otto Frankel, fundador do movimento sobre recursos genéticos, já chamava a atenção sobre o fato de que a erosão genética estava ameaçando as espécies essenciais, responsáveis pelo suprimento da alimentação mundial (Goedert, 2007). Nesse sentido, a conservação destes recursos em bancos de germoplasma tem sido estratégia imprescindível no processo de minoração dos efeitos da erosão genética (Wetzel & Ferreira, 2007).

O uso dos bancos de germoplasma, sobretudo com eficiência, ainda tem sido limitado (Wetzel & Ferreira, 2007). Entre as principais causas da subutilização do germoplasma cita-se a falta de interesse dos melhoristas, em razão de já possuírem as suas

coleções de trabalho (van Sloten, 1987; Valls, 2007). Outro fator é a escassez de informações a respeito de caracteres de importância para os acessos mantidos nesses bancos, e, às vezes, quando disponíveis, insuficientes ou inadequadas. No Brasil, a falta de melhoristas para uma grande quantidade de culturas ou até mesmo de programas de melhoramento, também tem sido sugerida como motivo de subutilização de alguns bancos de germoplasma (Wetzel & Ferreira, 2007; Valls, 2007).

Várias são as atividades desenvolvidas nesses bancos, demandando recursos financeiros e humanos. A importância da espécie cultivada é um fator que implica na disponibilização de recursos para a manutenção de seus bancos de germoplasma. As atividades de rotina num banco de germoplasma incluem a coleta, intercâmbio, conservação, caracterização, informação e documentação dos acessos (Wetzel & Ferreira, 2007).

A caracterização e avaliação dos bancos de germoplasma, além de gerarem informações sobre cada acesso, permitem mensurar a variabilidade genética e fenotípica e, ainda, verificar se há estruturação genética entre os acessos ali preservados. Analisando-se a magnitude da variabilidade genética, pode-se, por exemplo, constatar a necessidade de se incluir novos acessos no banco. O responsável por estas atividades é o curador do germoplasma; porém, a interação entre todos os especialistas envolvidos é fundamental para o sucesso dessa caracterização. O curador é responsável, ainda, pela organização e divulgação das informações geradas sobre os acessos (Valls, 2007). O autor acrescenta que tais informações permitem, por exemplo, a identificação de possíveis duplicações de acessos; o que assegura maior racionalização dos trabalhos no banco, evitando-se a repetição de atividades e reduzindo-se o tempo e os recursos financeiros dispendidos. Com a caracterização do germoplasma é possível, ainda, o estabelecimento de coleções nucleares; e, a formação destas coleções facilita o uso da diversidade nas avaliações subsequentes.

Segundo Valls (2007), o processo de caracterização e avaliação de acessos em bancos de germoplasma passa por cinco etapas. A primeira e a segunda são, respectivamente, a identificação botânica e a elaboração do cadastro dos acessos. Estas etapas são baseadas no preenchimento dos chamados descritores de passaporte. Já nesse momento pode ocorrer a identificação de possíveis duplicações entre os acessos. Muitas vezes, esses acessos foram reincidentemente introduzidos. A terceira etapa é a avaliação de caracteres de alta herdabilidade, facilmente visíveis ou mensuráveis, chamados descritores.

A quarta etapa consiste numa avaliação agrônômica preliminar, exigindo-se experimentos mais elaborados, haja vista a baixa herdabilidade dos caracteres avaliados. Devido à demanda de maior área e repetições, essa etapa requer maior disponibilidade de recursos financeiros e, ainda, menor número de acessos a ser avaliado. Na etapa final, faz-se uma avaliação aprofundada ou complementar, levantando-se dados de diversos caracteres, em vários locais e anos. Tratam-se de informações mais difíceis de serem obtidas e, portanto, que envolvem maior custo.

O processo de caracterização de um banco de germoplasma deve ser contínuo e envolver, simultaneamente, curadores e melhoristas. A caracterização fornecerá subsídios ao melhorista para a escolha de genitores para os seus cruzamentos. Ou seja, a análise de divergência genética auxilia na classificação de grupos divergentes e, conseqüentemente, na escolha de genitores para cruzamentos com maior probabilidade de sucesso ao melhoramento da espécie; pois genitores mais divergentes produzirão progênies com maior variância genética. Com isso, os esforços serão direcionados para acessos com maior probabilidade de resultar em combinações híbridas ou populações segregantes vantajosas (Queiroz & Lopes, 2007); e, ainda, se estimula o melhorista ao uso do germoplasma.

A mensuração da divergência genética geralmente é baseada em dados morfológicos, fisiológicos, agrônômicos e outros atributos fenotípicos importantes para a espécie (Oliveira et al., 2007). Os caracteres de maior interesse, entretanto, na sua maioria são controlados por muitos genes e apresentam baixa herdabilidade. Assim, os marcadores moleculares têm se despontado como uma potente ferramenta para auxiliar na caracterização de bancos de germoplasma. Com a identificação molecular dos acessos, é possível determinar as frequências gênicas e, então, a variabilidade genética das populações (Ferreira, 2003). Entre esses marcadores, os microssatélites (SSR) têm sido largamente utilizados na caracterização de germoplasma. A maior vantagem dessa técnica é o elevado polimorfismo revelado, especialmente em germoplasma aparentado e de baixa variabilidade, como é o caso da soja, que apresenta estreita base genética (Rongwen et al., 1995; Ferreira, 2003; Priolli et al., 2004).

A estrutura deficiente, seja em termos financeiros ou de recursos humanos, em muitos bancos de germoplasma, tem dificultado a caracterização de seus acessos. Dessa forma, a eficiência desses bancos depende largamente da força dos programas de melhoramento genético (Valls, 2007). Em contrapartida, o uso que os melhoristas fazem desses recursos genéticos, em geral, é pequeno; e isto não se restringe ao Brasil ou a países

em desenvolvimento (Nass et al., 1993). Segundo estes autores, a utilização do germoplasma de soja, por melhoristas no Brasil, apesar de baixa, é superior à de milho. Isso pode ser explicado por um maior nível de avaliação e caracterização dos acessos disponíveis de soja. Acrescentam que as atividades de avaliação e caracterização do germoplasma envolvem alto custo e retorno em longo prazo. A aproximação entre os programas de melhoramento, certamente, aumentará o conhecimento sobre o germoplasma e a extensão da variabilidade genética nas diferentes espécies. Logo, a organização e divulgação das informações sobre o germoplasma é fundamental para o melhor manejo de acessos nos programas de melhoramento (Nass et al., 1993; Valls, 2007).

O uso direto de acessos dos bancos de germoplasma pode não ser tão interessante para o melhorista, pois alguns deles apresentarem características indesejáveis. Assim, a transferência de genes desejáveis, a partir desses acessos, para cultivares adaptadas, seguida da sua inclusão nos programas de melhoramento, pode estimular mais esse uso. Para isso, é necessário o acúmulo de informações sobre os acessos, permitindo a identificação de características úteis em genitores específicos.

Para estimular a utilização da diversidade dos bancos de germoplasma tem sido proposta a construção de coleções nucleares, conceito introduzido por Frankel (1984). Coleção nuclear (*core collection*, em inglês) consiste de um grupo reduzido de acessos que representa a variabilidade genética da coleção de germoplasma, com mínima redundância. Essa estratégia permite priorizar e concentrar as atividades de caracterização e avaliação da coleção, formando uma base de informações mais completa sobre esse conjunto de acessos, podendo, então, estimular o seu uso. A coleção nuclear não é removida do banco de germoplasma, mas é designada como um conjunto de acessos dentro da coleção. Embora os acessos da coleção nuclear possam ter mais informações, isto não significa que o restante terá menor importância ou passará a ser descartável (Cordeiro & Abadie, 2007).

Em síntese, a representação do banco de germoplasma em coleções nucleares torna possível a concentração de esforços na caracterização e avaliação de recursos prioritários, e a maior disponibilização de informações sobre esses acessos poderá levar, efetivamente, à ampliação do uso do germoplasma. Várias coleções nucleares de culturas de importância alimentar têm sido definidas, tais como, arroz (Abadie et al., 2005), feijão (Zeven et al., 1999; Logozzo et al., 2007), milho (Malosetti & Abadie, 2001; Yu et al., 2004; Yao et al., 2008) e soja (Zhao et al., 2005; Wang et al., 2006; Oliveira, 2007; Oliveira et al., 2010). Para o estabelecimento de coleções nucleares, diferentes tipos de

informações e estratégias de amostragem têm sido utilizados (Ortiz et al., 1998; van Hintum et al., 2000; Chandra et al., 2002; Oliveira, 2007; Vasconcelos et al., 2007; Oliveira et al., 2010). Entre essas informações incluem-se os dados de passaporte, de caracteres quantitativos, qualitativos, sejam morfológicos ou agrônômicos, bem como os dados de marcadores moleculares (Rao, 2002).

Em soja, Oliveira et al. (2010) definiram, a partir da subcoleção de germoplasma do Departamento de Agricultura dos Estados Unidos (USDA), uma coleção nuclear com 1.600 acessos. Para constituir a coleção foi utilizada amostragem aleatória estratificada, proporcional ao tamanho dos estratos. Estes estratos foram formados a partir da análise de agrupamento pelo método de Tocher invertido, aplicada em dezoito características quantitativas. O autor orienta que deve haver prioridade na caracterização e avaliação dessa coleção para outros diversos caracteres e, ainda, atenta para a inclusão ou exclusão de acessos à medida que essas informações forem consideradas.

2.6 ANÁLISES DE COLEÇÕES DE GERMOPLASMA

2.6.1 Divergência genética

Em coleções de germoplasma dispõem-se de várias informações, tanto a partir de dados quantitativos, quanto qualitativos. O tratamento de cada informação separadamente dificulta a manipulação dos acessos de bancos de germoplasma. As técnicas de análise multivariada são instrumentos poderosos na análise de dados com múltiplas informações, como aqueles disponíveis nesses bancos (Ferreira, 2008). Assim, os métodos de estatística multivariada podem ser utilizados com o propósito de simplificar ou facilitar a interpretação dessas informações.

Vários métodos estatísticos multivariados podem ser utilizados para dados de bancos de germoplasma, prevalecendo-se as análises de componentes principais, de variáveis canônicas, de agrupamento e análise discriminante (Duarte, 1997). Segundo Cruz & Regazzi (2001), a escolha do método de mensuração da divergência genética depende da precisão desejada pelo pesquisador, da facilidade da análise e da forma como os dados são obtidos. O uso da análise multivariada em estudos sobre divergência genética tem sido assunto de extensas pesquisas em várias espécies vegetais: açaizeiro (Oliveira, 2005), trigo (Bertan et al., 2006), soja (Oliveira, 2007), pimentão (Sudré et al., 2005; Silva, 2008) e *Stylosanthes* (Karia, 2008).

Para dispor de variabilidade genética é necessário que o programa de melhoramento utilize genitores de base genética ampla. Embora exista consenso sobre a necessidade de se ampliar a base genética dos programas de melhoramento de soja, tem-se observado a utilização repetitiva de certos genótipos, com capacidade combinatória aparentemente alta e boas características agronômicas.

Até o ano de 1973, cerca de 80% das cultivares recomendadas para cultivo no Rio Grande do Sul eram resultantes de introdução ou seleção de materiais introduzidos. Nessa época, mais de setenta cultivares de soja nesse Estado eram derivadas de cruzamentos envolvendo as cultivares Hood ou Hill (Bonetti, 1981; Bonetti, 1983). O cenário não é muito diferente nas outras regiões brasileiras. Segundo Priolli et al. (2004), o germoplasma brasileiro de soja manteve a variabilidade genética constante nas últimas três décadas, mesmo sob expansão e melhoramento da cultura.

Hiromoto & Vello (1986) avaliaram a base genética do germoplasma de soja cultivado no Brasil, e compararam o grau de similaridade do germoplasma brasileiro aos Estados Unidos. Observaram que somente 26 ancestrais representavam as variedades de soja cultivadas no país. E ainda, onze destes genitores contribuíram com 89% do conjunto gênico dessas variedades. Então, concluíram que a base genética do germoplasma adaptado às condições brasileiras era muito restrita. Vello et al. (1988) afirmaram que das 74 variedades recomendadas para cultivo no ano agrícola 1983/1984, 15% eram oriundas de introduções de variedades norte-americanas, enquanto as 85% restantes foram obtidas de cruzamentos biparentais, também entre introduções norte-americanas. Gizlice et al. (1993) citam que, nos Estados Unidos, a base genética das variedades de soja recomendadas é constituída de pouco mais de quinze genitores.

Pequeno (2001) avaliou a diversidade genética entre 79 genótipos de soja, incluindo as cultivares recomendadas para o Estado de Goiás, no período de 1996 a 2001. Foram encontrados dezoito ancestrais nas genealogias de 39 cultivares avaliadas, sendo que doze contribuíram com 94% dos genes presentes nessas cultivares, indicando base genética estreita nesse germoplasma.

Mais recentemente, por intermédio de avaliação de genealogias, Wysmierski (2010) identificou sessenta ancestrais relacionados a 444 cultivares brasileiras de soja. O autor concluiu que quatro ancestrais (CNS, S-100, Nanking e Tokyo) contribuem com 55,26% da base genética; e apenas outros quatorze ancestrais tiveram contribuição acima

de 1%. Concluiu ainda que as cultivares analisadas apresentam base genética estreita, similar à dos Estados Unidos, e bem menor que as da China e do Japão.

Estudos de divergência genética são importantes para o conhecimento da variabilidade genética das populações e para o monitoramento dos bancos de germoplasma (Cruz & Carneiro, 2003). Para o direcionamento de genitores na formação de populações segregantes, buscam-se aqueles mais divergentes, pois a divergência genética nestas populações é função das distâncias entre os genitores (Ferreira, 2008). O grau de similaridade ou distância genética entre os acessos de uma coleção de germoplasma pode ser estimado a partir de dados qualitativos, como dados de marcadores moleculares, ou a partir de dados quantitativos. A análise de similaridade visa classificar um conjunto de acessos, avaliados por diferentes descritores ou caracteres, por meio da estimativa da distância entre eles (Oliveira et al., 2007).

As medidas de similaridade têm sido feitas a partir de distâncias ou coeficientes de similaridade. Uma medida de similaridade serve para indicar o grau de semelhança entre os acessos, e uma de dissimilaridade indica a diferença entre os acessos (Oliveira et al., 2007). Existem diversas medidas de dissimilaridade, as quais podem levar a distintos padrões de agrupamento (Cruz & Regazzi, 2001; Mingoti, 2007). A escolha adequada da medida de divergência e do método de agrupamento a serem utilizados depende dos objetivos e da natureza dos dados disponíveis. Neste trabalho são abordadas apenas medidas de similaridade ou dissimilaridade para variáveis quantitativas.

Entre os métodos de mensuração de distância genética, a partir de dados quantitativos, podem-se citar: distância euclidiana, distância de Mahalanobis e distância de Karl Pearson. Nesta última, a menos empregada em estudo de recursos genéticos vegetais, é feita uma generalização da distância com a matriz de variâncias das variáveis. Tal ponderação é indicada para variáveis em escalas diferentes (Mardia et al., 1979; Mingoti, 2007).

A distância euclidiana é uma das mais empregadas, embora apresente o inconveniente de crescer com o aumento do número de variáveis envolvidas no estudo. Nesse sentido, a distância euclidiana média retorna uma medida não mais influenciada por este número, pois envolve uma ponderação pelo número de variáveis (Mingoti, 2007). Esse tipo de distância é mais apropriado para variáveis que possuem escalas similares, pois, caso contrário, as variáveis com maior variabilidade irão dominar a classificação das

distâncias (Ferreira, 2008). Para contornar o problema de escala, tem sido recomendável a padronização dos dados (Cruz & Regazzi, 2001).

A distância euclidiana não leva em consideração as correlações residuais entre os caracteres. Assim, quando há dependência entre as variáveis, a distância de Mahalanobis é mais indicada, embora requeira dados com repetições (Cruz & Regazzi, 2001). Esta medida oferece a vantagem de levar em consideração a existência de correlações entre os caracteres, analisados por meio da matriz de variâncias-covariâncias residuais (Cruz & Regazzi, 2001).

2.6.2 Técnicas de agrupamento

A análise de agrupamento (*cluster analysis*, em inglês) tem como objetivo dividir os elementos de uma amostra ou população em grupos (Milligan & Cooper, 1985; Mingoti, 2007). Os métodos de agrupamento têm por objetivo classificar os elementos (indivíduos ou objetos), por exemplo, os acessos de um banco de germoplasma, de acordo com suas similaridades estimadas previamente. Os elementos semelhantes são alocados num mesmo grupo e, por conseguinte, elementos que pertencem a grupos diferentes são considerados dissimilares. O agrupamento busca homogeneidade entre os elementos dentro do grupo e dissimilaridade entre os grupos (Ferreira, 2008).

A análise de agrupamento é uma técnica distinta da análise discriminante. Essa última tem o objetivo de classificar elementos em grupos pré-existent; ou seja, neste caso já se conhecem os grupos e deseja-se designar uma nova informação (elementos) a um desses grupos. A análise de agrupamento, diferentemente, é uma técnica mais simples, em que não há pressuposições a respeito do número de grupos ou da sua estrutura (Johnson & Wichern, 1992).

O processo de agrupamento envolve duas etapas. A primeira é obter a matriz de distâncias ou de divergência entre os elementos (ex. genitores, acessos de germoplasma etc.). Para isso, no caso de dados quantitativos, a distância euclidiana ou a de Mahalanobis são as mais adotadas em melhoramento de plantas. A segunda etapa é a adoção de uma técnica de agrupamento. As técnicas de agrupamento são frequentemente classificadas em dois tipos: métodos hierárquicos; e métodos não-hierárquicos ou de otimização (Duarte, 1997; Cruz & Regazzi, 2001; Mingoti, 2007). Os métodos hierárquicos são classificados em: divisivos, nos quais se fazem divisões até obter os grupos; ou aglomerativos, em que

são feitas fusões até a determinação dos grupos. Os métodos hierárquicos são utilizados em análises exploratórias de dados com o intuito de se identificar possíveis agrupamentos e o provável valor do número de grupos. Já para o uso das técnicas não-hierárquicas é necessário que o número de grupos seja pré-estabelecido pelo pesquisador (Mingoti, 2007).

Os métodos hierárquicos aglomerativos partem do princípio de que no início do processo de agrupamento têm-se n conglomerados; ou seja, cada elemento do conjunto de dados observado é considerado como sendo um conglomerado isolado. Em cada passo do algoritmo, os elementos amostrais vão sendo agrupados, formando novos conglomerados, até o momento no qual todos os elementos estarão num único grupo. Em cada estágio do procedimento de agrupamento, os grupos vão sendo comparados através de uma medida de similaridade (ou dissimilaridade) previamente definida, e, então, aqueles mais similares são combinados para constituir um único conglomerado. Em cada estágio do algoritmo, cada novo conglomerado é formado a partir do agrupamento dos conglomerados anteriores. No método hierárquico, uma vez estabelecida a ligação entre dois elementos, esta não poderá mais ser rompida (Mardia et al., 1979; Mingoti, 2007).

Nos métodos hierárquicos podem ser aplicados os critérios de ligação simples, ligação completa, ligação média, centróide, Ward, entre outros (Mingoti, 2007). Segundo esta autora, os métodos de ligação simples, completa ou média podem ser utilizados tanto para variáveis quantitativas quanto qualitativas. Já os métodos do centróide e de Ward são apropriados apenas para variáveis quantitativas, já que tem como base a comparação de vetores de médias.

No método hierárquico de ligação simples (*Single Linkage Method*) ou do vizinho mais próximo, o critério de agrupamento se baseia na distância mínima entre elementos, isto é, a similaridade entre dois conglomerados é definida pelos dois elementos mais parecidos entre si (Mardia et al., 1979). No método hierárquico de ligação completa (*Complete Linkage Method*) ou método do vizinho mais distante, a similaridade entre um conglomerado e um elemento é definida pelos elementos que são mais distantes entre si (Duarte, 1997; Mingoti, 2007). O método hierárquico de ligação simples mostra-se incapaz de delinear grupos poucos separados. Já o método de ligação completa tende a isolar valores discrepantes nos primeiros passos do agrupamento (Mingoti, 2007).

No método hierárquico da ligação média (*Average Linkage*), a distância entre os conglomerados é dada pela média das distâncias entre todos os pares de elementos, envolvendo os dois conglomerados sob comparação. O agrupamento baseado na distância

média entre pares de observações tende a produzir grupos com pequenas variâncias e, apresenta uma tendência fraca de produzir grupos com mesma variância (SAS Institute, 2010). Esse procedimento divide-se em método da ligação média não-ponderada, conhecido como UPGMA (*unweighted pair group method with arithmetical*), e método da ligação média ponderada, conhecido como WPGMA (*weighted pair-group method using arithmetical average*). Este último foi desenvolvido, independentemente, por Sokal & Michener (1958) e McQuitty (1966). Nos programas computacionais, o método é também chamado de análise de similaridade de McQuitty, ou simplesmente método de McQuitty. A diferença entre UPGMA e WPGMA é que este último usa como peso o tamanho do respectivo grupo. Então, WPGMA tem sido recomendado quando os tamanhos dos grupos são irregulares (Mocellin & Rossi, 2006).

No método hierárquico do centróide a distância entre os grupos é definida como sendo a distância entre os vetores de médias, também chamados de centróides dos grupos que estão sendo comparados. A cada agrupamento um novo centróide é calculado e, para isso, a cada etapa é necessário voltar-se aos dados originais para o cálculo de nova matriz de distâncias, exigindo maior esforço computacional (Mingoti, 2007).

O método hierárquico de Ward, ou da mínima variância, fundamenta-se nas variâncias dentro e entre os grupos. Esse método apresenta semelhança com o método do centróide, no entanto, tende a produzir grupos com aproximadamente o mesmo número de elementos; isto é, conglomerados com tamanhos diferentes sofrem penalização. O método de Ward tem como base principal os princípios da análise de variância, embora não dependa do fato de os dados serem provenientes de população com distribuição normal (Mingoti, 2007).

Um procedimento comum para avaliar a eficiência das técnicas de agrupamento hierárquico é comparar a matriz de similaridade (ou dissimilaridade) original, com a matriz de distâncias resultante do agrupamento (gerada pelo dendrograma). A correlação entre estas duas matrizes é amplamente utilizada em estudos de taxonomia numérica, sendo conhecida como correlação cofenética (Sokal & Rolf, 1962; Holgersson, 1978; Aldenderfer & Blashfield, 1984; Romesburg, 2004). Esta medida é, portanto, um critério para verificar o quão bem o agrupamento realizado representa a matriz de similaridade (ou dissimilaridade) original. Correlações cofenéticas superiores a 0,75 são consideradas altas (McGarigal et al., 2000).

O método UPGMA tem sido citado como o mais eficiente na maximização da correlação cofenética. Isso porque esse método produz menores distorções na transformação da matriz de distâncias em dendrograma e, conseqüentemente, é o que melhor representa os dados originais (Farris, 1969; Romesburg, 2004). Moura et al. (2010) obtiveram agrupamentos pelos métodos UPGMA, Ward e vizinho mais próximo, tendo obtido os seguintes valores de correlação cofenética: 0,82, 0,59 e 0,62, respectivamente. Sesli & Yegenoglu (2010) compararam os métodos de agrupamento UPGMA, WPGMA, ligação simples e ligação completa. A maior correlação cofenética foi obtida a partir do agrupamento pelo método UPGMA, e a menor, pelo método da ligação completa. Em razão disso, o método UPGMA tem sido amplamente utilizado nos estudos sobre divergência em germoplasma de diversas culturas: arroz (Buso et al., 1998), soja (Li & Nelson, 2001), figo (Chatti et al., 2007), *Stylosanthes* (Karia, 2008), pimenta (Moura et al., 2010) e oliveira (Sesli & Yegenoglu, 2010).

Como na maioria dos métodos de agrupamento, as fontes de erro e de variação controlada não são formalmente consideradas, estes métodos tornam-se sensíveis à presença de valores discrepantes (*outliers*). Ademais, nos métodos hierárquicos, os objetos que forem possivelmente alocados de forma incorreta em estágios preliminares não serão realocados corretamente. Conseqüentemente, o agrupamento final deve ser examinado cuidadosamente, quanto à tal sensibilidade. Por isso, é aconselhável testar vários métodos de agrupamento para verificar se os resultados apresentam consistência. A estabilidade do agrupamento hierárquico pode ser verificada aplicando-se novo agrupamento aos dados, após alguma perturbação acrescentada a eles. Se os grupos são corretamente bem distinguidos, os agrupamentos antes e após a perturbação apresentarão concordância (Johnson & Wichern, 1992).

Uma forma de visualizar o agrupamento obtido é através de um dendrograma. O dendrograma é um diagrama com dois eixos em forma de árvore, em que um deles representa os elementos (ou objetos), e o outro, as distâncias (entre elementos ou grupos). A ramificação da árvore representa a organização de cada agrupamento (Mardia et al., 1979). Para a obtenção das matrizes de distâncias e posteriormente o agrupamento, pode-se contar com o auxílio de *softwares* como o SAS, NTSYS, Genes, R, entre tantos outros. Os dois últimos apresentam a vantagem de serem *softwares* livres.

Uma vez realizado o agrupamento, a questão seguinte é decidir sobre o número de grupos que melhor caracteriza o conjunto analisado. Tal como o tipo de medida de

similaridade e o critério de adequação da partição, a definição do número ideal de grupos é um questionamento difícil a ser resolvido pelo pesquisador. Segundo Cruz & Regazzi (2001), o estabelecimento desse número ainda é feito, normalmente, de forma subjetiva.

A definição do número de grupos a serem estabelecidos na análise, em termos práticos, representa a decisão acerca do momento de parada do algoritmo de agrupamento. E esta decisão, se tomada de maneira equivocada, pode levar a um de dois tipos de erros. O primeiro ocorre quando o número de grupos formado, após a parada do algoritmo, for maior do que realmente é. E o segundo, é quando ocorre a formação de menos grupos do que os dados originais apresentam. Embora a severidade de ambos os erros pode variar dependendo do contexto, segundo Milligan & Cooper (1985), o último deles pode ser considerado mais grave, devido à união de grupos distintos e, conseqüentemente, a perda de informação.

Alguns critérios para auxílio na determinação do número de grupos são discutidos por Mingoti (2007). Um destes é a análise do comportamento dos níveis de fusão. Por este critério se compara a distância entre os grupos (nível de distância), em cada passo do agrupamento, através de um gráfico cartesiano (X,Y), "passo" *versus* "nível de distância". Por esse gráfico avalia-se o comportamento da distância em cada estágio do agrupamento, permitindo visualizar e identificar passos sucessivos de maior distanciamento ou "salto". Esse ponto indica o momento ideal de parada do algoritmo, pois sugere que os grupos começaram a perder a similaridade. Outro critério similar é a análise do comportamento do nível de similaridade. Neste, observa-se o nível de similaridade entre os grupos, procurando detectar pontos que apresentam decréscimo acentuado na similaridade dos conglomerados unidos.

Outro critério baseia-se no coeficiente de determinação (R^2), resultante da análise das somas de quadrados entre e dentro de grupos, passível de estimação a cada passo do algoritmo. Este coeficiente representa a relação entre a soma de quadrados total entre grupos e a soma de quadrados total corrigida para a média global. Numa determinada análise, quanto maior o número de grupos, maior será a soma de quadrados entre os grupos (e menor aquela dentro de grupos), implicando em maiores valores de R^2 . Assim, deve-se fazer um gráfico (X,Y), passos do agrupamento *versus* valores de R^2 , para se visualizar algum ponto de "salto" relativamente grande em relação aos demais, que reflete aumento substancial da variabilidade interna nos grupos, devido ao agrupamento de elementos heterogêneos. O ideal seria conciliar este critério com os dois citados anteriormente;

embora não haja garantias de concordância entre eles. Mingoti (2007) acrescenta que a melhor estratégia seria determinar uma região de valores plausíveis para o número de grupos, baseado em um dos critérios anteriores, e refinar a escolha nessa região com base no critério R^2 .

A estatística Pseudo F, sugerida por Calinski & Harabasz (1974), também representa um critério adicional para se determinar o ponto de parada num algoritmo de agrupamento. Este critério é análogo à estatística F de uma análise de variância univariada, e, segundo seus propositores, já havia sido usado por Edwards & Cavalli-Sforza (1965), similarmente ao teste F, numa análise multivariada. Por este critério, a melhor partição dos dados corresponderá ao número de conglomerados que resultar no maior valor de Pseudo F. Este valor está relacionado à significância do teste de médias populacionais dos grupos, resultando na partição com a maior heterogeneidade entre grupos (Mingoti, 2007).

Outro critério com mesma finalidade baseia-se na estatística Pseudo T^2 , proposta por Duda & Hart (1973), que se assemelha ao Pseudo F; pois também está relacionado ao uso de um teste de hipótese. Aplica-se como se em cada passo do algoritmo estivesse sendo feito um teste para comparar os vetores de médias dos dois grupos que se uniram para formar um novo grupo. Assim, em cada passo é calculado um valor de Pseudo T^2 , buscando-se também o maior valor da estatística. Este é o valor que maximiza a probabilidade de se rejeitar a hipótese de igualdade entre os vetores de médias; e, então, os dois conglomerados não deveriam ser unidos para formar um único grupo. Assim, indica a interrupção do algoritmo no passo anterior a esse agrupamento (Mingoti, 2007).

Milligan & Cooper (1985) compararam, via simulação, trinta métodos para estimar o número de grupos em análise de conglomeração. Os três critérios que se destacaram foram Pseudo F, Pseudo T^2 e outro critério denominado CCC (*Cubic Clustering Criterion*), que é disponível no aplicativo computacional SAS. Segundo Mingoti (2007), o critério CCC, cuja partição final está relacionada a valores da estatística CCC maiores que 3,0, não é usualmente abordado na literatura de estatística multivariada. Mufti et al. (2005) avaliaram a estabilidade de três critérios de agrupamento, em um dos dois conjuntos de dados utilizados, o Pseudo F mostrou-se o mais adequado.

Por fim, outro critério utilizado para determinação do número de grupos é o chamado coeficiente de correlação semiparcial (SPRSQ). Nesse critério, valores pequenos indicam que os dois conglomerados em análise podem ser considerados um só; enquanto valores altos indicam que os dois grupos unidos são provavelmente diferentes.

Semelhantemente ao Pseudo T^2 , a correlação semiparcial observa a distância entre os elementos de dois conglomerados unidos num determinado passo do algoritmo. Para observar a variação da correlação semiparcial, como em outros métodos, elabora-se um gráfico (X,Y) passos *versus* valor observado do coeficiente e verifica-se a ocorrência de saltos na curva obtida (Mingoti, 2007).

2.6.3 Análises de componentes principais e de variáveis canônicas

A análise de componentes principais (ACP) é uma técnica estatística multivariada que pode ser utilizada para compressão de dados, redução da dimensionalidade do conjunto de dados multivariados, extração de características e projeção de dados multivariados. Através da ACP é possível obter uma combinação linear de um conjunto de variáveis, que retenha o máximo possível da informação (variabilidade) contida nas variáveis originais. A idéia básica é buscar variáveis latentes que representem combinações lineares do grupo de variáveis em estudo (Ferreira, 2008).

A análise de variáveis canônicas (AVC) é um procedimento alternativo à ACP, que apresenta como diferença única, o uso adicional da matriz de covariância residual para a obtenção das combinações lineares das variáveis originais (Cruz & Regazzi, 2001). Com isso, agrega vantagens comparativas (tratadas mais adiante), embora possua os mesmos princípios teóricos da análise de componentes principais. Logo, o desenvolvimento que se segue, embora esteja dirigido para ACP, é igualmente válido para AVC.

O objetivo principal da ACP é explicar a estrutura de variância e covariância do matriz composta pelas variáveis aleatórias, através da construção de combinações lineares das variáveis originais. Esse novo conjunto é chamado de componentes principais. A informação contida nas variáveis originais é substituída pela informação contida nas componentes principais, as quais são não correlacionadas entre si (Mingoti, 2007). Os componentes principais são extraídos na ordem da mais explicativa para a menos explicativa. Teoricamente, o número deles é sempre igual ao número de variáveis, entretanto, alguns poucos componentes podem ser responsáveis por grande parte da explicação total (Ferreira, 2008).

Quando a distribuição de probabilidades do vetor aleatório em estudo é normal, os componentes principais, além de não correlacionadas, são independentes e têm

distribuição normal. Entretanto, a suposição de normalidade não é pré-requisito para a aplicação da técnica de análise em componentes principais (Mingoti, 2007).

Uma vez determinados as componentes principais, os seus valores numéricos, denominados escores, podem ser calculados para cada elemento amostral. Quando o objetivo é a redução da dimensionalidade do espaço amostral, é necessário, então, estabelecer critérios de escolha do número de componentes principais que serão mantidos no sistema. A utilidade prática da análise decresce com o aumento do número de componentes principais mantidas no sistema (Mingoti, 2007).

Um critério para determinação do número de componentes principais é a representatividade das componentes mantidas, em relação à variância total (Mingoti, 2007). Se for necessário um número muito grande de componentes principais para explicar proporção mínima da variação total, em muitas aplicações, não haverá vantagem de se utilizar a ACP. Isso porque a representação gráfica será de melhor qualidade num espaço de até três dimensões; isto é, se a maior parte da variação total estiver representada em até três primeiros componentes principais. A recomendação geral para que se tenha boa visualização gráfica numa ACP, sem grandes distorções, é utilizar o número de componentes principais que contemplem pelo menos 70% da variação total. Com um pouco de mais rigor, o número dessas componentes será aquele que contempla 80% da variação total (Cruz & Regazzi, 2001; Ferreira, 2008).

A idéia geométrica por trás dessa análise é realizar uma rotação rígida no sistema de eixos coordenados, fazendo com que os novos eixos resultantes sejam posicionados no sentido da maior variabilidade entre os indivíduos. Pode-se, então, reduzir as dimensões originais em um sistema de coordenadas, sem que se tenha perdido muita informação do sistema original. Assim, o emprego da ACP permite, por exemplo, estudos sobre divergência genética entre acessos de bancos de germoplasma. Isto é possível através da análise da dissimilaridade dos genótipos, em gráficos de dispersão elaborados com as primeiras componentes principais (Ferreira, 2008).

A obtenção das componentes principais envolve a decomposição da matriz de covariâncias. A ordem de grandeza das variáveis originais influencia o resultado da análise. Assim, quando existe discrepância muito grande entre as variâncias das variáveis, cada componente passa a ser extremamente dominado por uma variável em particular. Essa discrepância pode, muitas vezes, ser causada pela diferença das unidades de medida das variáveis. Por isso, recomenda-se padronizar as variáveis pelas respectivas médias e

desvios padrão, gerando-se novas variáveis centradas em zero e com variâncias iguais a um. Nesse caso, a ACP aplicada aos dados padronizados (standardizados) é equivalente à obtenção dos componentes principais a partir da matriz de correlações das variáveis originais. É importante ressaltar ainda que os coeficientes dos componentes principais, obtidos por meio da decomposição da matriz de correlações, não são iguais aos coeficientes obtidos através da decomposição da matriz de covariâncias (Mingoti, 2007).

A principal vantagem de se utilizar a matriz de correlação para definir os componentes principais numa ACP é que os resultados de diferentes conjuntos de variáveis aleatórias são mais diretamente comparáveis do que as análises com base na matriz de covariância. As componentes principais têm as mesmas implicações para comparação entre diferentes matrizes de correlação de mesma dimensão, mas não para diferentes matrizes de covariância. Também, os padrões dos coeficientes das componentes principais podem ser facilmente comparados para diferentes matrizes de correlação, permitindo verificar se as duas matrizes de correlação estão gerando resultados similares. A vantagem geral de se usar a matriz de covariâncias ao invés da matriz de correlação é que a inferência estatística sobre as componentes principais da população, com base nas componentes da amostra, é mais fácil. Isso é relevante quando as componentes principais são usadas num contexto em que a inferência estatística é importante. No entanto, na prática é mais comum utilizarem-se as componentes principais como estatística descritiva, o que torna a vantagem do uso da matriz de covariância irrelevante (Jolliffe, 2002).

Essa análise tem sido muito utilizada na área de melhoramento de plantas, devido ao grande volume de dados disponíveis acerca de grupos de genótipos sob avaliação. Além de possibilitar o estudo de diversidade genética, essa técnica permite inferir sobre a importância de cada variável estudada sobre a variação total disponível. O interesse nesse tipo de avaliação reside na possibilidade de se descartarem caracteres que pouco contribuem para a discriminação do material avaliado (Cruz & Regazzi, 2001). Assim, a análise em componentes principais tem sido aplicada também na seleção de descritores, em coleções de germoplasma de várias espécies como: *Stylosanthes* (Karia, 2008), açaizeiro (Oliveira, 2005), cajueiro (Barros, 1991), cacaueteiro (Bekele, et al., 2006) e *Capsicum* (Silva, 2008).

Assim como a ACP, a análise de variáveis canônicas (AVC) também permite avaliar o grau de similaridade entre elementos em gráficos de dispersão. Enfim, todos os resultados obtidos numa AVC podem ser interpretados à semelhança de uma análise em

componentes principais. A análise de variáveis canônicas, entretanto, tem a vantagem de manter o princípio do processo de agrupamento com base na distância de Mahalanobis, pois leva em conta as correlações residuais existentes entre as médias dos genótipos. Para estimação das variáveis canônicas é necessário, então, dispor também da matriz de correlação residual. Em muitas condições experimentais, como aquelas inerentes à avaliação de germoplasma, a matriz de correlação residual pode não estar disponível (Cruz & Regazzi, 2001).

A técnica de variáveis canônicas tem sido amplamente difundida no Brasil, por meio do aplicativo computacional Genes (Cruz, 1998). À semelhança da ACP, neste procedimento também se obtêm as combinações lineares das variáveis originais, agora chamadas de variáveis canônicas. Essa análise fornece também a proporção da variância total que cada variável canônica contempla. Na sequência, os escores de cada elemento podem ser usados em análises gráficas sobre as primeiras variáveis canônicas, buscando explicar grande parte da variação dos dados (Cruz, 2006).

Aplicações da AVC em estudos de divergência entre elementos ou populações são reportadas em culturas como: mandioca (Vidigal et al., 1997), feijão (Fonseca & Silva, 1999; Bonett et al., 2006; Moreira et al., 2009), coqueiro-gigante-do-brasil (Ribeiro et al., 1999), aveia (Benin et al., 2003), pimenta e pimentão (Sudré et al., 2005), maracujá (Negreiros et al., 2008), tomate (Marim et al., 2009), espinheira-santa (Mariot & Barbieri, 2010) e mamona (Neto et al., 2010). Além disso, tal como a ACP, a técnica tem sido utilizada também na identificação das variáveis mais importantes num conjunto de descritores, os chamados “descritores mínimos”, com consequente descarte daquelas menos relevantes ou redundantes para a caracterização do germoplasma (Fonseca & Silva, 1999; Ribeiro et al., 1999).

3 MATERIAL E MÉTODOS

3.1 DELINEAMENTO EXPERIMENTAL

Nesta pesquisa foi avaliada uma subcoleção do banco germoplasma (BAG) de soja mantido pela Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), com apoio do Centro Tecnológico para Pesquisas Agropecuárias (CTPA). A coleção completa tem cerca de 2.800 entradas, incluindo introduções, cultivares comerciais e antigas, além de linhagens experimentais derivadas dos ciclos avançados de seleção, conduzidos no âmbito do programa de melhoramento vinculado às referidas instituições. Foi feita amostragem estratificada por julgamento (Gressler, 2004; Malhotra, 2004), cobrindo-se uniformemente os três ciclos de maturação (precoce, médio e tardio) e evitando-se acessos redundantes em termos da genealogia disponível (acessos provenientes de mesmos cruzamentos ou com genitores em comum). Também não foram incluídas linhagens transgênicas, para se garantir manejo uniforme do experimento. O tamanho da subcoleção foi de 527 acessos (quase 20% da coleção), com 177 de ciclo precoce, 175 de ciclo médio e 175 de ciclo tardio, o que se assumiu como amostra representativa do germoplasma preservado no BAG. A relação dos acessos avaliados, com as respectivas informações de ciclo e genealogia, encontra-se no Apêndice A.

A avaliação em campo foi conduzida em novembro de 2007, na Unidade Experimental da Seagro-GO (16°43'10'' de latitude sul, 49°07'05'' de longitude oeste e 765 m de altitude), no município de Senador Canedo-GO. O delineamento experimental adotado foi em blocos incompletos com testemunhas comuns (blocos aumentados), sendo que, para facilitar o manejo das plantas, os blocos foram reunidos em grupos (superblocos) de tratamentos, conforme o ciclo dos acessos. A área foi dividida em trinta blocos com vinte parcelas cada, com duas testemunhas por bloco. Cada parcela foi constituída de uma fileira de plantas, de 5,0 m de comprimento (5,0 m x 0,5 m). Como testemunhas foram escolhidas duas cultivares: Conquista, de ciclo médio, e Monsoy 6101, de ciclo precoce. Ambas apresentavam grande disponibilidade de sementes com altas taxas de germinação. No plantio, feito no espaçamento de 0,5 m entre fileiras, primeiramente as testemunhas

foram alocadas aleatoriamente em parcelas de cada bloco, com uma repetição por bloco; e, na sequência, os acessos foram distribuídos, também por sorteio, nas dezoito parcelas remanescentes de cada bloco.

Na colheita, as plantas foram trilhadas separadamente por parcela, e suas sementes, ensacadas, identificadas e enviadas ao laboratório para as análises físico-químicas do grão. As análises foram realizadas no Laboratório de Genética e Melhoramento de Plantas, da Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás. Apenas sementes íntegras, isto é, livres da presença de insetos, danos mecânicos ou podridões, foram aproveitadas. Estas foram moídas em moinho tipo Wiley, peneira Mesh 20/30, crivo circular. O pó recolhido (farinha) foi armazenado em recipiente plástico, o qual foi devidamente tampado e acondicionado em “freezer”, até a execução das análises laboratoriais.

Os caracteres avaliados foram: teores de óleo e de proteína, expressos em g/100 g de grãos, e teores de ácidos graxos (palmítico, esteárico, oleico, linoleico e linolênico), expressos em g/100 g de óleo. Ambas as unidades podem ser interpretadas como porcentagem. Foram determinados, ainda, os teores de ácidos graxos saturados (soma dos teores dos ácidos palmítico e esteárico) e de ácidos graxos poli-insaturados (soma dos teores dos ácidos linoleico e linolênico). Estes teores permitem inferir sobre a estabilidade e a qualidade do óleo de soja.

3.2 ANÁLISES FÍSICO-QUÍMICAS

3.2.1 Análise do teor de proteína

Na determinação do teor de proteína total no grão, o método empregado foi o de Kjeldahl, que se baseia na determinação do nitrogênio total, após digestão completa da amostra em ácido sulfúrico (Instituto Adolfo Lutz, 2008). Para isto foi utilizada placa digestora e destilador micro-Kjeldahl. Os reagentes utilizados foram: ácido sulfúrico PA; hidróxido de sódio 40% (p/v); solução de ácido sulfúrico 0,1N; indicador de Tashiro (vermelho de metila + azul de metileno); solução de ácido bórico 4%; e mistura catalítica de sulfato de sódio e sulfato de cobre, 10:1. Foram pesados, diretamente em tubos de micro-Kjeldahl, aproximadamente 0,25 g de farinha, tomados em cada amostra previamente seca. A esta massa foi adicionado 1,0 g da mistura catalítica, juntamente com

5,0 mL de ácido sulfúrico concentrado e, então acoplado ao aparelho digestor à temperatura inicial de 40°C. Este foi aquecido gradativamente até atingir a temperatura de 400°C, a qual foi mantida até o aparecimento de solução límpida de coloração esverdeada, que indica o ponto de completa digestão da matéria orgânica. Após esse processo, os tubos foram retirados do aparelho digestor e resfriados. Fez-se a lavagem da parede dos tubos com algumas gotas de água destilada. Se ainda houvesse matéria orgânica na parede do tubo, a amostra era submetida à temperatura de 400°C, novamente, até a completa digestão.

Para a destilação do nitrogênio, foi utilizado um aparelho de Kjeldahl analógico. O produto da destilação foi recolhido em erlenmeyer contendo 20 mL de solução de ácido bórico 4%, com cinco gotas do Indicador de Tashiro. Inicialmente, a solução apresentava coloração avermelhada, indicando acidez, e após o recebimento da solução destilada tornava-se esverdeada, indicando a presença de nitrogênio em forma de amônia. Após o recolhimento de 50 mL do destilado, a solução foi titulada com ácido sulfúrico 0,1N. Para a obtenção do teor de proteína total, foi utilizado o seguinte cálculo:

$$\% \text{ proteína} = \frac{(\text{Volume titulação} \times 0,1 \times F_c \times 0,014) \times 6,25 \times 100}{\text{Peso da amostra}}$$

em que: o volume da titulação, em mL, foi determinado pela quantidade de ácido sulfúrico 0,1N necessária para se atingir o ponto de equivalência, subtraída do volume da titulação da prova em branco; 0,1 x Fc é a normalidade real do ácido sulfúrico usado na titulação, multiplicada pelo fator de correção do ácido para a normalidade real; 0,014 é o equivalente-grama de nitrogênio por mL de titulação; 6,25 refere-se ao fator de transformação de nitrogênio em proteína (conforme Resolução RDC nº 360, de 23 de dezembro de 2003) e o valor 100 é usado para se expressar o conteúdo de proteína em porcentagem (g/100g).

3.2.2 Análise do teor de óleo

A técnica de extração de óleo utilizada foi do tipo Soxhlet indicada no método oficial da *Association of Official Analytical Chemists* (Deutsch, 1995), adaptada por Oliveira (2003). Esse método envolve a extração dos lipídios por meio de solventes. O solvente empregado foi o éter de petróleo.

Pesaram-se 2,5 g de farinha de soja, em cartucho de papel celulose qualitativo (Oliveira, 2003). O cartucho foi acoplado ao aparelho (Soxhlet), suspenso por uma haste.

Adicionaram-se 100 mL de éter de petróleo ao frasco de extração, previamente seco e pesado. Logo depois o frasco foi conectado ao extrator, sob temperatura de 120°C. A extração do óleo, em cada amostra procedeu-se por três horas e trinta minutos. Após esse período, o cartucho foi suspenso para que se recuperasse parte do solvente. O frasco de extração foi retirado do extrator e levado à estufa a 70°C, até a completa evaporação do solvente. Em seguida o frasco foi resfriado em dessecador, até a obtenção de peso constante. O peso do óleo foi obtido pela diferença, em gramas, entre os pesos do frasco antes e depois da extração. Assim, o teor de óleo, em g/100 g, foi obtido por:

$$\% \text{ óleo} = \frac{(P_2 - P_1)}{P_a}$$

em que: P_1 e P_2 são os pesos, em gramas, do frasco seco e resfriado, antes e depois da extração; e P_a é o peso inicial da amostra, em gramas.

3.2.3 Análise dos teores de ácidos graxos

Da fração lipídica extraída na determinação do extrato etéreo, fez-se a transformação do óleo em ésteres metílicos, de acordo com as normas do Instituto Adolfo Lutz (2008). Esse procedimento analítico é conhecido como FAME (*Fatty Acid Methyl Esters*). Nele os ésteres metílicos de ácidos graxos são separados, identificados e quantificados por cromatografia em fase gasosa. Foram quantificados os seguintes ácidos graxos: palmítico (16:0), esteárico (18:0), oleico (18:1), linoleico (18:2) e linolênico (18:3). A opção por estes ácidos veio do fato de serem os de maiores concentrações no óleo de soja.

Para o processo de esterificação pesaram-se, num tubo com tampa em rosca, 100 mg de óleo, aos quais foram adicionados 5,0 mL de solução de hidróxido de sódio (NaOH) em metanol (MeOH) (12,0 g NaOH, em 600 mL de MeOH). A mistura foi agitada vigorosamente por um minuto, em aparelho tipo Vortex. Os tubos foram levados ao banho-maria à temperatura de 80°C, por vinte minutos, sendo posteriormente resfriados. Na sequência, adicionaram-se 5,0 mL de solução de Cloreto de amônio (NH₄Cl) em metanol acidificado. Os tubos, tampados, foram então agitados e levados novamente ao banho-maria à temperatura de 80°C, por cinco minutos. Foram então resfriados, tendo sido adicionados 5 mL de hexano (grau cromatógrafo). Na sequência foram agitados por três minutos, para a recuperação dos ésteres metílicos. Após a separação do conteúdo em duas

fases, a fase superior foi recolhida e colocada em microtubo de 5,0 mL, para quantificação em cromatógrafo gasoso.

As condições das análises em cromatógrafo gasoso (Focus GC Master) foram: temperatura de forno, 250°C; temperatura do injetor, 200°C; e temperatura do detector de ionização em chama, 250°C. Utilizou-se capilar com dimensões de 25 m x 0,32 mm x 0,25 µm, com fase 70% Cyanopropyl polysilphenylene-siloxane. O fluxo dos gases foi de 30 mL por minuto, seja para nitrogênio (gás de arraste), hidrogênio ou para o ar sintético (mistura de oxigênio e nitrogênio). O tempo de corrida de cada amostra foi de dez minutos.

Os cromatogramas (exemplo no Apêndice B) foram gerados a partir de injeção manual de 1,0 µL da solução esterificada, no cromatógrafo, com microseringa agulha fixa de 10 µL. Os teores de cada ácido graxo foram obtidos após a sua detecção no ionizador de chamas, a partir de cromatogramas, tendo sido expressos também em porcentagem (neste caso, em g/100g de óleo). A quantificação baseou-se nos padrões de ésteres metílicos de ácidos graxos, para efeito de comparação com as amostras em análise.

3.3 ANÁLISES ESTATÍSTICAS

3.3.1 Estimação das médias genotípicas dos acessos

Na análise estatística dos dados, inicialmente foram estimadas as médias dos 527 acessos, para cada variável ou caráter. Estas resultaram do ajustamento do seguinte modelo estatístico, assumido como aleatório:

$$Y_{ijk} = m + c_k + b_{j(k)} + g_i + e_{ijk} \quad (1)$$

em que:

Y_{ijk} é o valor observado numa parcela que recebeu o acesso ou genótipo i , no bloco j de um grupo k (precoce, médio ou tardio);

m é a média geral do experimento;

c_k é o efeito do k -ésimo grupo;

$b_{j(k)}$ é o efeito do j -ésimo bloco dentro do k -ésimo grupo;

g_i é o efeito do genótipo i , incluindo-se as cultivares testemunhas; e

e_{ijk} é o erro experimental associado à ijk -ésima observação.

Logo, excetuando-se “ m ”, todos os demais efeitos foram assumidos como aleatórios, independentes, homocedásticos e com distribuição normal.

Os componentes de variância foram estimados pelo método da máxima verossimilhança restrita (REML). Os valores genotípicos associados aos acessos foram preditos pela abordagem de modelos mistos, via estimação REML, método de Quadrados Mínimos Generalizados (GLS) e “melhor predição linear não-viciada” (BLUP). Mais apropriadamente utilizou-se o chamado EBLUP (*Empirical Best Linear Unbiased Predictor*), haja vista que os componentes de variâncias no modelo não eram conhecidos, mas sim estimados dos próprios dados. Por fim, a média genotípica ajustada, para cada variável e acesso, foi obtida somando-se a estimativa da média geral à respectiva predição do valor genotípico do acesso. Para tudo isso, foi utilizado o aplicativo computacional SAS – Statistical Analysis System (SAS Institute, 2010), em seu procedimento Proc Mixed, conforme preconiza Duarte (2000).

Após a obtenção das médias genotípicas, foi identificado um conjunto de acessos com melhor perfil de estabilidade oxidativa no óleo. Para isso, selecionaram-se aqueles com índices de ácidos graxos poli-insaturados abaixo da média de cultivares comerciais, 59,9% (Sanibal & Mancini Filho, 2004), e com os maiores índices de ácido oleico (monoinsaturado). O valor mínimo para este ácido (30%) foi estabelecido com base nos menores índices encontrados entre linhagens, descritas na literatura como menos susceptíveis à oxidação durante o processo de refino, armazenamento e fritura; isto é, como potenciais fontes doadoras desta característica (Rahman et al., 1994; Rahman et al., 1996; Burton et al., 2006; Rahman et al., 2001).

3.3.2 Análises de componentes principais e de variáveis canônicas

Com intuito de representar graficamente os acessos ($n = 527$) num espaço de reduzido número de dimensões, aplicaram-se as análises de componentes principais e de variáveis canônicas. O uso das duas técnicas teve como propósito avaliar a concordância entre elas. Para essas análises, consideraram-se, simultaneamente, as sete variáveis ($p = 7$) padronizadas (teores de óleo, proteína e dos ácidos graxos palmítico, esteárico, oleico, linoleico e linolênico). Ambas as técnicas permitem a análise exploratória da diversidade entre os acessos em gráficos de dispersão dos elementos no espaço das variáveis. Além disso, permitem verificar as variáveis originais com maior peso ou contribuição para a discriminação dos acessos. Na análise de componentes principais utilizou-se a matriz de correlações fenotípicas, o que é equivalente à análise dos dados a partir das variáveis

padronizadas. Na análise de variáveis canônicas, além da matriz de correlações fenotípicas, utilizou-se a matriz de variâncias-covariâncias residuais estimadas do ajustamento do modelo (1). A diferença entre as duas análises está na vantagem desta última técnica em considerar as correlações residuais (Cruz & Regazzi, 2001). Na representação gráfica optou-se pela técnica que discriminou os acessos com menor número de combinações lineares das variáveis originais.

Para a construção da matriz de variâncias-covariâncias residuais, $\Sigma_{(p)}$, foram calculados os quadrados médios residuais (QM_{res}) para cada variável, bem como o QM_{res} para as somas das variáveis tomadas duas a duas. As covariâncias residuais (COV_{res}) entre duas variáveis quaisquer (var_1 e var_2) foram estimadas pela seguinte equação (Cruz & Regazzi, 2001):

$$COV_{res(var_1, var_2)} = \frac{QM_{res(var_1+var_2)} - QM_{res(var_1)} - QM_{res(var_2)}}{2}$$

A análise de componentes principais foi executada nos aplicativos computacionais SAS (Proc Princomp) e R (R Development Core Team, 2010), e a análise de variáveis canônicas, no aplicativo Genes (Cruz, 1998). A dispersão gráfica dos acessos e a representação dos coeficientes de ponderação das variáveis foram feitas apenas para as variáveis canônicas, utilizando-se o aplicativo R. O número de variáveis canônicas escolhido foi aquele que representou, no mínimo, 70% da variação total dos dados (Jolliffe, 2002).

A análise em variáveis canônicas também foi aplicada ao conjunto dos acessos selecionados como de melhor perfil de estabilidade oxidativa no óleo. Com intuito de visualizar o grau de separação entre esses acessos, fez-se também a sua dispersão gráfica nos eixos canônicos.

3.3.3 Análise de agrupamento

Para avaliar a divergência entre os acessos e identificar possíveis grupos de similaridade entre eles, foi utilizado o método de agrupamento hierárquico aglomerativo, associado aos critérios de ligação pelo vizinho mais próximo, vizinho mais distante, média UPGMA (*unweighted pair group method with arithmetical*) e método de Ward. O método foi aplicado à matriz de dissimilaridade entre os acessos, obtida com base na distância Mahalanobis, considerando-se, simultaneamente, as sete variáveis padronizadas (teores de

óleo, proteína e dos ácidos graxos palmítico, esteárico, oleico, linoleico e linolênico). Para sua obtenção, utilizou-se o aplicativo Genes, que retorna a matriz de distância entre as linhas ($D_{(n)}$) de uma matriz de dados, a partir das médias ajustadas ($X_{n \times p}$) e da matriz de variâncias-covariâncias residuais ($\Sigma_{(p)}$). Após obter a matriz $D_{(n)}$, a análise de agrupamento e o respectivo dendrograma foram implementados via a função *hclust*, no aplicativo R.

O dendrograma foi produzido somente a partir do critério UPGMA, que resulta em máxima correlação cofenética. Ainda assim, as correlações cofenéticas associadas aos demais critérios foram obtidas, a fim de compará-las à do critério UPGMA. Os valores cofenéticos foram obtidos a partir das distâncias entre os acessos no dendrograma. Esses valores são representados pela distâncias em que os acessos, tomados dois a dois, se tornam membros de um mesmo conglomerado no processo de agrupamento (Legendre & Legendre, 1998). Assim fica definida a matriz de distâncias cofenéticas, cujos valores foram correlacionados com aqueles da matriz de distâncias originais. As correlações cofenéticas foram calculadas no aplicativo R, através de suas funções *cophenetic* e *mantel*. A significância dessa correlação foi testada pelo teste de Mantel, utilizando-se 10.000 permutações aleatórias. As instruções computacionais para uso das funções do aplicativo R, aqui descritas, estão no Apêndice C.

No dendrograma, os acessos foram identificados também quanto aos seus níveis de óleo, proteína e de ácidos graxos saturados e poli-insaturados. Para isso, foram categorizados em três classes, com níveis equidistantes (amplitude dividida por três): Baixo, Médio e Alto. A identificação destes níveis em cada acesso foi feita utilizando-se a função *A2Rplot* do pacote *A2R*, também no aplicativo R.

Para orientar a definição do número de grupos, foram calculados os valores da estatística Pseudo F, a cada passo do agrupamento. Assim, adotaram-se os valores máximos dessa estatística como indicadores de maior heterogeneidade entre os conglomerados (Mingoti, 2007). Para isto utilizou-se a função *calinhara* do pacote *fpc*, também do aplicativo R.

Após a definição dos grupos de similaridade, via análise de agrupamento, estes foram identificados graficamente também na dispersão dos acessos por análise de variáveis canônicas. Isto permite verificar a concordância entre essas técnicas e, ainda, observar a similaridade entre os acessos de um mesmo grupo na dispersão. Também foram realizadas análises de variância univariadas, no modelo inteiramente casualizado, assumindo-se cada

grupo como um tratamento experimental. O objetivo foi avaliar a discriminação entre os grupos identificados, com respeito a cada atributo avaliado do grão.

Todas essas análises foram também aplicadas ao conjunto de acessos selecionados (com melhor perfil em estabilidade oxidativa). Obteve-se o respectivo dendrograma e, com o auxílio dos valores de Pseudo F, definiram-se os grupos de similaridade dentro desse conjunto. Estes grupos também foram identificados graficamente na dispersão por análise de variáveis canônicas.

3.3.4 Análise discriminante canônica

Após a definição dos grupos de similaridade, aplicou-se a análise discriminante canônica, com intuito de identificar as variáveis que melhor discriminaram esses grupos. A idéia principal dessa análise é maximizar a separação entre classes dos elementos, neste caso os grupos de similaridade, baseando-se numa combinação linear das variáveis, que tem as maiores correlações possíveis com as classes. Assim, a análise permite visualizar o grau de separação entre os grupos por meio de suas projeções no espaço das variáveis canônicas (SAS Institute, 2010).

Buscando-se discriminar os ciclos de maturação, no que tange aos caracteres avaliados, mesmo admitindo-se o confundimento de seus efeitos com os de grupos (superblocos), aplicou-se novamente a análise discriminante canônica. Nesse caso, as classes dos elementos que foram separadas são os ciclos de maturação. Foi feita também a dispersão das variáveis discriminantes canônicas com o intuito de visualizar o grau de separação entre os ciclos.

Para essa análise e em ambos os casos, utilizou-se o procedimento Proc Candisc do aplicativo SAS. Este procedimento permite, ainda, o teste da hipótese de igualdade das médias entre as classes, para cada variável, usando-se sua opção *anova*.

4 RESULTADOS E DISCUSSÃO

4.1 VARIABILIDADE DO GERMOPLASMA: ENFOQUE UNIVARIADO

O banco de germoplasma da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (BAG/Seagro-GO) mostrou variabilidade fenotípica entre os acessos, para todos os caracteres avaliados (teores de óleo, proteína e dos ácidos graxos palmítico, esteárico, oleico, linoleico e linolênico no grão) (Apêndice A). Priolli et al. (2004) já haviam relatado que as cultivares da Emgopa/Emater-GO (atual Seagro-GO) apresentavam o segundo maior nível de diversidade genética, atrás somente daquelas do Instituto Agronômico de Campinas (IAC). Os autores caracterizaram por marcadores microsatélites 186 cultivares brasileiras, incluindo também os programas de melhoramento da Coodetec, Monsoy, Embrapa e Indusem.

Os teores de proteína no grão variaram entre 35,7% (Bag 645) e 44,0% (Bag 308), com média de 39,0%, e os de óleo, entre 14,8% (Bag 308) e 18,5% (Bag 524), com média de 16,7% (Tabela 1). O acesso Bag 308 (PI-219653), que apresentou o maior teor médio de proteína e o menor teor de óleo, é uma linhagem introduzida (PI); assim como os dois acessos que a seguem em teor proteico, Bag 146 (PI-88788) e Bag 313 (PI-240667). Estes, no entanto, não apresentaram teores de óleo tão baixos.

Comparando-se os teores médios encontrados no banco de germoplasma da Seagro-GO aos da coleção base de germoplasma de soja do Departamento de Agricultura dos Estados Unidos (USDA), com 16.981 acessos, verifica-se que a primeira apresenta baixa variabilidade em conteúdo de óleo e proteína. Contrariamente à coleção nacional, os acessos da coleção americana apresentaram teores de proteína entre 31,7% e 57,9%, e de óleo variando entre 8,1% e 25,4% (Oliveira et al., 2010). Iqbal et al. (2008) avaliaram 139 acessos do Centro Nacional de Pesquisa Agropecuária do Paquistão e, também, encontraram maior variação nos teores de óleo, de 17,4% a 22,9%. Nos dois casos, verifica-se que, além da maior variabilidade, essas coleções dispõem de acessos com níveis desejáveis bem superiores aos da presente pesquisa, em ambos os caracteres. Para aumentar a variabilidade genética nesses caracteres e, conseqüentemente, o potencial de

uso dos recursos genéticos do BAG/Seagro-GO, seria necessário, portanto, a introdução de acessos com maiores teores de óleo e, ou, de proteína no grão. Deve-se ressaltar, ainda, que as menores amplitudes observadas nesta subcoleção, em relação às coleções anteriormente referidas, embora possa ser estatisticamente justificada pelo seu menor tamanho, não deixam de caracterizar sua limitação em acessos com teores desejáveis.

Tabela 1. Valores mínimos, máximos e médio de sete caracteres (teores de óleo, proteína e dos ácidos graxos palmítico, esteárico, oleico, linoleico e linolênico), com respectivos desvios padrão (DP), coeficientes de variação genética (CV_g) e ambiental (CV_e), bem como relação entre eles, para 527 acessos do banco de germoplasma de soja da Seagro-GO¹.

Estimativas ²	PROT	OL	PAL	EST	OLC	LIN	LIL
Mínimo	35,7	14,8	9,6	1,4	17,6	31,3	3,6
Máximo	44,0	18,5	17,4	3,8	48,9	62,1	8,1
Média	39,0	16,7	11,9	2,1	25,3	52,6	6,5
DP	1,1	0,5	0,8	0,3	4,7	4,1	0,6
CV_g	3,8	5,7	7,4	18,3	20,1	8,2	10,5
CV_e	3,1	9,7	3,5	14,6	8,1	2,4	7,4
CV_g/ CV_e	1,2	0,6	2,1	1,3	2,5	3,4	1,4
USDA ³	31,7 - 52,9	8,1 - 25,4	8,0 - 18,4	1,1 - 10,6	11,1 - 50,3	30,3 - 65,0	3,4 - 18,1

¹ - Seagro-GO: Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás.

² - OL: teor de óleo em g/100g; PROT: teor de proteína em g/100g; PAL: teor de ácido graxo palmítico em g/100g; EST: teor de ácido graxo esteárico em g/100g; OLC: teor de ácido graxo oleico em g/100g; LIN: teor de ácido graxo linoleico em g/100g e LIL: teor de ácido graxo linolênico em g/100g.

³ - Teores mínimos e máximos para as sete variáveis, na coleção base de soja do Departamento de Agricultura dos Estados Unidos (USDA) (Oliveira et al., 2010).

Wehrmann et al. (1987) e Scott & Kephart (1997) registram o uso de introduções na transferência de elevados teores de proteína para cultivares com alta produtividade. O desafio do melhoramento é aproveitar genes de interesse, provenientes desse tipo de material, e introduzi-los em cultivares elites sem perder os padrões agrônomicos já alcançados. Para se aumentar a utilização dos recursos genéticos disponíveis, Nass et al. (2007) sugerem a criação de programas de pré-melhoramento. Nestes programas seriam feitas a identificação de genitores com características promissoras e, posteriormente, a incorporação de suas características em materiais agronomicamente adaptados.

Deve-se destacar que os teores relativamente baixos de óleo observados na presente pesquisa não podem ser atribuídos somente ao potencial genético dos acessos da coleção. Isso porque seis desses acessos foram também avaliados em outra pesquisa

(Bonato et al., 2000), porém, sempre com teores de óleo mais elevados. Tais acessos, com os respectivos teores médios observados nas duas pesquisas (o primeiro valor refere-se ao presente estudo), são: Bag 594 (15,7% e 18,9%), Bag 168 (16,3% e 19,5%), Bag 625 (16,7% e 19,1%), Bag 367 (17,1% e 19,1%), Bag 599 (17,4% e 20,3%) e Bag 728 (17,4% e 20,4%). Outro acesso, Bag 3, apresentou apenas 16,3% de óleo no presente estudo, porém, acumulou 22,9% de óleo no grão, nos experimentos de Brogin et al. (2010). É possível inferir, portanto, que as condições de avaliação da presente pesquisa desfavoreceram a manifestação do caráter, rebaixando os teores médios de óleo da maioria dos acessos.

Quanto aos teores de proteína, os valores médios observados naqueles mesmos acessos também foram inferiores aos respectivos valores encontrados por Bonato et al. (2000), exceto nos acessos Bag 599 (37,8%) e Bag 728 (38,8%), embora com pequena diferença. Esse resultado não era esperado, hajam vista os baixos teores de óleo observados e a correlação negativa amplamente reportada entre conteúdos de óleo e proteína no grão de soja (Bonato et al., 2000; Feng et al., 2004). Neste caso, os teores médios observados nas duas pesquisas (primeiro valor referente ao presente estudo) foram: Bag 594 (41,3% e 42,3%), Bag 168 (38,1% e 40,3%), Bag 625 (39,2% e 40,8%), Bag 367 (37,6% e 39,3%), Bag 599 (38,0% e 37,8%) e Bag 728 (40,3% e 38,3%).

Diante disso, além das condições ambientais que podem ter desfavorecido a manifestação dos caracteres, sobretudo o teor de óleo, deve-se admitir também a possibilidade de problemas de ordem metodológica nas análises físico-químicas; reiterando-se, em especial, na determinação do teor de óleo no grão, para o qual as diferenças foram mais nítidas. A granulometria da farinha de soja utilizada para as análises pode ser um ponto crítico de controle, que deve ser reavaliado. Outro ponto crítico pode ser o tempo de extração utilizado na quantificação do óleo (três horas e trinta minutos). Na presente pesquisa esse tempo não foi aumentado, pois isto implicaria em maior degradação de ácidos graxos, os quais seriam determinados a partir do óleo extraído. Alternativamente, poder-se-ia elevar a temperatura de extração, porém, isto também implicaria na maior degradação. A granulometria da farinha, se alta, pode dificultar a extração do óleo e, ainda, provocar amostragem heterogênea de partes do grão nas diferentes análises. Sob tais condições, tempo maior do que o preconizado pode ser necessário para a completa extração do óleo; o que ainda não é explicitado claramente na literatura.

Segundo as normas da AOAC (Deutsch, 1995; Nielsen, 2003) e do Instituto Adolfo Lutz (2008), o tempo de extração do extrato etéreo pode variar de quatro a dezesseis horas. Em soja, podem-se citar variações no tempo de extração do óleo: três horas, em Ribeiro (2006); oito horas, em Lacerda et al. (2003); e dez horas, em Moraes et al. (2006). Oliveira (2003), determinando o teor de óleo em milho (5% do óleo), utilizou três horas de extração. Monteiro et al. (2004) utilizaram de quatro a cinco horas na determinação do teor de lipídeos em carcaças de ratos. Em outras pesquisas, que envolveram determinação de teores de óleo em soja, pelo método de Soxhlet, esse tempo não foi explicitado (Mascarenhas et al., 1996; Bonato et al., 2000; Café et al., 2006).

Entre os ácidos graxos avaliados, os teores dos ácidos palmítico e esteárico representam o grupo dos ácidos graxos saturados, cujos níveis desejáveis tanto em dietas para alimentação humana quanto para produção de biodiesel devem ser baixos. Os teores de ácido palmítico variaram de 9,6% (Bag 312) a 17,4% (Bag 51), com média de 11,9% (Tabela 1). Os acessos com menores teores desse ácido foram as introduções Bag 312 e Bag 517. O destaque fica para essa última, por apresentar também o segundo maior teor de ácido oleico, que é um constituinte de interesse em óleos vegetais para alimentação.

Tomando-se outra vez como exemplo a grande coleção base do USDA, nos EUA, Oliveira et al. (2010) reporta variação entre 8,0% e 18,4%, para o ácido palmítico. Logo, a variabilidade observada nas duas coleções é similar, indicando boa diversidade na coleção mantida pela Seagro-GO, para esse atributo. A média destes acessos para esse ácido foi de 11,9%, o que, indesejavelmente, é superior à média de cultivares comerciais, 10,8% (Sanibal & Mancini Filho, 2004). Contudo, ainda restam na coleção 45 acessos com teores desejáveis, isto é, abaixo dessa última média. Uma vez que o ácido palmítico representa o maior conteúdo de ácidos saturados, vislumbra-se a possibilidade de uso deste germoplasma para algum melhoramento para qualidade de óleo destinado à alimentação. Embora haja registros de linhagens com teores de ácido palmítico inferiores a 7% (Burton et al., 1983; Wilcox & Cavins, 1990; Fehr et al., 1991; Burton et al., 1994; Stojsin et al., 1998).

Os teores de ácido esteárico variaram de 1,4% (Bag 352) a 3,8% (Bag 909), com média 2,1%. Os níveis observados, na maioria dos acessos, são desejáveis à indústria de alimentos ou de biodiesel, com apenas doze acessos apresentando teores acima de 3,0% para esse atributo. O maior índice encontrado esteve próximo à média de cultivares comerciais, que é de 3,3% (Sanibal & Mancini Filho, 2004), evidenciando que a maior

parte do germoplasma apresenta níveis de ácido esteárico reduzidos. No germoplasma americano a variação desse atributo foi de 1,1% a 10,6% (Oliveira et al., 2010). Logo, apenas o teor mínimo observado no germoplasma da Seagro-GO é semelhante ao daquela coleção de referência, cujos acessos atingem níveis quase três vezes superiores. Embora limites assim tão altos não sejam atualmente demandados, sob a perspectiva de conservação genética, isto revela certa carência de variabilidade.

Os teores de ácido oleico variaram de 17,6% (Bag 514) a 48,9% (Bag 553), com média de 25,3% (Tabela 1). O ácido oleico é importante para a indústria de óleo de soja, por ser o ácido graxo insaturado mais estável à oxidação. Esse ácido é o principal representante do grupo dos ácidos monoinsaturados no óleo de soja. No germoplasma do USDA, a amplitude para o teor deste ácido graxo foi um pouco maior, de 11,1% a 50,3% (Oliveira et al., 2010), embora seu valor máximo encontra-se próximo ao respectivo valor na coleção em estudo. A média observada (25,3%) também foi superior à de cultivares comerciais, 21,5% (Sanibal & Mancini Filho, 2004). Ademais, entre os acessos, 418 apresentaram teores de ácido oleico superiores a essa média, ratificando o grande potencial da coleção mantida pela Seagro-GO, para uso no melhoramento da soja neste importante caráter. É importante destacar que existem linhagens de soja com teores de ácido oleico acima de 50% (Rahman et al., 1994; Burton et al., 2006; Rahman et al., 2001; Anai et al., 2008), que representam, portanto, recursos genéticos para o enriquecimento do germoplasma em relação à qualidade do óleo.

O acesso Bag 553, além de se destacar com o maior teor de ácido oleico (48,9%), apresentou o menor nível de ácido linoleico (31,3%), o que implicaria na produção de óleo com alta estabilidade oxidativa. Este acesso é a cultivar Sharkey (Apêndice A), de origem americana, a qual é derivada da cultivar Peking (Ha et al., 2004), muito utilizada em cruzamentos para incorporação de resistência ao nematoide de cisto da soja (*Heterodera glycines*) (Dias et al., 2009). Logo, tal característica apresenta-se como vantagem adicional ao se utilizar este acesso em cruzamentos visando a obtenção de genótipos que produzam óleo de maior estabilidade.

Os ácidos graxos linoleico e linolênico são os principais representantes do grupo de ácidos poli-insaturados. Altos níveis destes ácidos são responsáveis pela redução da estabilidade do óleo de soja. Os acessos avaliados nesta pesquisa apresentaram variações de 31,3% (Bag 553) a 62,1% (Bag 586), com média de 52,6%, para o ácido linoleico, e de 3,6% (Bag 278) e 8,1% (Bag 514), com média de 6,5%, para o ácido

linolênico. O teor médio destes ácidos em cultivares comerciais é 55,1% e 4,8%, respectivamente (Sanibal & Mancini Filho, 2004). A maior parte do germoplasma apresentou teores de ácido linoleico abaixo da média de cultivares comerciais, mas, teores abaixo de 40% ficaram restritos a sete acessos. A menor média, conforme já comentado, foi do Bag 553 (cultivar Sharkey), que se destacou com o maior teor de ácido oleico; já a segunda menor, 35,3%, foi do Bag 278 (BRGO99-4095-05), que, entretanto, apresentou baixo teor de ácido oleico. Não se pode inferir sobre a qualidade do óleo para indústria avaliando-se somente a presença de reduzidos teores de ácidos poli-insaturados. Isso porque, aliada a essa característica, são desejáveis também altos teores de monoinsaturados.

As variações encontradas no germoplasma do USDA foram de 30,3% a 65,0%, para o ácido linoleico, e de 3,4% a 18,1% para o ácido linolênico (Oliveira et al., 2010). Portanto, para o ácido linoleico, verificam-se amplitudes bastante similares nas duas coleções. Porém, no que concerne ao ácido linolênico, apenas os limites mínimos são similares; o teor máximo deste ácido na coleção americana é mais que duas vezes o máximo observado no germoplasma da Seagro-GO. Isto, porém, não se constitui problema se o objetivo é obter cultivares que produzam óleos com maior estabilidade oxidativa, pois neste caso buscam-se valores baixos de ambos os ácidos; a menos que o interesse esteja, por exemplo, em genótipos com teores muito baixos de ácido linolênico, isto é, inferiores a 2%, como afirmam Clemente & Cahoon (2009). Nestes casos, para se aumentar o potencial de obtenção de linhagens melhoradas (com teores baixos desses ácidos), a partir desse germoplasma, ter-se-ia que recorrer a novas introduções. Há relatos de linhagens obtidas por mutagênese, com teores abaixo de 3% de ácido linolênico no óleo (Hammond & Fehr, 1983; Fehr et al., 1992; Fehr & Hammond, 1996), bem como com teores menores que 33% para o ácido linoleico (Rahman et al., 1994; Rebetzke & Pantalone, 2006; Anai et al., 2008).

Outra abordagem para os teores de ácidos graxos poli-insaturados é em relação à sua contribuição na qualidade da dieta. Estes ácidos graxos, conhecidos como PUFA (do inglês, *polyunsaturated fatty acids*), podem ser classificados em dois grupos: ômega-6 e ômega-3. Os ácidos linoleico e linolênico são, respectivamente, os principais representantes destes grupos, inclusive no óleo de soja (Topfer et al., 1999; Visentainer & Franco, 2006). Os ácidos ômega-6 e ômega-3 são essenciais ao organismo humano (somente adquiridos na dieta) e, ainda, controlam e regulam muitos processos corporais

importantes (Katan, 2009). Sabe-se também que a redução da proporção entre ômega-6 e ômega-3 na dieta, mais especificamente entre os ácidos linoleico e linolênico, pode exercer efeito supressor no surgimento de doenças cardiovasculares, câncer e inflamatórias (Simopoulos, 2008; Deckelbaum, 2010; Farzaneh-Far et al., 2010). Entre os dois grupos, o tipo ômega-3 ainda se destaca por ser um elemento nutracêutico, isto é, com propriedades funcionais (Topfer et al., 1995); de modo que a Anvisa (Agência Nacional de Vigilância Sanitária), em sua Resolução nº 18, de 30 de abril de 1999 (Brasil, 1999), recomenda, para a manutenção de níveis saudáveis de triglicerídeos, alimentos que possuam esse tipo de propriedade.

Nesse contexto, se houver interesse na obtenção de produtos de soja com elevados teores de ácidos graxos poli-insaturados, os resultados revelam que a subcoleção da Seagro-GO possui variabilidade genética satisfatória nos teores de ácido linoleico. Semelhantemente ao discutido para o teor de ácido esteárico, o limite superior para o ácido linolênico esteve abaixo dos índices encontrados em germoplasma de soja. Isto deve ser considerado quando o interesse prioritário é a conservação de recursos genéticos da espécie. Por exemplo, se admitirmos a possibilidade de que altos índices desse ácido possam ser demandados no futuro; uma vez que ácidos graxos ômega-3 (ácido linolênico) são mais desejáveis na dieta que os ômega-6 (ácido linoleico).

Avaliando-se simultaneamente as variações relativas dos caracteres, em nível genético e ambiental (CV_g e CV_e), observa-se que, exceto para o caráter teor de óleo, os valores de CV_g foram superiores àqueles de CV_e (Tabela 1), o que comprova a existência de variabilidade genética na coleção de germoplasma. Os baixos valores de CV_g observados para os teores de proteína e óleo no germoplasma também corroboram a baixa variabilidade discutida anteriormente. Os maiores coeficientes CV_e foram encontrados nos teores de ácido esteárico (14,6%), óleo (9,7%) e ácido oleico (8,1%). Esses valores, em geral, foram baixos, embora para o ácido esteárico a estimativa mostrou-se relativamente discrepante aos demais caracteres. Isso pode ser atribuído aos baixos índices deste caráter no óleo de soja, o que, inclusive, dificulta a sua mensuração. O valor de CV_e para teor de proteína foi de 3,1%, semelhante ao encontrado por Marega Filho et al. (2001), 3,9%. Estes autores, contudo, obtiveram menor valor de CV_e para o teor de óleo, 5,9%. O teor de ácido oleico foi o caráter que apresentou o maior coeficiente de variação genética (20,1%), seguido dos teores dos ácidos esteárico (18,3%) e linolênico (10,5%). Esses valores indicam maior variabilidade genética disponível no germoplasma, para estes caracteres.

Apesar disso, para os ácidos esteárico e linolênico, a subcoleção apresentara limites superiores abaixo daqueles encontrados em germoplasma de soja (Oliveira et al., 2010).

A relação CV_g/CV_e reflete melhor a referida variação entre os acessos, representando valiosa informação adicional para o melhorista (Vencovsky & Barriga, 1992). As variáveis com os maiores valores para esta relação foram: ácido linoleico (3,4); oleico (2,5); e palmítico (2,1). Segundo Vencovsky & Barriga (1992), valores de CV_g/CV_e superiores à unidade refletem uma situação muito favorável à seleção. Na presente pesquisa, a relação apresentou valor abaixo da unidade apenas para o caráter teor de óleo, o que indica baixo potencial do germoplasma para seleção direta neste caráter; nos demais, há, portanto, maior potencial genético a ser explorado.

4.2 HERDABILIDADE E CORRELAÇÃO DE CARACTERES

As estimativas de herdabilidade encontradas foram relativamente altas, exceto para o teor de óleo (0,26). Os maiores valores observados foram: 0,92 para ácido linoleico; 0,86 para ácido oleico; e 0,82 para ácido palmítico (Tabela 2). Coerentemente, esses caracteres foram aqueles que apresentaram também os maiores valores para a relação CV_g/CV_e . A possibilidade de ganhos genéticos obtidos por uma estratégia de seleção é função direta da herdabilidade do caráter (Cruz & Regazzi, 2001). Dessa forma, os valores observados indicam que a seleção direta para os teores dos ácidos linoleico, oleico e palmítico representa uma situação bastante favorável para obtenção de genótipos superiores. Bachlava et al. (2008) encontraram valores de herdabilidade um pouco menores para os ácidos linoleico (0,78) e oleico (0,76), porém, um valor maior para o ácido linolênico (0,92).

O valor de herdabilidade estimado para o teor de óleo (0,26) foi inferior aos relatados em outras pesquisas (Johnson et al., 1955; Kwon & Torrie, 1964; Smith & Weber, 1968; Shorter et al., 1976; Silva, 2008). Panthee et al. (2005) encontraram herdabilidade de 0,54; e, mais recentemente, Bachlava et al. (2008), no estudo de três populações segregantes, observaram valores ainda maiores, entre 0,33 e 0,69. Para teor de proteína, a herdabilidade estimada foi de 0,59, indicando que a maior parte da variação observada no caráter ainda é de origem genética. Esse valor foi pouco menor que a herdabilidade de 0,66, encontrada por Panthee et al. (2005).

Observando-se a matriz de correlações entre os caracteres avaliados, verifica-se correlação genética ($p < 0,05$) de cada ácido graxo com, pelo menos, outra variável relacionada à qualidade do grão (Tabela 2). Analisando-se primeiramente a correlação dos ácidos graxos saturados com os outros caracteres, nota-se que o ácido palmítico mostrou correlação genética negativa com os teores de óleo (-0,77) e de ácido oleico (-0,41); e correlação positiva com o ácido esteárico (0,52). As correlações negativas observadas também foram relatadas por Horejsi et al. (1994), Ndzana et al. (1994) e Rebetzke et al. (1998). O ácido esteárico, embora correlacionado positivamente com o ácido palmítico, mostrou correlação positiva com o teor de óleo (0,59). Um estudo adicional por análise de trilha (Cruz & Regazzi, 2001) poderia ajudar no entendimento dos efeitos diretos e indiretos envolvidos nessas correlações.

Tabela 2. Coeficientes de correlação genotípica (abaixo da diagonal) e fenotípica (acima da diagonal) entre sete caracteres ligados à qualidade do grão de soja e herdabilidade (h^2), em nível de médias, avaliados em 527 acessos de soja do banco de germoplasma da Seagro-GO¹.

Caracteres ²	PROT	OL	PAL	EST	OLC	LIN	LIL
PROT	-	-0,29	-0,10	-0,13	0,24	-0,25	-0,01
OL	-0,11	-	-0,23	-0,07	0,20	-0,11	-0,39*
PAL	-0,12	-0,77**	-	0,36	-0,51**	0,28	0,33
EST	-0,34	0,59**	0,52**	-	-0,40*	0,26	0,02
OLC	0,48**	-0,15	-0,41*	-0,31	-	-0,95**	-0,66**
LIN	-0,47**	-0,27	0,24	0,27	-1,00**	-	0,58**
LIL	-0,33	-0,36	0,27	-0,10	-0,64**	0,60**	-
h^2	0,59	0,26	0,82	0,61	0,86	0,92	0,67

** e *: valores significativos pelo teste t-Student, a 1% e a 5% de probabilidade, respectivamente.

¹- Seagro-GO: Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás.

²- OL: teor de óleo em g/100g; PROT: teor de proteína em g/100g; PAL: teor de ácido graxo palmítico em g/100g; EST: teor de ácido graxo esteárico em g/100g; OLC: teor de ácido graxo oleico em g/100g; LIN: teor de ácido graxo linoleico em g/100g e LIL: teor de ácido graxo linolênico em g/100g.

A correlação fenotípica para esses mesmos pares de caracteres somente foi significativa entre os teores de ácido palmítico e ácido oleico (-0,51). Os ácidos palmítico e esteárico são os principais ácidos graxos saturados no óleo de soja; e, os resultados obtidos sugerem que reduções nos teores de ácido palmítico tendem a ser acompanhadas por redução também nos teores de ácido esteárico. Logo, se o objetivo do melhoramento é reduzir teores de ácidos graxos saturados, isto poderia ser alcançado mais rapidamente, já

que a redução em um destes ácidos leva à redução também no outro, acumulando ganhos na direção daquele objetivo. Entre os teores dos ácidos esteárico e oleico observou-se correlação fenotípica negativa (-0,40); e, não tendo sido detectada correlação genética entre estes caracteres, sugere-se uma influência inversa dos fatores ambientais sobre a concentração desses ácidos graxos. Os resultados também indicam que incrementos nos teores de ácido oleico podem ser alcançados sem que se aumentem os teores de ácidos saturados no grão de soja. Do ponto de vista industrial, altos teores de ácido oleico garantem melhor estabilidade ao óleo de soja; e, ainda, baixos teores de ácidos graxos saturados (palmítico e esteárico) têm sido desejados na indústria de alimentos, devido à preocupação em reduzir o consumo destes ácidos e, conseqüentemente, os riscos associados a doenças cardiovasculares (Erkkila et al., 2008). Portanto, os resultados de correlação indicam que a seleção visando maior estabilidade do óleo pode trazer, como consequência, melhoria também na qualidade nutricional do óleo de soja.

Analisando-se as correlações dos ácidos graxos insaturados com as outras variáveis (Tabela 2), verifica-se que o ácido oleico apresentou correlação genética positiva com o teor de proteína (0,48) e negativa com os teores dos ácidos linoleico (-1,00) e linolênico (-0,64). Bachlava et al. (2008) também encontraram correlação genética alta entre os ácidos oleico e linoleico (-0,97). A correlação praticamente perfeita entre os teores dos ácidos linoleico e oleico pode ser explicada pela dependência da síntese dos ácidos poli-insaturados, a partir dos ácidos monoinsaturados. Na biossíntese de ácidos graxos, os ácidos linoleico e linolênico são obtidos a partir da conversão do ácido oleico, por meio da atividade de enzimas desaturases. A enzima delta-12 desaturase é responsável pela conversão do ácido oleico em linoleico, e a delta-15 desaturase, pela conversão deste último em ácido linolênico (Topfer et al., 1995; Ji et al., 2009; Zhang et al., 2009).

Outra correlação de grande interesse aplicado é aquela entre o teor de ácido oleico e o teor de óleo. Bachlava et al. (2008) encontraram correlação genética negativa de -0,40; e Clemente & Cahoon (2009) alegam que o plantio comercial de linhagens com maiores teores desse ácido ainda não se materializou, principalmente devido à esta correlação. No presente estudo, porém, esta correlação, embora tenha apresentado o mesmo sinal, não se revelou significativa. Isto revela que não há evidências definitivas de que a seleção visando aumentar os teores de ácido oleico só pode ocorrer em detrimento dos teores de óleo e proteína no grão de soja.

O teor de ácido linoleico apresentou, ainda, correlação genética negativa com o teor de proteína (-0,47), e positiva com o ácido linolênico (0,60) (Tabela 2). Entre esses pares, a correlação fenotípica só se confirmou entre os teores dos ácidos linoleico e linolênico (0,58). Entre teor de ácido linolênico e teor de óleo observou-se significância somente para a correlação fenotípica, que foi negativa (-0,39). A correspondente correlação genética, embora com mesmo sinal (-0,36), não atingiu significância estatística a 5% de probabilidade ($p=0,06$).

Esses resultados sugerem que aumentos no teor de ácido oleico, em geral, são acompanhados por redução na concentração de ácidos graxos saturados e poli-insaturados no grão de soja. A forte correlação negativa entre os ácidos oleico e linoleico motiva esforços na tentativa de melhorar a qualidade oxidativa do óleo de soja. Isto se daria, principalmente, pelo aumento do teor de ácido oleico e redução do teor de ácidos poli-insaturados; e, sabidamente, o ácido linoleico é o ácido graxo poli-insaturado de maior concentração no óleo de soja (30% a 60%). Baixas concentrações de ácido linoleico provocam melhor estabilidade oxidativa no óleo e, conseqüentemente, em seus subprodutos como, por exemplo, o biodiesel (Fehr, 2007; Garcia & Romeiro, 2009; Graef et al., 2009). Devido à elevada correlação negativa, espera-se que reduções substanciais nos teores de ácido linoleico acarretem aumentos significativos nos teores de ácido oleico, e vice-versa. Logo, a correlação entre estes caracteres colabora com o objetivo do melhoramento genético visando melhoria na qualidade do óleo de soja (Alt et al., 2005; Bachlava et al., 2008).

Segundo Vencovsky & Barriga (1992), em certas situações, o melhoramento genético num caráter Y pode ser conseguido mais eficientemente selecionando-se em outro (X). Isto seria possível quando, teoricamente, a resposta correlacionada em Y, pela seleção em X, fosse maior que o ganho genético obtido com a seleção direta em Y. Os autores acrescentam que isso não é muito comum, mas pode acontecer quando o produto da raiz quadrada da herdabilidade de X pela correlação entre X e Y for maior que a raiz quadrada da herdabilidade de Y. Baseado nisso, pode-se aventar a possibilidade de maior resposta correlacionada para o teor de ácido oleico, via seleção indireta no teor de ácido linoleico. Vale acrescentar que o valor da relação CV_g/CV_e para o ácido linoleico foi superior ao do ácido oleico (Tabela 1), o que reforça a melhor condição de seleção no teor de ácido linoleico. Segundo Cruz & Regazzi (2001), a seleção indireta é interessante quando se deseja obter ganhos em caracteres de grande importância, mas, por questões de ordem

prática, econômica ou ambiental, a seleção precisar ser praticada em caracteres auxiliares. No caso presente esse interesse não se materializa, pois a metodologia de quantificação dos ácidos graxos retorna, simultaneamente, os teores de todos eles. Contudo, numa situação em que a quantificação desses ácidos venha ser feita separadamente, a seleção indireta pode se justificar com impacto favorável na redução do volume de análises.

4.3 DIVERGÊNCIA GENÉTICA ENTRE ACESSOS: ENFOQUE MULTIVARIADO

4.3.1 Estrutura da diversidade entre acessos

Na análise de divergência entre os acessos por componentes principais, as três primeiras componentes explicaram, respectivamente, 43%, 20% e 15%. Jolliffe (2002) sugere que o modelo reduzido explique 70% a 90% da variância total. Dessa forma, seriam necessárias, no mínimo, três componentes para explicar 78% da variância total. Situações em que o número de variáveis é grande implicam em dificuldades de se reter variâncias superiores a 70%, com poucas componentes principais. Para essas situações sugerem-se outros métodos que auxiliam na decisão sobre quantas componentes principais devem ser utilizados no modelo reduzido (Mardia et al., 1979; Jolliffe, 2002; Ferreira, 2008).

Esse resultado, aliado à possibilidade de uso da informação da matriz de covariâncias residuais, justificou a opção definitiva pela análise de variáveis canônicas, na análise da dispersão dos acessos em espaço vetorial de dimensão reduzida. As três primeiras variáveis canônicas acumularam, respectivamente, 58,5%, 81,8% e 97,6% da variância total. Assim, duas variáveis canônicas foram suficientes para explicar grande parte da variação (81,8%), facilitando o estudo de divergência entre os acessos por meio da dispersão gráfica em apenas duas dimensões. Na análise de componentes principais, entretanto, para se explicar quase 80% dessa variação seriam necessários três componentes; isto é, a análise gráfica deveria ser feita numa representação tridimensional, o que dificulta a interpretação.

A dispersão dos escores genotípicos nas duas primeiras variáveis canônicas não evidencia separação nítida de grupos de acessos (Figura 1). Ou seja, os acessos não apresentaram tendência de agrupamento evidente na dispersão, em relação às duas primeiras variáveis canônicas. Se isso aconteceu com as duas primeiras variáveis canônicas, então, não seria esperado encontrar melhor tendência adicionando-se a terceira

variável canônica. O objetivo da análise, como já comentado, é a redução da dimensão dos dados e, conseqüentemente, obtenção de facilidade na interpretação. Vale lembrar que, semelhantemente à análise de componentes principais, a análise de variáveis canônicas é uma ferramenta exploratória dos dados e não uma técnica de agrupamento. Por isso, adotou-se também esse último procedimento, na expectativa de se identificar agrupamentos mais nítidos de acessos, por meio da similaridade entre eles, estimada com base nas variáveis mensuradas.

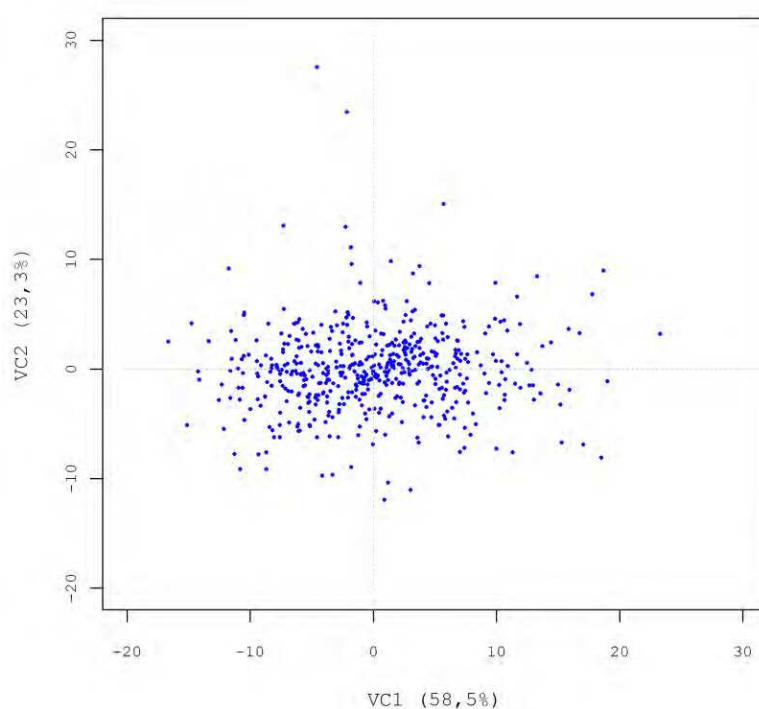


Figura 1. Dispersão de 527 acessos do germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO) nas duas primeiras variáveis canônicas (VC1 e VC2).

Na análise de agrupamento (método hierárquico aglomerativo) aplicada à matriz de distâncias de Mahalanobis, de fato, identificou-se mais claramente a formação de grupos de similaridade entre os acessos. No que se refere à representação da matriz de distâncias originais no dendrograma obtido, destacou-se, como esperado, o critério da ligação média UPGMA, com correlação cofenética igual a 0,79 ($p < 0,0001$). Os demais critérios resultaram nas seguintes correlações: vizinho mais próximo (0,54); vizinho mais distante (0,44) e método de Ward (0,44). Segundo McGarigal et al. (2000), valores desta correlação acima de 0,75 são considerados relativamente altos. Também por isso,

justificou-se a escolha do critério UPGMA como estratégia mais adequada à análise de agrupamento aplicada ao conjunto de dados desta pesquisa.

Os valores de Pseudo F, estimados para auxiliar na definição do número de grupos de similaridade, apresentaram picos mais expressivos nos estágios correspondentes à formação de dezessete e cinco grupos (Figura 2). Esses picos apontam os passos do algoritmo de agrupamento em que houve a formação de grupos com maior heterogeneidade entre si (Mingoti, 2007). É evidente que um dos problemas maiores em análise de agrupamento deve-se à decisão do número de grupos que melhor descreve o conjunto de dados. Se o algoritmo levar a uma partição que não seja a melhor para aqueles dados, implicará em inferências erradas. Assim, a decisão sobre o melhor número destes grupos tem sido assunto de algumas pesquisas (Halkidi & Vazirgiannis, 2001; Halkidi et al., 2001; Halkidi et al., 2002).

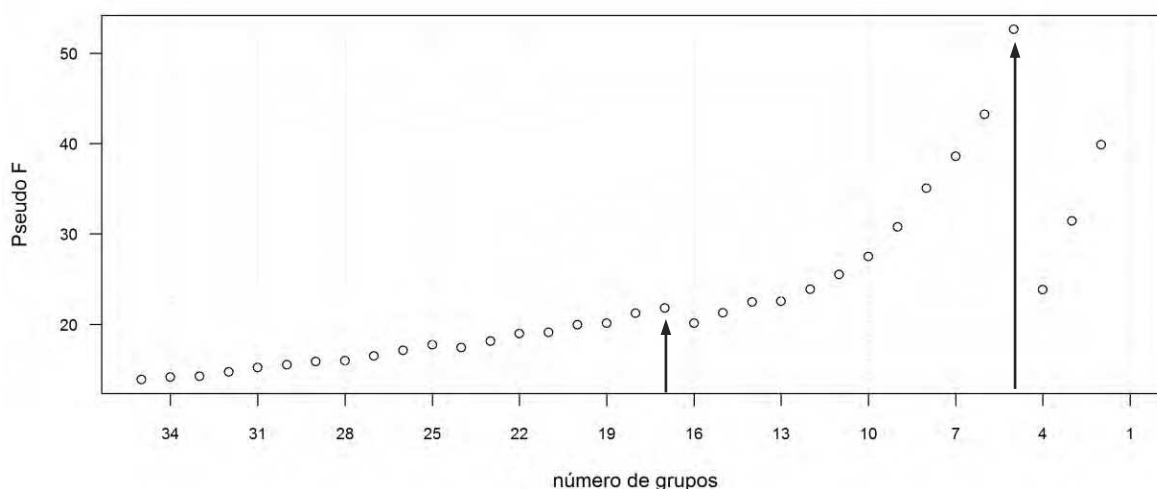


Figura 2. Representação gráfica dos valores de Pseudo F, em relação ao número de grupos formados em cada passo do algoritmo de agrupamento (análise aglomerativa hierárquica pelo critério UPGMA).

Ainda que existam critérios que auxiliem na determinação do número de grupos, essa decisão dependerá do objetivo da análise de agrupamento. Em bancos de germoplasma, a organização dos acessos em grupos de similaridade é fundamental, uma vez que a busca por genitores geneticamente contrastantes está ligada à probabilidade de se encontrar genótipos segregantes superiores na descendência de seus cruzamentos. A complementaridade nos locos que controlam determinado caráter só é possível se os

genitores apresentarem divergência nestes locos. Dessa forma, a estratificação do germoplasma em grupos de similaridade pode orientar o melhorista na escolha de genitores para hibridação, de tal forma que em suas gerações segregantes, se tenha maior possibilidade de recuperar genótipos superiores.

Para motivar a utilização do germoplasma, não é desejável que se formem muitos grupos ou estratos. O excesso deles pode dificultar a tomada de decisão dos melhoristas, na escolha de genitores. Por outro lado, ao se identificarem estratos em número muito reduzido, pode acontecer de se reunirem acessos divergentes num mesmo grupo, com conseqüente perda de informação (Milligan & Cooper, 1985). Na presente pesquisa, decidiu-se, então, pela estratificação dos acessos em poucos grupos, possibilitando, entretanto, a subdivisão destes grupos em estratos menores. Essa hierarquização da divergência entre os acessos poderá ser útil na escolha de genitores para programas de melhoramento que visem a atender demandas específicas da indústria de óleo de soja.

Com base nessa orientação e nos valores indicativos da estatística Pseudo F, os acessos do banco de germoplasma de soja da Seagro-GO foram estratificados, primeiramente, em cinco grupos (A, B, C, D e E). Estes, por sua vez, foram hierarquicamente subdivididos em outros quinze (subgrupos), sendo oito no grupo A (A1, A2 ... A8), cinco no grupo B (B1, B2 ... B5) e dois no grupo C (C1 e C2); os grupos D e E não tiveram tal subdivisão (Figura 3 e Apêndice D). A identificação numérica dos subgrupos dentro de grupos seguiu a ordem decrescente de seus tamanhos (número de acessos dentro de cada um). Como se observa, os grupos apresentam tamanhos bastante irregulares, com concentração de grande parte dos acessos nos grupos A (426 acessos) e B (92 acessos). Isto, de certa forma, revela baixa divergência genética entre a maioria dos acessos, no que tange aos caracteres ligados à qualidade do grão avaliados. Contrariamente, os grupos C (seis acessos), D (dois acessos) e E (apenas um acesso) são constituídos por poucos acessos, com características singulares nesses atributos, sendo também fundamentais na estruturação da divergência do germoplasma. Semelhantemente a essa estruturação, Iqbal et al. (2008) sugeriram dois níveis de estratificação para 139 acessos de soja do Centro Nacional de Pesquisa Agropecuária do Paquistão. Formaram-se, então, dois grupos maiores e cinco grupos menores. O agrupamento foi baseado no teor de óleo e em outras sete características associadas ao rendimento de grãos.

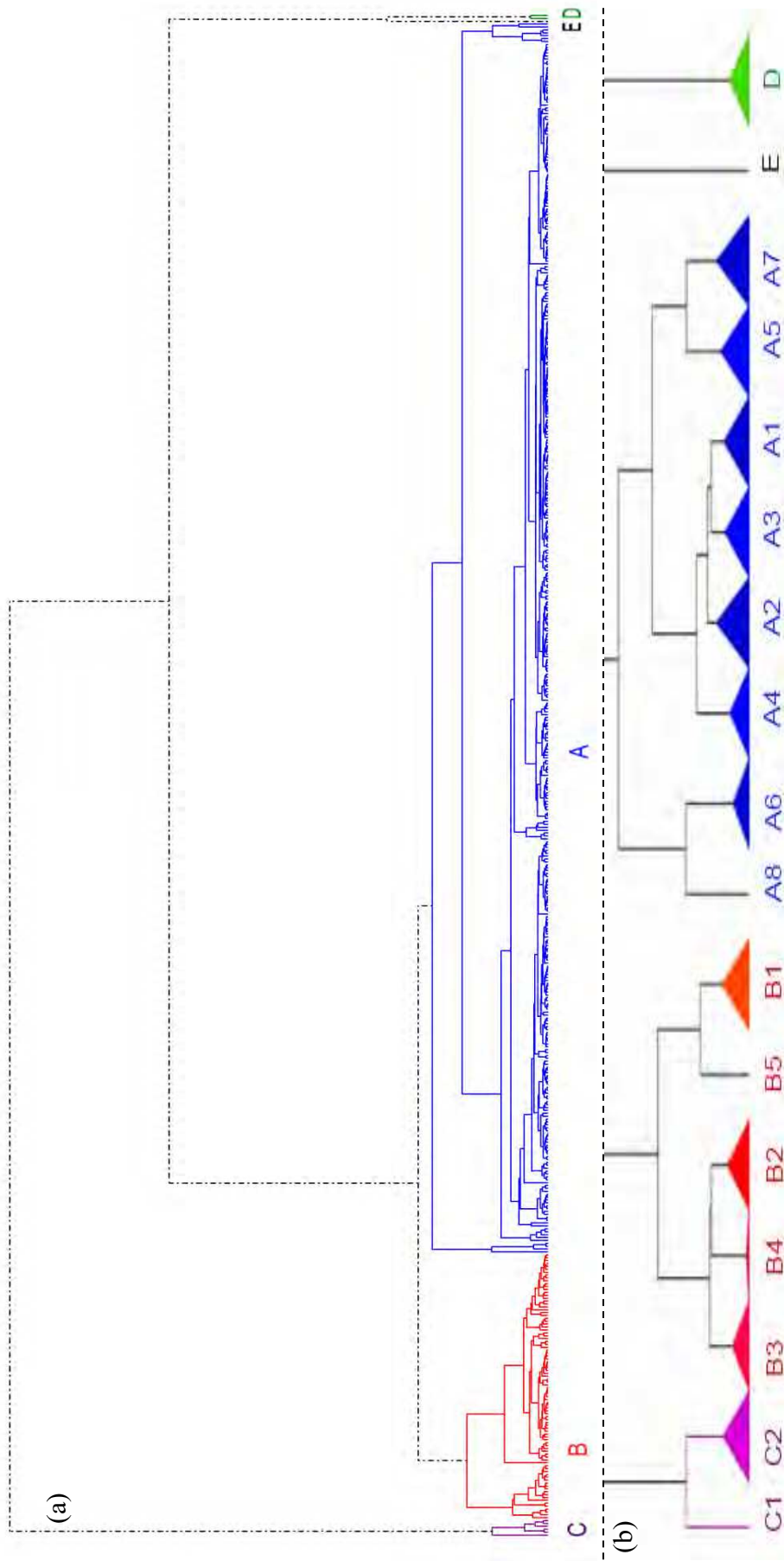


Figura 3. Dendrograma do agrupamento hierárquico aglomerativo pelo critério de UPGMA, de 527 acessos de soja do germoplasma da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em cinco estratos ou grupos de similaridade, em relação à distância de Mahalanobis, baseada em sete variáveis relacionadas à qualidade do grão (a); em (b) está representada a estratificação dos acessos em subgrupos de similaridade dentro dos estratos principais.

Nas análises de variância univariadas, assumindo os grupos como diferentes tratamentos experimentais, excetuando-se o caráter teor de óleo, verificaram-se diferenças significativas ($p < 0,01$) entre os grupos e entre os subgrupos em todos os demais caracteres (Tabela 3). Isso indica que a análise de agrupamento aglomerativo pelo critério de UPGMA permitiu boa discriminação dos acessos em grupos de similaridade, mesmo quando as variáveis são analisadas de forma independente. Evidencia também a presença de heterogeneidade entre e dentro dos grupos para maioria das variáveis avaliadas.

Tabela 3. Resultados das análises de variância univariadas, para sete variáveis relacionadas à qualidade do grão, comparando-se cinco e dezessete subgrupos de similaridade formados a partir da análise de agrupamento pelo método hierárquico, baseado no critério de UPGMA, a partir de 527 acessos de soja.

Variáveis ¹	Grupos ²		Subgrupos ³	
	QM	Valor F	QM	Valor F
OL	0,324	1,45 ^{ns}	0,330	1,04 ^{ns}
PROT	11,241	9,81**	4,430	3,95**
PAL	30,886	81,16**	10,450	34,39**
EST	1,342	16,49**	1,282	23,92**
OLC	1850,2925	219,37**	529,455	81,02**
LIN	1459,791	246,74**	449,384	131,91**
LIL	6,240	23,87**	2,217	8,98**

** : valores significativos pelo teste F, a 1% de probabilidade.

¹- OL: teor de óleo em g/100g; PROT: teor de proteína em g/100g; PAL: teor de ácido graxo palmítico em g/100g; EST: teor de ácido graxo esteárico em g/100g; OLC: teor de ácido graxo oleico em g/100g; LIN: teor de ácido graxo linoleico em g/100g e LIL: teor de ácido graxo linolênico em g/100g.

²- Grupos: A, B, C, D e E, (resíduo com 522 graus de liberdade).

³- Subgrupos: A1, A2 ... A8, B1, B2 ... B5, C1, C2, D e E, (resíduo com 510 graus de liberdade).

Com a identificação dos grupos (A, B, C, D e E) na dispersão dos acessos em duas primeiras variáveis canônicas (Figura 4), verifica-se que o agrupamento produzido faz sentido também nessa dispersão, embora as evidências mais claras da estratificação sejam prontamente visíveis apenas para os grupos D e E. Os acessos dos grupos D e E apresentaram escores com valores discrepantes em relação aos demais, o que ratifica os resultados da análise de agrupamento, que também identificou estes grupos entre os mais divergentes. Isso reforça também a necessidade do emprego de mais de uma técnica de análise estatística para uma interpretação mais segura de conjuntos de dados dessa natureza.

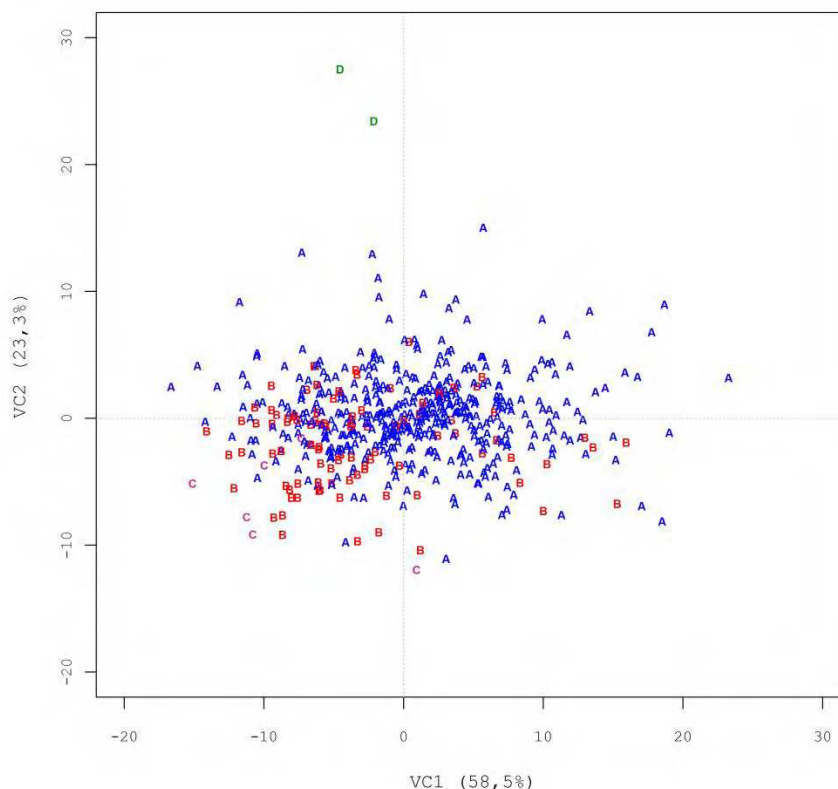


Figura 4. Representação gráfica dos escores de 527 acessos de soja do germoplasma da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em relação às duas primeiras variáveis canônicas, distribuídos em cinco grupos de similaridade estabelecidos pelo agrupamento via UPGMA, em relação a sete variáveis relacionadas à qualidade do grão.

4.3.2 Caracterização dos grupos de similaridade

O grupo A foi o que apresentou maior variabilidade para os teores de óleo e de proteína, e, ainda, os acessos com os valores extremos superiores para teor de óleo (Figura 5). Os teores médios dos ácidos oleico e poli-insaturados para esse grupo foram, respectivamente, 23,6% e 60,6% (Tabela 4). Estes valores são semelhantes aos encontrados em óleo de soja comercial (Sanibal & Mancini Filho, 2004) e representam, também, a menor média em ácido oleico e a maior em ácidos poli-insaturados dos grupos identificados. De fato, os acessos desse grupo não apresentam médias atraentes no que diz respeito à estabilidade oxidativa do óleo de soja. Entretanto, os ácidos poli-insaturados podem ser desejados na dieta, por estarem associados a efeitos benéficos à saúde. Dessa forma, esses acessos poderiam atender nichos específicos para a produção de óleo de soja,

visando a elaboração de alimentos com melhores propriedades nutricionais. O detalhamento da similaridade/divergência entre estes acessos é apresentado no Apêndice E.

Tabela 4. Resultado do agrupamento¹ de 527 acessos do banco de germoplasma de soja da Seagro-GO², em grupos e subgrupos de similaridade, com a quantidade de acessos e os valores médios estimados para cada um, segundo sete atributos relacionados à qualidade do grão³.

Grupos	Subgrupos	Nº acessos	PROT	OL	PAL	EST	OLC	LIN	LIL	SAT	POLI
A	A1	269	38,9	16,7	12,0	2,2	23,0	54,9	6,5	14,2	61,5
	A2	134	38,9	16,7	12,0	2,2	25,3	52,3	6,6	14,1	58,9
	A3	7	38,6	16,9	12,5	3,4	20,9	54,1	6,0	15,9	60,0
	A4	5	37,6	16,7	10,9	2,0	21,4	57,7	6,9	12,9	64,6
	A5	5	39,0	16,2	13,9	2,2	24,8	50,7	6,4	16,0	57,1
	A6	3	40,4	16,7	11,7	2,1	20,2	58,4	6,3	13,8	64,8
	A7	2	38,5	17,1	13,7	3,2	26,4	48,0	5,9	16,9	53,9
	A8	1	40,4	17,0	10,0	2,1	19,2	62,1	6,2	12,1	68,2
	Geral	426	38,9	16,7	12,0	2,2	23,6	54,1	6,5	14,2	60,6
B	B1	72	39,5	16,7	11,4	2,0	30,8	47,9	6,2	13,4	54,1
	B2	13	40,2	16,7	10,8	1,8	36,5	43,1	6,0	12,6	49,0
	B3	4	39,2	16,3	11,8	1,8	34,5	42,7	7,0	13,5	49,7
	B4	2	38,8	16,2	12,3	2,0	34,3	43,3	6,2	14,3	49,5
	B5	1	39,1	17,4	11,3	3,2	34,4	43,9	4,8	14,5	48,7
	Geral	92	39,6	16,7	11,4	2,0	31,9	46,9	6,2	13,3	53,0
C	C1	5	39,8	16,9	10,1	1,7	43,3	37,4	5,5	11,9	42,9
	C2	1	41,1	16,6	10,3	1,4	48,9	31,3	5,2	11,7	36,5
	Geral	6	40,1	16,8	10,2	1,7	44,2	36,3	5,5	11,8	41,8
D	-	2	38,4	16,2	17,0	2,2	20,4	51,2	7,0	19,2	58,2
E	-	1	37,9	17,5	16,6	2,9	32,4	35,5	3,6	19,5	39,1

¹- Método aglomerativo hierárquico pelo critério UPGMA, baseado na distância de Mahalanobis.

²- Seagro-GO: Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás.

³- OL: teor de óleo em g/100g; PROT: teor de proteína em g/100g; PAL: teor de ácido graxo palmítico em g/100g; EST: teor de ácido graxo esteárico em g/100g; OLC: teor de ácido graxo oleico em g/100g; LIN: teor de ácido graxo linoleico em g/100g e LIL: teor de ácido graxo linolênico em g/100g; SAT: teor de ácidos saturados (PAL+EST) em g/100g; POLI: teor de ácidos poli-insaturados (LIN+LIL) em g/100g.

O grupo B apresentou variabilidade comparável ao grupo A em teores de óleo e de proteína (Figura 5). O teor médio de óleo nesse grupo B foi igual ao do grupo A (16,7%), porém, com menor amplitude (Tabela 4). Já em teor de proteína, o grupo B apresenta média levemente superior, e contém os acessos com os valores extremos superiores neste atributo (acima de 42%). Além dos acessos com os maiores teores proteicos, o grupo B tem vantagem sobre o grupo A, por apresentar maior teor médio de

ácido oleico (31,9%) e menor teor dos ácidos poli-insaturados (53,0%) (Figura 5 e Tabela 4). Isto é, em média, os acessos do grupo B apresentam características mais atraentes à qualidade do óleo de soja do que o grupo A. A similaridade ou divergência entre os acessos deste grupo é detalhada no Apêndice F.

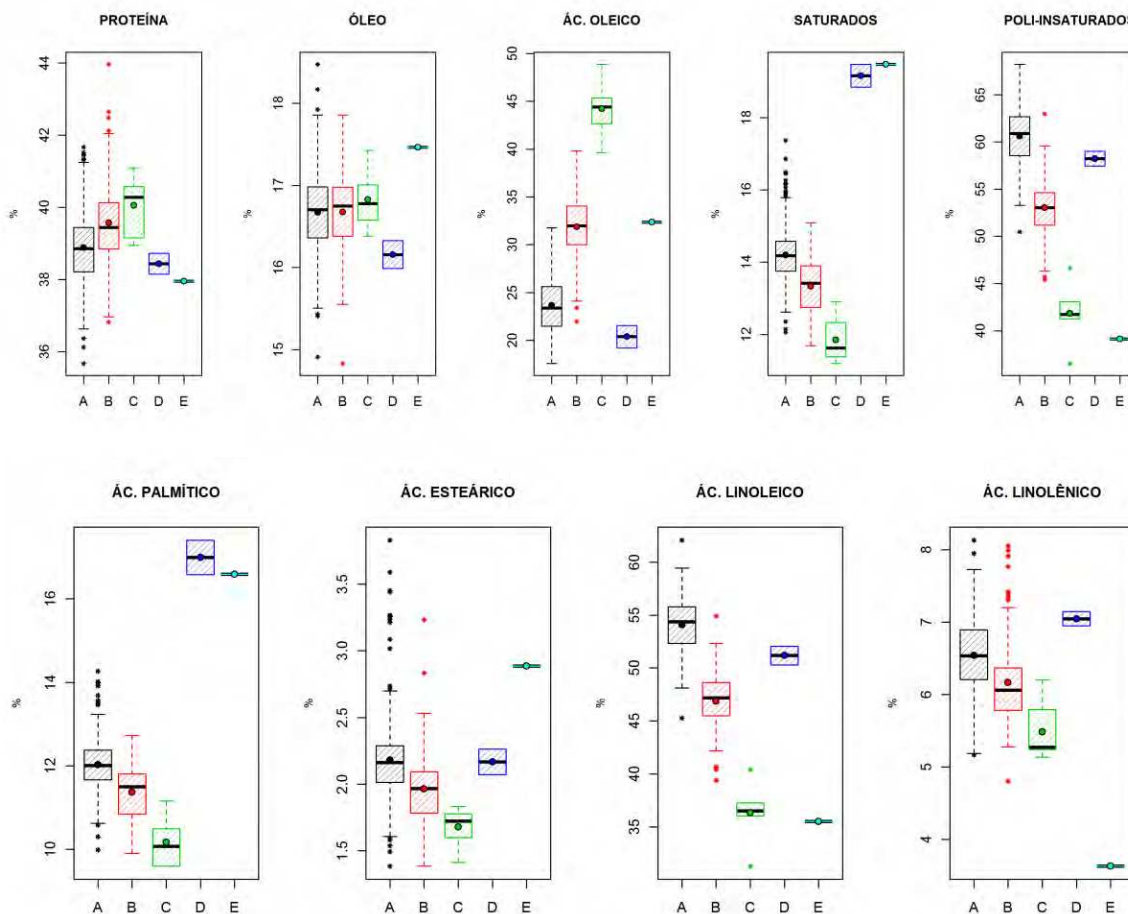


Figura 5. Boxplots de cinco grupos de similaridade obtidos por análise de agrupamento (UPGMA) de 527 acessos do banco de germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em relação aos teores percentuais de óleo e proteína no grão, e dos ácidos oleico, saturados, poli-insaturados, palmítico, esteárico, linoleico e linolênico no óleo.

O grupo C destacou-se, entre todos os grupos, por apresentar o melhor perfil em relação à estabilidade oxidativa do óleo de soja. Isto se deve principalmente aos maiores teores de ácido oleico, com média de 44,2%, e baixos teores de ácidos poli-insaturados, com média de 41,8% (Figura 5 e Tabela 4). Deve ser dada atenção ao fato de que, na subcoleção avaliada, menos de dez acessos apresentaram teores de ácido oleico iguais ou superiores a 40%; entre estes, estão os seis acessos que compõem o grupo C (Apêndice A). Além dessas características, vale destacar que o grupo C apresenta também

a maior média para teor de proteína (40,1%). As divergências entre acessos dentro deste grupo (C), assim como nos grupos D e E, podem ser inspecionadas no Apêndice G.

O grupo D ficou constituído por apenas dois acessos (Bag 51 e Bag 588), que são caracterizados pelos altos teores de ácidos graxos saturados, principalmente ácido palmítico, além de baixos teores de ácido oleico (Figura 5). O teor médio em ácidos saturados foi de 19,2%, um dos maiores entre os grupos, e em ácido oleico, 20,4%, o menor deles. O acesso Bag 588 (FT Manacá) é uma cultivar e o Bag 51 (GOBR94-09543-00.02), uma linhagem desenvolvida pelo programa de melhoramento de soja da Seagro-GO e seus parceiros. A cultivar FT Manacá, com o segundo maior teor de ácido palmítico (16,6%), é um dos progenitores da linhagem GOBR94-09543-00.02 (genealogias no Apêndice A), que apresentou o mais alto teor observado para este ácido graxo (17,4%). Isto pode sugerir que o caráter tem herança relativamente simples, o que pode ser reforçado pela sua elevada estimativa de herdabilidade, 82% (Tabela 2).

O grupo E ficou representado por apenas um acesso (Bag 278). Este é também uma linhagem desenvolvida pelo programa de melhoramento de soja da Seagro-GO e seus parceiros, cuja identificação é BRGO99-4095-05. A linhagem, de fato, apresentou-se bem distante dos demais acessos na dispersão de variáveis canônicas (Figura 5). Sua divergência em relação aos demais acessos deve-se, principalmente, ao seu alto teor de ácidos graxos saturados (19,5%), associado ao teor muito baixo dos ácidos poli-insaturados (39,1%), com destaque para o ácido linolênico (Figura 5 e Tabela 4). No geral, este acesso apresentou característica de alta estabilidade no óleo, associada à maior média observada entre os grupos, em conteúdo de óleo no grão (17,5%). No entanto, o teor expressivo de ácidos saturados não é desejável, tanto para a indústria alimentícia, quanto para a indústria de biodiesel. Para a produção de alimentos, altos níveis de saturação no óleo têm sido relacionados com altos índices de colesterol no sangue e com problemas cardiovasculares. Na produção de biodiesel, o alto nível de ácidos graxos saturados pode trazer complicações ao motor, devido ao aumento do ponto de névoa e de congelamento do combustível (Knothe, 2010). Logo, a linhagem tem potencial, principalmente, como genitor em programas de cruzamento para transferir a característica de baixo teor de ácidos graxos poli-insaturados aos genótipos segregantes. O teor de óleo no grão deste acesso foi acima da média dos outros grupos; no entanto, seu teor médio de proteína foi o menor. Considerando-se a correlação negativa entre esses dois caracteres (Bonato et al., 2000; Feng et al., 2004), e a possível subestimação dos teores de óleo neste estudo, o acesso

ainda poderia ter potencial como genitor para melhoramento da soja em teor de óleo no grão.

Uma vez que alguns grupos foram organizados em subgrupos de similaridade, foi necessário também caracterizá-los. O grupo A ficou dividido em oito subgrupos, denominados A1 (269 acessos), A2 (134 acessos), A3 (sete acessos), A4 (cinco acessos), A5 (cinco acessos), A6 (três acessos), A7 (dois acessos) e A8 (um acesso) (Figura 6 e Tabela 4). A maioria dos acessos dos subgrupos A1 e A2 e todos de A5 e A7 apresentaram teores de ácido oleico superiores à média de cultivares comerciais, 21,5% (Sanibal & Mancini Filho, 2004) (Figura 6). Os teores médios de óleo e de proteína foram semelhantes para esses subgrupos, porém, devem-se destacar os subgrupos A1 e A2, por incluírem os acessos com os maiores valores para essas variáveis. O subgrupo com maior teor médio de ácido oleico foi A7 (26,4%) e, na sequência, A2 (25,3%), A5 (24,8%) e A1 (23,0%). Vale ressaltar, entretanto, que os subgrupos A1 e A2, juntos, contêm cerca de vinte acessos com valores acima daqueles encontrados no subgrupo A7. No que tange a outra característica importante para a qualidade do óleo, o teor de ácidos poli-insaturados, os acessos dos subgrupos A7 e A5, e grande parte daqueles de A2, apresentam teores inferiores aos encontrados em cultivares comerciais, 59,9% (Sanibal & Mancini Filho, 2004). Assim, em geral, os acessos dos subgrupos A7 e A5 apresentam melhor perfil na qualidade do óleo.

Os subgrupos A4, A6 e A8 destacam-se no grupo A por apresentarem os maiores teores de ácidos graxos poli-insaturados e os menores de ácidos saturados. O destaque se confirma, comparando-se esses teores às respectivas médias de cultivares comerciais, 59,9% e 14,1% (Sanibal & Mancini Filho, 2004). Entre eles, destaca-se o subgrupo A8 (Bag 586), com 12,1% de ácidos saturados e 68,2% de poli-insaturados. Este acesso (cultivar UFV-3), como se sabe foi desenvolvido principalmente para a região Norte de Minas Gerais (Sediyama et al., 1980). A cultivar ainda apresentou teores de óleo e de proteína acima da média do grupo A. Por suas características, incluindo-se o fato de já ser um material comercial, essa cultivar talvez pudesse ser indicada para a obtenção de óleo de soja com fins nutracêuticos. Conforme reporta a literatura, o ácido linoleico, além de ser essencial ao organismo humano, quando consumido em níveis maiores que os ácidos saturados, pode reduzir os riscos de doenças cardiovasculares (Simopoulos, 2008; Katan, 2009; Deckelbaum, 2010; Farzaneh-Far et al., 2010). Como já comentado, o ácido linoleico é o principal representante, no óleo de soja, do grupo de ácidos poli-insaturados do tipo ômega-6. Deve-se atentar, entretanto, conforme advertência da Anvisa (Agência

Nacional de Vigilância Sanitária), que o ômega-6, diferentemente do ômega-3, mesmo sendo um constituinte que apresenta benefícios à saúde, não é um alimento funcional; pois foi excluído da lista das alegações de propriedade funcional aprovadas, em sua última atualização (Brasil, 2005).

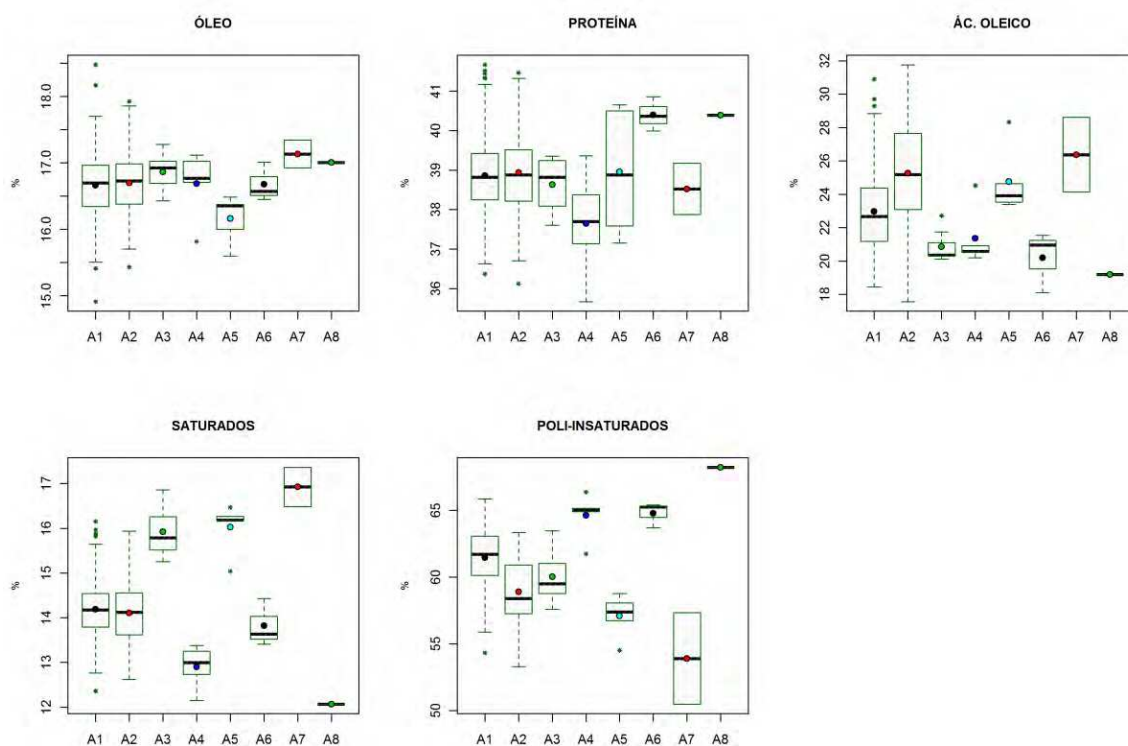


Figura 6. Boxplots de subgrupos de similaridade (A1 a A8) resultantes da divisão do grupo A (426 acessos do banco de germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás / Seagro-GO), em relação aos teores de óleo, proteína, ácido oleico e ao total de ácidos graxos saturados e poli-insaturados no grão.

No grupo B foram identificados cinco subgrupos de similaridade, e no C, dois subgrupos (Figura 7). Uma característica comum aos dois grupos é que todos os seus acessos apresentam teores de ácido oleico superiores à média de cultivares comerciais, 21,5% (Sanibal & Mancini Filho, 2004). O subgrupo B1, com 72 entradas, destaca-se por conter acessos com os maiores teores de proteína da coleção, embora sua média tenha sido a segunda maior no grupo B (Tabela 4). Os outros subgrupos, B2 a B5, apresentam características semelhantes para os ácidos oleico e poli-insaturados; a principal diferença entre eles é devido aos teores de ácidos graxos saturados (Figura 7). Entre esses subgrupos, B2, com treze entradas, apresenta menor média de ácidos graxos saturados (12,6%), caracterizando-se, portanto, como o de melhor perfil para a indústria de alimentos.

O grupo C ficou caracterizado por apresentar perfil favorável à estabilidade no óleo de soja, pelos menores índices de ácidos poli-insaturados e, simultaneamente, maiores índices de ácido oleico (Figura 7 e Tabela 4). A formação dos dois subgrupos de similaridade permitiu discriminar o Bag 553 (subgrupo C2), dos outros cinco acessos (subgrupo C1). O Bag 553, cultivar Sharkey, destacou-se como o de melhor perfil para produção de óleo com boa estabilidade oxidativa. O teor médio de ácido oleico desta cultivar foi 48,9%, e o de ácidos poli-insaturados, 36,5%. Além disso, esse acesso apresenta elevado teor médio de proteína (41,1%), tendo apresentado a maior média no grupo. No subgrupo C1, o acesso com melhor estabilidade de óleo seria o Bag 517 (PI-90406), com 45,4% de ácido oleico e 41,3% de ácidos poli-insaturados. Ambos os acessos apresentaram quase o dobro de ácido oleico em relação ao teor médio de cultivares comerciais, 21,5% (Sanibal & Mancini Filho, 2004).

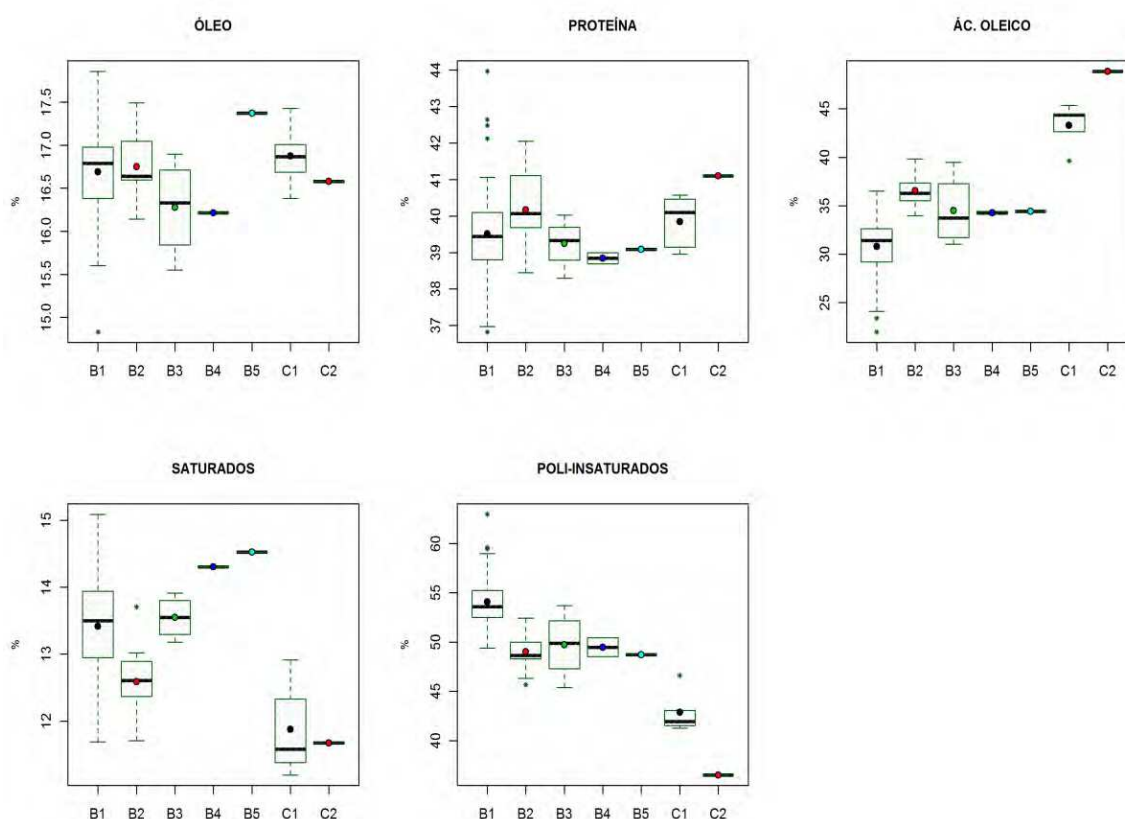


Figura 7. Boxplots de subgrupos de similaridade (B1 a B5, C1 e C2) resultantes da subdivisão dos grupos B (92 acessos) e C (seis acessos), do germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em relação aos teores de óleo, proteína, ácido oleico e do total de ácidos graxos saturados e poli-insaturados no grão.

Em síntese, os acessos do grupo C apresentam melhor perfil para qualidade de óleo (estabilidade oxidativa) e, ainda, certa variabilidade para os caracteres teor de óleo e teor de proteína. Logo, é possível indicar acessos com perfis ideais de qualidade de óleo, aliado a desempenho relativamente bom para teor de proteína, como, por exemplo, o Bag 553 (cultivar Sharkey). Aliando-se médias superiores em teor de óleo, pode-se acrescentar os acessos Bag 2591 (GOBR03-2397-13) e Bag 517 (PI-90406). Esses acessos representam, portanto, genitores com potencial de uso em programas de melhoramento de soja, visando cultivares com melhor qualidade industrial de óleo.

Com a categorização dos níveis de cada caráter (baixo, médio e alto), associada ao dendrograma (Tabela 5 e Figura 8), primeiramente confirma-se a nítida diferenciação entre os grupos nos teores de ácidos oleico e poli-insaturados; além da diferenciação menor entre eles nos teores de óleo, proteína e ácidos graxos saturados. Também observa-se que, excetuando-se o ciclo ou grupo de maturação, nas outras cinco variáveis consideradas nessa representação, os grupos mostram-se homogêneos internamente e com padrões de variabilidade relativamente semelhantes. Ademais, foi possível constatar que altos níveis de óleo, proteína e ácido oleico e baixos níveis de ácidos poli-insaturados são características restritas a poucos acessos no germoplasma. Ainda, observa-se que todos os grupos têm, pelo menos, um acesso com altos teores de óleo. Para índices mais elevados de óleo e proteína, simultaneamente, destacam-se Bag 491 (PI-262180), do subgrupo A2, e Bag 318 (Santa Rosa) e Bag 527 (BRS 217 - Flora), do subgrupo A1. Cultivares com altos teores de proteína têm sido preferidas pelos produtores de extrato de soja e de queijo tofu (Sediyama et al., 2005). Paralelamente a essa demanda, o melhoramento genético certamente deverá também persistir na busca de genótipos com alto teor de óleo no grão. Dessa forma, esses acessos, com destaque para Bag 491 (PI-262180), apresentam potencial genético para serem explorados na obtenção de genótipos superiores para ambos os objetivos.

Tabela 5. Limites inferiores e superiores (g/100g) utilizados para a categorização dos acessos em três níveis de conteúdo (com classes equidistantes), baixo, médio e alto, para sete atributos relacionados à qualidade do grão de soja.

Variáveis	Baixo	Médio	Alto
Óleo	14,8 - 16,0	16,1 - 17,3	17,4 - 18,5
Proteína	35,7 - 38,4	38,5 - 41,2	41,3 - 44,0
Ácido oleico	17,6 - 28,0	28,1 - 38,4	38,5 - 48,9
Ácidos saturados	11,2 - 14,0	14,1 - 16,8	16,9 - 19,5
Ácidos poli-insaturados	36,5 - 47,1	47,2 - 57,6	57,7 - 68,2

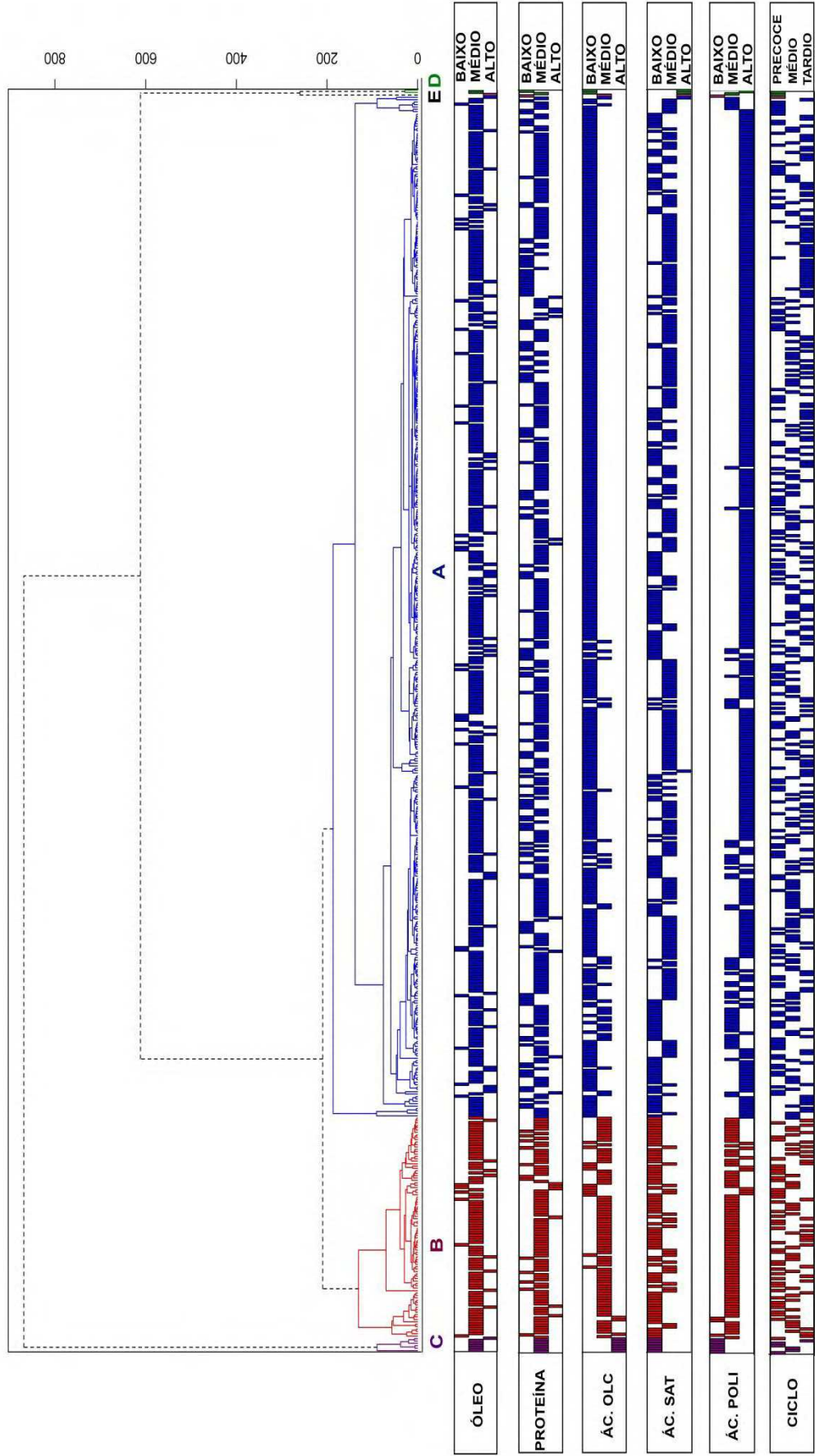


Figura 8. Dendrograma do agrupamento de 527 acessos do banco de germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em relação aos níveis de: óleo; proteína; ácidos graxos oleico (OLC), saturados (SAT), poli-insaturados; e ao ciclo de maturação (cada um categorizado em três classes). Os acessos foram divididos em cinco grupos de similaridade, os quais estão representados por diferentes cores.

4.3.3 Divergência entre acessos com óleo de melhor perfil em estabilidade oxidativa

A estabilidade oxidativa do óleo está relacionada aos baixos índices de insaturações na cadeia dos ácidos graxos (Knothe, 2010). A demanda por óleos com este perfil está relacionada à obtenção de produtos com maior vida útil e menor degradação por oxidação. Para que o óleo de soja apresente níveis reduzidos de ácidos graxos poli-insaturados, espera-se que, em contrapartida, os níveis de ácidos graxos monoinsaturados aumentem. Isto porque não é desejável, tanto para a indústria alimentícia, quanto para a indústria de biodiesel, que se aumentem os índices de ácidos saturados em função da redução dos ácidos graxos poli-insaturados.

Com intuito de avaliar a estruturação da diversidade somente entre acessos com melhor perfil de estabilidade oxidativa, selecionaram-se acessos com reduzidos índices de ácidos poli-insaturados (<60%) e, dentro destes, aqueles com os maiores índices (>30%) de ácidos monoinsaturados (ácido oleico). Isto resultou num total de 82 acessos (Tabela 6). Conforme já descrito, a referência para o teor máximo de poli-insaturados foi a média de cultivares comerciais, 59,9% (Sanibal & Mancini Filho, 2004), e o valor mínimo para o ácido oleico foi tomado com base nos índices apresentados por linhagens descritas como fonte de características de interesse associadas a este ácido, por exemplo, menor susceptibilidade à oxidação durante o processo de refino, armazenamento e fritura (Rahman et al., 1994; Rahman et al., 1996; Rahman et al., 2001; Burton et al., 2006).

A análise de agrupamento via UPGMA permitiu discriminar os 82 acessos, em cinco grupos de similaridade, identificados de G1 a G5 (Figura 9 e Tabela 6). A correlação cofenética (0,86) foi alta e significativa ($p < 0,0001$), revelando boa representação da matriz de distâncias originais no dendrograma. A representação dos grupos de similaridade na dispersão de variáveis canônicas (Figura 10) mostra concordância muito boa entre esta técnica e a análise de agrupamento (Figura 9). É possível notar que dois grupos mostraram-se mais divergentes, ambos com um acesso cada: G1, constituído pelo acesso Bag 278 (BRGO99-4095-05), e G2, pelo acesso Bag 553 (cultivar Sharkey). Isto já era esperado, pois no agrupamento geral (Figura 3) estes dois acessos já ficaram isolados dos demais. Naquela análise (Tabela 4), o Bag 553 foi o único representante do subgrupo C2 e o Bag 278, o único do grupo E.

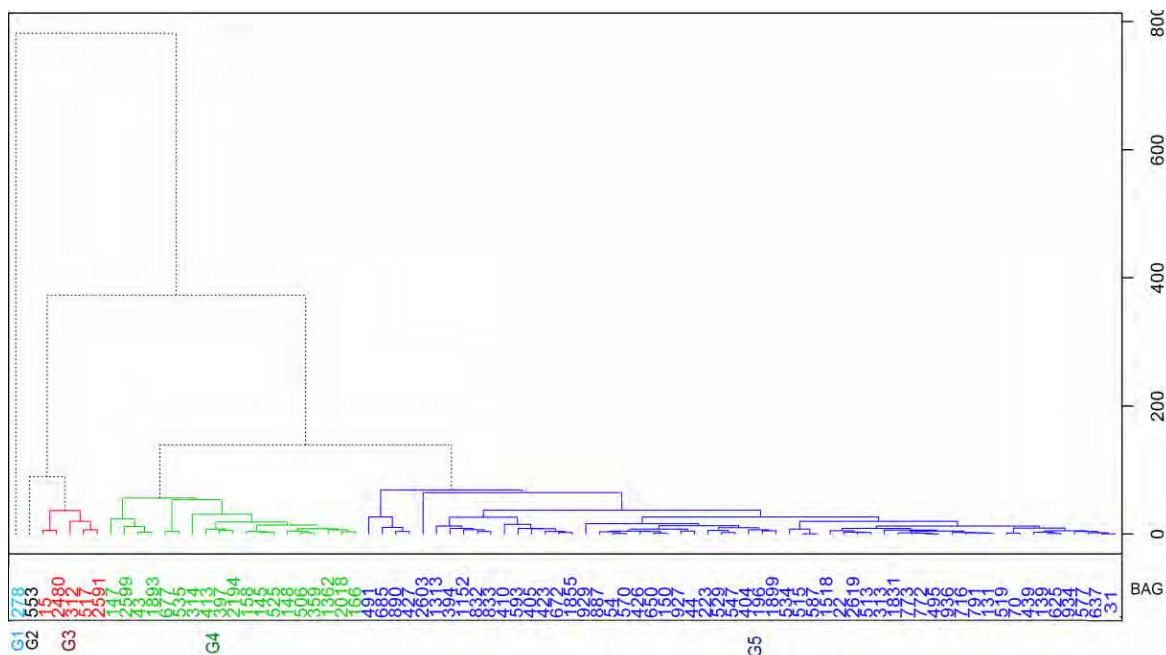


Figura 9. Dendrograma do agrupamento de 82 acessos, com óleo de melhor perfil em estabilidade oxidativa, mantidos no banco de germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em relação à distância de Mahalanobis baseada em sete atributos relacionados à qualidade do grão.

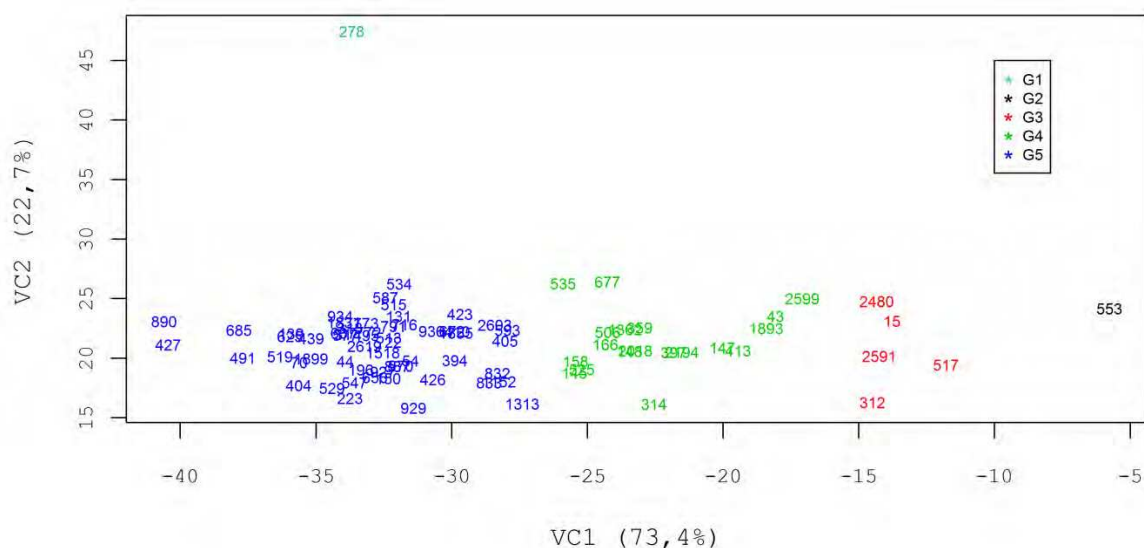


Figura 10. Representação gráfica dos escores de 82 acessos do banco de germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em relação às duas primeiras variáveis canônicas, distribuídos em cinco grupos de similaridade (G1 a G5) estabelecidos por agrupamento UPGMA, em relação a sete atributos relacionados à qualidade do grão, com número Bag de identificação.

Tabela 6. Relação de 82 acessos de soja de melhor perfil em estabilidade oxidativa, pertencentes à subcoleção mantida pela Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), com informações¹ ciclo de maturação (CM) e teores médios (g/100 g) de proteína (PROT), óleo (OL) e dos ácidos graxos palmítico (PAL), esteárico (EST), oleico (OLC), linoleico (LIN) e linolênico (LIL), com a respectiva classificação em grupos de similaridade (G).

Bag ¹	Acesso	CM ²	PROT	OL	PAL	EST	OLC	LIN	LIL	G ³
15	CSM-1	M	16,7	40,5	10,5	1,8	44,5	36,2	5,3	G3
22	PINTADO	T	17,1	39,4	11,6	1,7	31,6	47,8	6,3	G5
31	GOBR97-058163	T	16,8	39,7	11,7	1,8	31,1	48,2	6,1	G5
43	BR95-15305	T	16,9	39,3	11,9	2,0	35,1	42,4	6,8	G4
44	BR92-249	T	16,9	38,5	11,2	2,1	30,7	49,1	6,1	G5
54	GOBR93-1483 (LUZIÂNIA)	M	16,8	39,7	10,9	2,0	33,1	47,2	5,9	G5
70	GOBR97-057130	M	17,0	39,8	11,3	1,8	30,3	49,7	6,2	G5
131	GOBR96-013519	P	15,6	41,0	11,7	2,1	31,4	45,8	5,9	G5
139	GOBR97-056191 (IPAMERI)	T	16,3	40,1	11,6	1,9	30,8	48,5	6,0	G5
145	CENTENNIAL	P	16,4	39,7	10,3	1,8	36,3	43,8	5,8	G4
147	PI-90763	P	15,6	40,0	11,7	1,5	31,0	45,7	8,0	G4
148	PICKET	P	17,0	41,2	11,0	1,5	35,6	43,5	6,4	G4
150	LEE	P	17,3	39,8	10,8	1,6	31,9	48,0	6,2	G5
158	INA	P	16,5	38,4	10,8	1,8	34,0	44,6	6,2	G4
166	GOBR99-4102	P	16,6	38,8	10,9	2,1	36,3	43,0	5,7	G4
196	GO 00-366-01-01	M	17,1	40,0	10,7	2,2	32,6	47,4	5,5	G5
223	GOBR98-127058	T	16,8	39,8	10,4	1,6	31,8	48,8	6,1	G5
278	BRGO99-4095-05	P	17,5	37,9	16,6	2,9	32,4	35,5	3,6	G1
312	PI-200490	P	16,9	40,1	9,6	1,8	39,6	40,4	6,2	G3
313	PI-240667	T	16,0	42,5	11,7	1,9	31,4	47,8	6,2	G5
314	PI-94159 FRF	M	16,1	40,1	10,1	1,6	34,7	45,6	6,8	G4
359	TAINUNG 4 (PI 368039)	P	16,6	41,1	11,3	1,7	36,4	42,3	6,1	G4
394	GOBR97-14346 (ARAÇU)	P	16,2	38,0	10,9	2,0	31,8	46,5	6,0	G5
397	GOBR97-14723	P	17,2	39,7	10,3	1,7	39,7	40,4	5,3	G4
404	GOBR97-08728	T	16,8	40,3	10,6	1,9	30,9	49,0	5,8	G5
405	GOBR98-085064	T	16,7	40,0	10,9	1,8	36,5	43,9	5,5	G5
410	GOBR99-143010	M	16,7	42,1	11,7	2,1	31,5	46,5	6,3	G5
413	GOBR99-971004	M	16,7	40,2	10,5	1,9	39,8	40,7	5,7	G4
423	GOBR99-961001	M	16,2	39,0	11,8	2,0	32,9	45,5	6,0	G5
426	GOBR99-945001	T	17,2	39,5	10,8	1,8	32,3	47,5	6,2	G5
427	GOBR99-671019	T	17,0	39,6	11,4	1,8	30,5	49,7	5,5	G5
439	GOBR99-671061	T	17,0	40,1	11,2	1,8	33,9	47,0	5,4	G5

(Continua)

Tabela 6. Continuação.

Bag ¹	Acesso	CM ²	PROT	OL	PAL	EST	OLC	LIN	LIL	G ³
491	PI-262180	P	17,4	41,5	11,0	2,2	31,8	48,1	5,2	G5
495	PI-200492	P	16,6	39,8	11,5	2,0	31,4	47,2	5,9	G5
506	BRS ROSA	P	16,6	40,5	11,1	1,7	36,6	42,4	5,9	G4
513	FT-LIDER	P	17,1	38,8	11,7	2,1	30,5	47,5	6,1	G5
515	BRS ANHUMAS	T	17,1	38,8	11,9	2,0	33,3	45,5	5,5	G5
517	PI-90406	P	17,0	39,0	9,6	1,6	45,4	36,0	5,3	G3
519	BRSMS TAQUARI	T	17,3	38,9	11,3	1,8	30,8	49,2	5,9	G5
525	CD 207	M	17,5	39,2	10,7	2,0	35,5	44,9	6,0	G4
529	CD 206	P	16,8	38,6	10,3	2,0	32,1	48,0	5,5	G5
534	BRQ 94-7277	P	16,2	41,1	12,6	2,5	30,5	46,5	6,1	G5
535	BRQ 95-1873	T	16,2	39,0	12,4	1,9	34,1	44,1	6,4	G4
547	BED FORD	P	16,6	38,0	10,6	2,1	30,8	48,6	5,8	G5
553	SHARKEY	P	16,6	41,1	10,3	1,4	48,9	31,3	5,2	G2
570	FT 100	M	16,6	39,3	11,0	2,0	32,0	48,2	6,2	G5
577	DM RAINHA	M	15,9	39,1	11,3	1,8	33,3	47,2	5,8	G5
587	FT EUREKA	P	16,6	38,9	12,1	2,2	31,8	45,8	5,7	G5
593	CD 202	P	16,9	37,0	11,4	2,0	34,0	44,5	5,7	G5
625	BR 4 RC	M	16,7	39,2	11,7	1,8	30,2	48,9	6,1	G5
637	CAJEME	P	16,8	39,5	11,6	1,8	31,5	47,4	5,9	G5
650	UFV 14	T	16,9	39,7	10,7	1,9	32,5	48,1	5,9	G5
672	CORDELL	P	16,5	39,4	11,1	2,0	35,0	44,4	5,4	G5
677	BRQ 95-1136	M	16,2	38,7	12,3	2,1	34,4	42,5	6,1	G4
685	MONSOY 7603	M	16,4	38,9	11,7	1,9	30,4	49,2	5,8	G5
716	J 200	M	16,8	39,9	11,6	1,9	33,1	46,3	5,8	G5
772	PRATA	M	16,5	40,5	11,4	2,0	32,5	46,7	5,7	G5
773	EMBRAPA 34 (TEREZINA)	M	16,8	40,5	11,5	2,0	33,4	45,9	5,5	G5
791	PI-200490	T	17,0	41,0	11,7	1,9	32,6	46,4	5,9	G5
832	GOBR00-0466-013	M	16,9	39,1	10,3	1,9	36,4	45,1	5,5	G5
833	GOBR00-0464-048	P	17,4	39,1	10,5	1,4	34,4	46,0	6,0	G5
887	BRGO99-4313-03	T	16,4	38,3	11,0	1,8	32,0	48,1	6,2	G5
890	GOBR00-345-028	T	16,4	39,8	11,7	1,8	30,9	48,9	5,4	G5
927	BRGO99-409659-03-01	T	16,6	37,7	10,7	2,3	32,5	47,9	5,6	G5
929	BRGO99-4101-51-03-01	T	17,5	39,4	9,9	1,8	35,1	47,3	5,6	G5
934	GOBR01-0978	P	16,2	39,0	11,8	2,2	31,4	47,3	5,8	G5
936	GO 02-2118	M	16,6	40,8	11,5	2,0	33,2	46,1	5,9	G5
1152	GOBR03-2866-30	P	17,6	38,3	10,8	2,1	31,9	46,5	6,0	G5
1313	BRB02-854-03-01	M	16,2	40,3	10,1	1,7	33,2	47,0	6,5	G5

(Continua)

Tabela 6. Continuação.

Bag ¹	Acesso	CM ²	PROT	OL	PAL	EST	OLC	LIN	LIL	G ³
1362	BRB02-977-03-01	M	16,6	41,1	11,5	2,2	34,5	43,8	6,2	G4
1518	BRB01-122/123-03-01	M	16,8	39,6	10,9	1,7	34,1	46,2	5,6	G5
1831	GOBR99-08-018	P	16,3	40,2	11,7	2,3	30,9	47,2	5,7	G5
1855	BRGO01-01833	T	16,8	38,7	11,5	2,1	32,8	46,0	5,9	G5
1893	BR 00-923	M	16,5	39,4	11,9	1,8	32,4	43,3	7,3	G4
1899	GO 02-2052-03-33	T	17,1	40,7	11,0	2,0	32,2	48,0	5,6	G5
2018	GOBR00-665.14.04.01	M	17,1	39,9	10,7	2,2	37,4	42,7	5,6	G4
2194	GOBR02-2203-1	M	16,7	42,0	10,7	1,7	38,3	42,2	6,1	G4
2480	BRGO02-2172-3	M	16,4	40,6	11,2	1,8	42,6	37,3	5,8	G3
2591	GOBR03-2397-13	T	17,4	39,1	9,9	1,7	44,4	36,8	5,1	G3
2599	GOBR03-2682-1	T	16,1	38,3	11,5	1,9	39,5	39,4	6,0	G4
2603	GOBR03-3061-22	T	17,4	39,1	11,3	3,2	34,4	43,9	4,8	G5
2619	GOBR03-2668-55	M	17,5	37,9	11,5	1,8	31,2	47,7	5,9	G5

¹- Bag: número usado como código de cada acesso no banco de germoplasma.

²- CM (ciclo de maturação): com classificação correspondente aos ciclos precoce (P), médio (M) e tardio (T).

³- Os grupos de similaridade foram obtidos por análise de agrupamento hierárquica e aglomerativa via critério de ligação UPGMA, tendo resultado em cinco grupos (G1, G2, G3, G4 e G5).

A cultivar Sharkey (Bag 553) apresenta melhor desempenho para estabilidade oxidativa, pois resultou no maior teor de ácido oleico (48,9%) e no menor teor de ácidos poli-insaturado (36,5%). Neste atributo, o acesso é seguido da linhagem BRGO99-4095-05 (Bag 278), com 39,1%, a qual, no entanto, apresenta um dos menores índices de ácido oleico do conjunto (32,4%), apesar do maior teor de óleo (Figura 11). Outra diferença entre esses acessos está no alto índice de ácidos saturados (19,5%) desta linhagem, porém, esta característica não a prejudica em relação à estabilidade oxidativa. Altos níveis de ácidos graxos saturados no óleo são indesejáveis, tanto na indústria de biodiesel, por complicações causadas ao motor (Knothe, 2010), quanto na indústria de alimentos, por problemas relacionados à saúde (Tyson & McCormick, 2006).

Os outros oitenta acessos foram discriminados em três grupos de similaridade: G3 (cinco acessos), G4 (dezenove acessos) e G5 (56 acessos) (Figura 9 e Figura 10). Os grupos G3, G4 e G5 ordenam-se decrescentemente quanto à estabilidade oxidativa do óleo; isto é, com índices cada vez menores de ácido oleico e índices crescentes de ácidos poli-insaturados (Figura 11). Os teores de óleo e de proteína, também nesse subconjunto do germoplasma, discriminaram pouco os grupos de similaridade. Isso foi confirmado nas análises de variância univariadas, em que não se verificou efeito significativo ($p > 0,05$) de grupos nestes caracteres. Para os ácidos graxos (palmítico, esteárico, oleico, linoleico e

linolênico), contrariamente, verificou-se significância deste efeito ($p < 0,01$). Logo, mesmo restringindo-se as médias dos acessos entre 30,2% e 48,8% para o ácido oleico e de 36,5% a 55,9% para os ácidos poli-insaturados, ainda foi possível identificar grupos de similaridade com diferenças significativas entre si naquelas variáveis.

Nos grupos G4 e G5 é possível optar por acessos com diferenças em até 2,0% para o teor de óleo e, para o teor de proteína, em até 3,7% e 5,5%, respectivamente. Mesmo com a baixa variabilidade para esses caracteres, vislumbra-se a possibilidade de escolha de acessos divergentes, com algum incremento em relação à média dentro desses grupos.

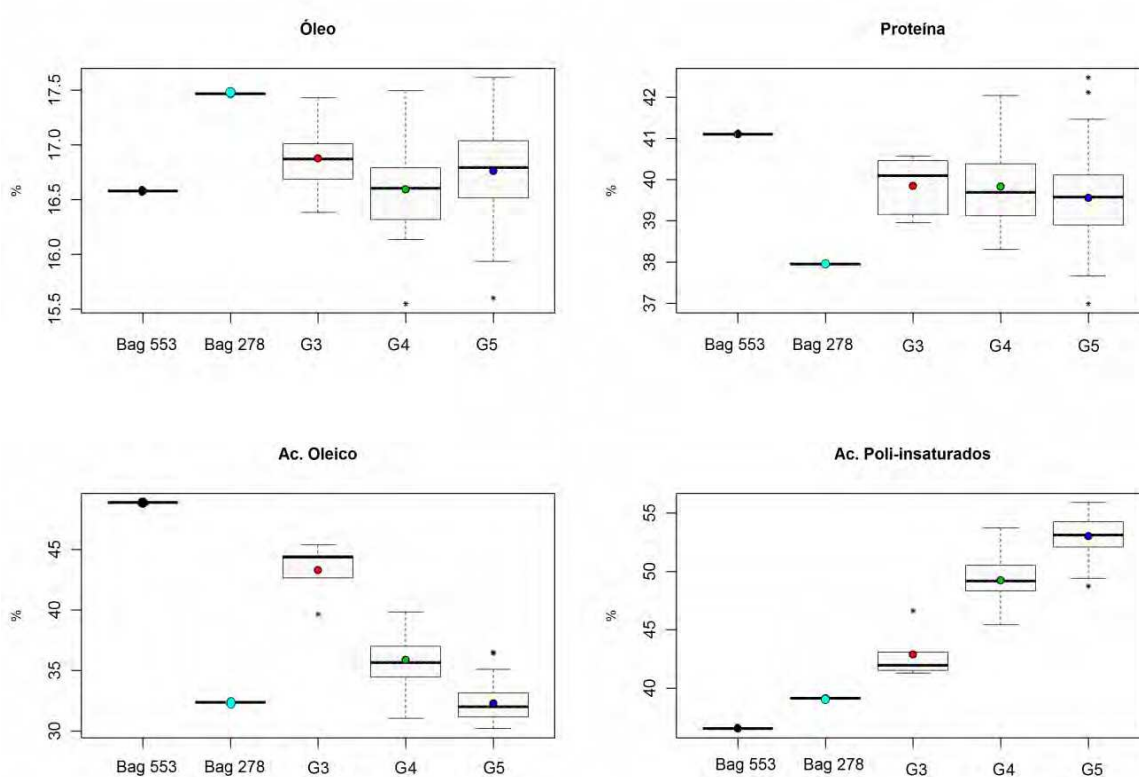


Figura 11. Boxplots de cinco grupos de similaridade obtidos por análise de agrupamento (UPGMA) de 82 acessos, com melhor perfil para estabilidade do óleo, em relação aos teores percentuais de óleo e proteína no grão, e dos ácidos oleico e poli-insaturados.

4.3.4 Importância relativa dos caracteres na discriminação dos acessos

A dispersão dos coeficientes de ponderação associados às duas primeiras variáveis canônicas permite observar a contribuição de cada caráter na dispersão dos

acessos (Figura 12). Quanto maior o coeficiente de uma variável maior é a sua influência sobre o grau de dispersão dos acessos. Isso porque, cada uma delas representa uma combinação linear das variáveis originais, e os pesos das variáveis são os próprios coeficientes desta combinação linear (Esbensen, 2002). Pode-se notar que, em geral, os teores de ácidos graxos têm maior contribuição, na distribuição dos acessos nas variáveis canônicas, do que os teores de óleo (OL) e de proteína (PROT), pois esses dois últimos caracteres apresentaram baixa correlação com ambos os eixos. Logo, infere-se que na coleção de germoplasma em estudo, os teores de óleo e de proteína não são caracteres que discriminam bem os acessos. Isso confirma os resultados já discutidos, em nível multivariado (ex. Figura 8) ou mesmo univariado (reduzidos coeficientes de variação genética: $CV_g(\text{óleo}) = 5,7\%$ e $CV_g(\text{proteína}) = 3,8\%$).

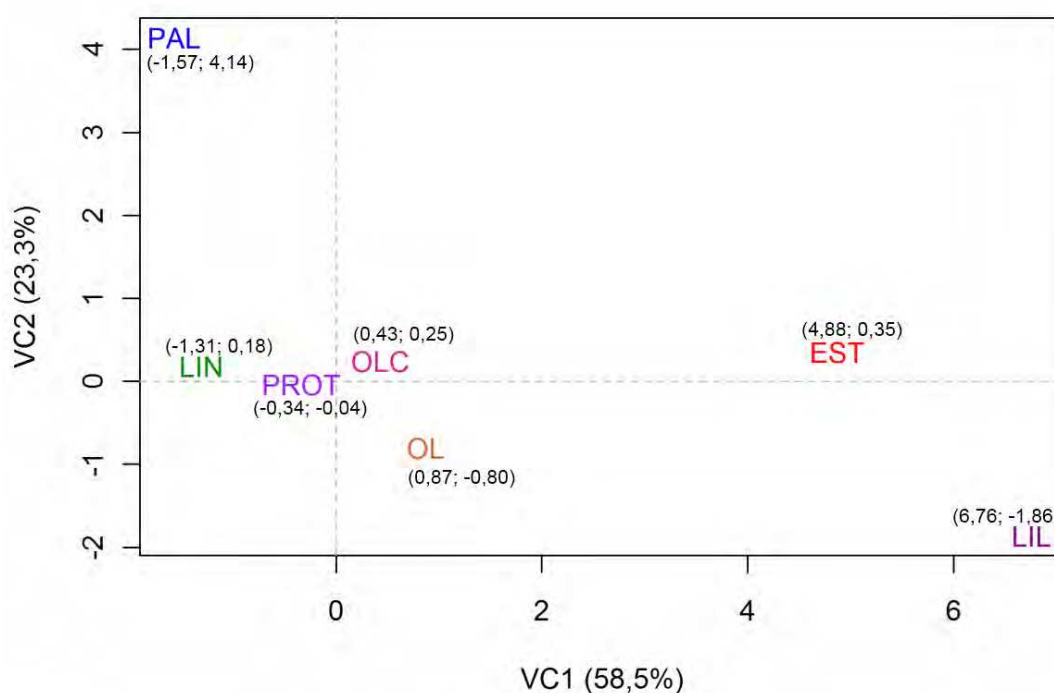


Figura 12. Coeficientes de ponderação das duas primeiras variáveis canônicas (VC1 e VC2), correspondentes às variáveis teor de óleo (OL), de proteína (PROT) e dos ácidos graxos palmítico (PAL), esteárico (EST), oleico (OLC), linoleico (LIN) e linolênico (LIL) no grão, em análise aplicada aos dados de 527 acessos do banco de germoplasma de soja da Secretaria da Agricultura Pecuária e Abastecimento do Estado de Goiás (Seagro-GO).

A participação dos coeficientes de ponderação em cada variável canônica pode ser medida também pela proporção do respectivo coeficiente (valor absoluto), em relação ao maior deles, também em valor absoluto (Jolliffe, 2002). Sendo assim, os coeficientes de

ponderação mais importantes, com proporções maiores que 50%, em relação à primeira variável canônica foram dos teores dos ácidos linolênico (LIL) e esteárico (EST), com pesos de 6,76 e 4,88, respectivamente (Figura 12). Já o teor de ácido palmítico (PAL) apresentou peso intermediário (proporção menor que 25%), com valor igual a -1,57. Na segunda variável canônica destacou o teor de ácido palmítico com alto coeficiente (4,14), seguido do ácido linolênico, com coeficiente intermediário (-1,86).

Em resumo, os caracteres que tiveram maior importância na discriminação dos acessos, na primeira variável canônica, são aqueles relacionados aos teores de ácidos graxos, sejam eles poli-insaturados (ácidos linoleico e linolênico) ou saturados (ácidos esteárico e palmítico). Ambos os grupos são de interesse do melhoramento da soja para a obtenção de óleos com melhor qualidade oxidativa e, ou, nutricional. Logo, a discriminação dos acessos levou em consideração os principais atributos relacionados à qualidade do óleo de soja. Como já comentado, o alto nível de ácidos graxos saturados tem duas implicações básicas. A primeira é em relação à utilização do óleo para produção de biodiesel, que provoca aumento do ponto de névoa e de congelamento deste combustível, trazendo complicações ao motor (Knothe, 2010). A outra diz respeito aos problemas de saúde associados às dietas com elevados índices desses ácidos graxos (Tyson & McCormick, 2006). Quanto aos ácidos poli-insaturados, sua implicação está relacionada à estabilidade do óleo, a qual melhora sob reduzidos teores destes ácidos (Knothe, 2010). Então, para melhor qualidade do óleo em ambos os setores, é desejável genótipos com menores teores de ácidos saturados e de ácidos poli-insaturados.

Para ratificar a importância dos ácidos graxos saturados e poli-insaturados na discriminação dos acessos por variáveis canônicas, foi feita outra representação gráfica a partir dos escores obtidos nessa análise (Figura 13). Com a utilização dos níveis baixo, médio e alto para esses caracteres, foi possível observar estruturação dos acessos quanto aos ácidos poli-insaturados e, mais claramente, quanto aos ácidos saturados (pouca sobreposição dos pontos associados aos diferentes níveis). Após esta análise pode-se perceber que acessos com menores teores de ácidos graxos saturados e poli-insaturados, simultaneamente, estão distribuídos na parte esquerda inferior da dispersão. O destaque para esta região deve-se ao fato de que essas características são importantes para a indústria de óleo. Esse padrão na dispersão ratifica o grupo C como o de melhor perfil para qualidade do óleo (Figura 4).

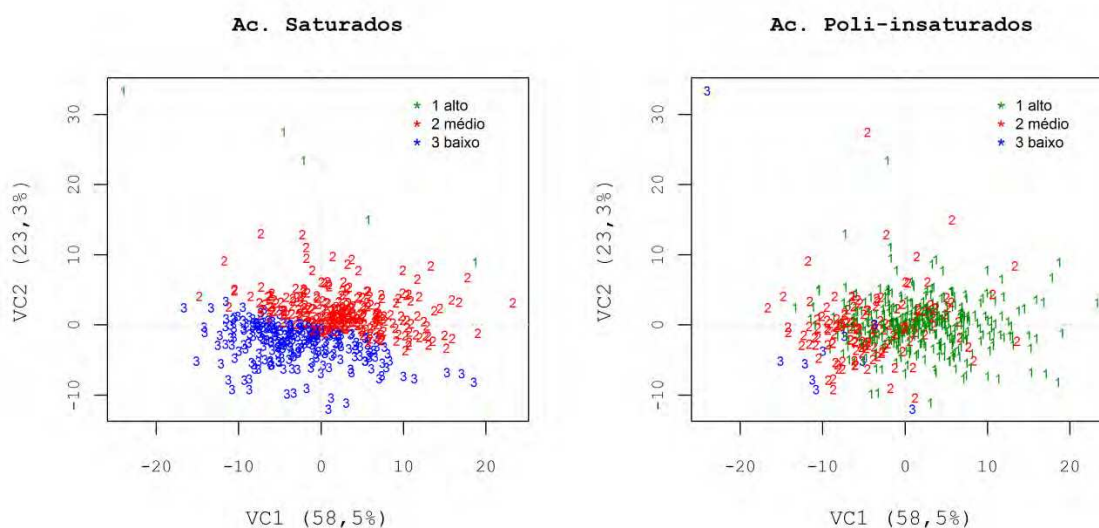


Figura 13. Representação gráfica dos escores de 527 acessos de soja do germoplasma da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em relação às duas primeiras variáveis canônicas (VC1 e VC2), para os teores dos ácidos graxos saturados e poli-insaturados, classificados segundo os níveis 1 (alto), 2 (médio) e 3 (baixo).

Outro tipo de análise foi implementado para avaliar a importância dos caracteres relacionados à qualidade do grão na discriminação dos grupos de similaridade (A, B, C, D e E). Neste caso, utilizando-se uma análise discriminante desses grupos foi possível maximizar a separação entre eles (Figura 14). Os escores da primeira variável discriminante canônica, que permite tal separação, resultam da seguinte função:

$$\text{Can1} = +0,02X_1 - 0,14X_2 - 0,54X_3 + 0,52X_4 + 0,24X_5 - 0,25X_6 + 1,32X_7$$

em que: as variáveis X_1 a X_7 são, respectivamente, os teores médios de proteína, óleo e dos ácidos graxos palmítico, esteárico, oleico, linoleico e linolênico.

Esta primeira discriminante canônica representa 78,5% da variação total entre os grupos de similaridade, e, por conseguinte, a separação entre os grupos nesse eixo foi devida, principalmente, aos teores dos ácidos linolênico, palmítico e esteárico, que tiveram, respectivamente, os pesos 1,32, -0,54 e 0,52. Os baixos pesos para os teores de óleo (-0,14) e proteína (0,02) confirmam, novamente, a baixa importância relativa desses caracteres na discriminação do germoplasma.

Ratificam-se, portanto, os resultados sobre os caracteres mais importantes na discriminação dos acessos, sejam nas variáveis canônicas (Figura 12) ou pela função discriminante dos grupos de similaridade obtidos pelo agrupamento via UPGMA. A

concordância obtida, resultante da combinação de três técnicas de análise multivariada, corrobora o fato de que os ácidos graxos saturados e poli-insaturados são os caracteres mais importantes na estruturação da diversidade desse germoplasma.

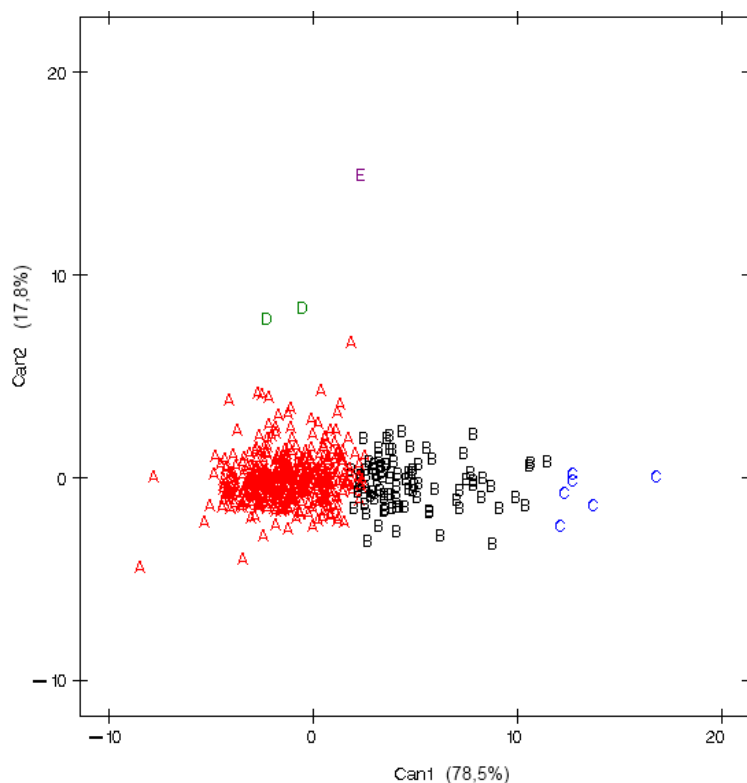


Figura 14. Análise discriminante canônica de 527 acessos do banco de germoplasma de soja da Secretaria da Agricultura Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), classificados quanto aos grupos de similaridade (A, B, C, D e E) obtidos pelo agrupamento via UPGMA, em relação a sete atributos relacionados à qualidade do grão.

Com intuito de investigar a importância das variáveis na discriminação dos ciclos de maturação (precoce, médio e tardio), ainda que admitindo algum confundimento de efeitos, obtiveram-se também as funções discriminantes para esta classificação. Neste caso, a primeira variável canônica explicou quase 95% da variação entre os ciclos, tendo permitido melhor discriminar os ciclos precoce e tardio (Figura 15). Como se observa, o ciclo médio teve a representação dos acessos sobreposta aos outros dois. Os escores dessa primeira variável discriminante canônica são obtidos pela seguinte função:

$$\text{Can1} = +0,03X_1 - 0,21X_2 + 1,33X_3 + 2,94X_4 + 1,47X_5 + 1,48X_6 + 1,30X_7$$

em que: as variáveis X_1 a X_7 são, respectivamente, os teores médios de proteína, óleo e dos ácidos graxos palmítico, esteárico, oleico, linoleico e linolênico.

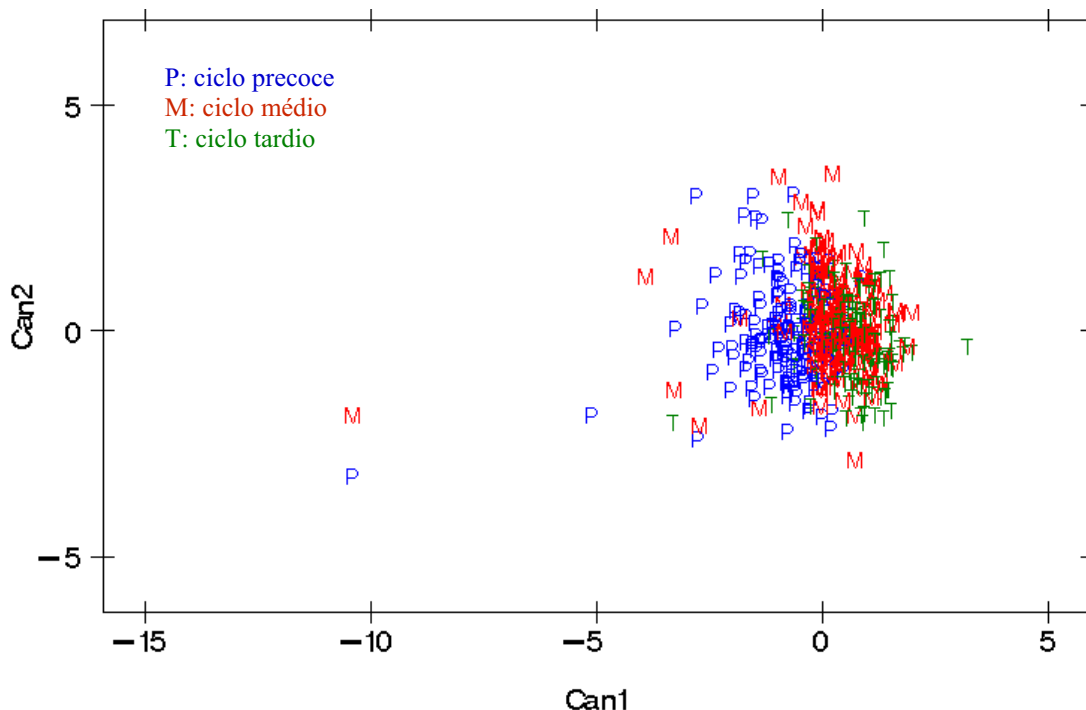


Figura 15. Análise discriminante canônica de 527 acessos do banco de germoplasma de soja da Secretaria da Agricultura Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), classificados quanto ao ciclo de maturação (precoce, médio e tardio), em relação a sete atributos relacionados à qualidade do grão.

A separação entre os acessos de ciclos precoce e tardio foi devida, principalmente, ao teor de ácido esteárico, com peso de 2,94 na função discriminante; seguido dos ácidos linoleico e oleico, com pesos 1,48 e 1,47, respectivamente. Os teores de óleo e de proteína, mais uma vez, foram as variáveis de menor participação na discriminação dos acessos quanto aos três ciclos de maturação. Contrariamente, a análise de variância univariada mostrou diferenciação ($p < 0,05$) entre os ciclos de maturação, quanto aos teores de proteína no grão; assim como em relação aos teores dos ácidos oleico e linoleico (Tabela 7). Em média, relativamente ao ciclo tardio, os acessos de ciclo precoce apresentam maior teor de ácido oleico e menor teor de ácido linoleico. Como já informado, tais características são importantes para a melhor estabilidade oxidativa do óleo de soja, o que favorece os acessos de ciclo precoce.

Tabela 7. Médias dos teores¹ de óleo, proteína e dos ácidos palmítico, esteárico, oleico, linoleico e linolênico, de 527 acessos do banco de germoplasma de soja da Secretaria da Agricultura Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), em relação aos ciclos de maturação², com a diferenciação estatística entre estes.

Ciclo	PROT	OL	PAL	EST	OLC	LIN	LIL
Precoce	39,0	16,7	12,0	2,1	25,8	51,5	6,4
Médio	39,2	16,7	11,9	2,1	25,6	52,5	6,5
Tardio	38,9	16,7	11,9	2,2	24,6	53,7	6,5
P-valor F	0,03	0,98	0,08	0,49	0,04	<0,01	0,60

¹- OL: teor de óleo em g/100g; PROT: teor de proteína em g/100g; PAL: teor de ácido graxo palmítico em g/100g; EST: teor de ácido graxo esteárico em g/100g; OLC: teor de ácido graxo oleico em g/100g; LIN: teor de ácido graxo linoleico em g/100g e LIL: teor de ácido graxo linolênico em g/100g.

²- A diferenciação entre ciclos está confundida com os efeitos de variação local referentes à alocação dos acessos em grupos de maturação (superblocos), realizada para facilitar o manejo de materiais de ciclos distintos.

Outra informação obtida em relação aos ciclos de maturação foram as estimativas de variância genética em cada ciclo. Essas variâncias, para os teores de óleo e dos ácidos graxos palmítico, oleico, linoleico e linolênico, foram maiores entre acessos do ciclo precoce. Já para o teor de proteína, as maiores variâncias genéticas foram observadas entre acessos de ciclo médio; e para ácido esteárico, entre os de ciclo tardio. Isto sugere que, de modo geral, os acessos de ciclo precoce são mais divergentes nos caracteres ligados ao óleo no grão.

No dendrograma obtido pelo agrupamento aglomerativo via UPGMA (Figura 8), observa-se que não houve padrão de distribuição dos acessos nos grupos e subgrupos, em relação aos ciclos de maturação. Isto sugere que acessos similares, considerando-se simultaneamente as sete variáveis analisadas, não apresentam necessariamente o mesmo ciclo. Mesmo tomando-se apenas as variáveis que melhor discriminaram os grupos de similaridade (teores de ácido linolênico e de ácidos saturados), valores baixos ou altos destas variáveis não se restringem a um determinado grupo de maturação. Isto abre a perspectiva de se utilizar acessos potenciais do banco de germoplasma para a obtenção de genótipos melhorados em atributos relacionados à qualidade do grão, com os diferentes perfis de qualidade identificados, em quaisquer dos grupos de maturação.

O perfil ideal do óleo de soja é atribuído à presença elevada de ácidos graxos monoinsaturados, presença reduzida de ácidos poli-insaturados e moderado conteúdo de ácidos saturados; isto, para a produção de alimentos e de biodiesel. Na subcoleção em estudo, foi possível identificar pelo menos 82 acessos com esse perfil (Tabela 6). O uso

desse material genético pode se dar de maneira mais direta via cultivo de genótipos que reúnam as propriedades de interesse, ou utilizando-os como genitores em cruzamentos estratégicos com cultivares comerciais. No primeiro caso, poder-se-ia lançar mão de linhagens já melhoradas desenvolvidas no âmbito do Programa soja da Seagro-GO e seus parceiros, haja vista apresentarem boa adaptação às condições de cultivo da soja na região.

Outro perfil para o óleo de soja, contrário ao demandado atualmente pela indústria, está relacionado a menores índices de ácidos saturados e maiores de poli-insaturados. Isso porque índices elevados de saturação no óleo têm sido comumente associado a altos níveis de colesterol no sangue e a problemas cardiovasculares. Por isso, dietas ricas em ácidos poli-insaturados, principalmente ácidos ômega-3, podem exercer efeito supressor no surgimento de doenças cardiovasculares, inflamatórias e câncer (Simopoulos, 2008; Deckelbaum, 2010; Farzaneh-Far et al., 2010). Se admitirmos a possibilidade de que esse perfil venha ser demandado em maior escala no futuro, é possível identificar acessos com tais características no banco de germoplasma de soja da Seagro-GO, e, mais especificamente em seus grupos de similaridade A e B, nesta ordem de importância (Figura 8).

A partir dos grupos de divergência identificados na subcoleção do banco de germoplasma de soja da Seagro-GO é possível, ainda, estabelecer uma coleção nuclear com base nos caracteres de interesse industrial aqui avaliados. Dessa forma, será possível concentrar esforços na avaliação específica de outros caracteres de interesse, mesmo envolvendo-se avaliações mais onerosas e elaboradas, bem como a inclusão de vários ambientes, partindo-se um conjunto de genótipos já devidamente caracterizados para atributos importantes relacionados à qualidade físico-química do grão.

5 CONCLUSÕES

- i)* Existe variabilidade genética na coleção de germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), para os teores de proteína e óleo no grão, bem como para os principais ácidos graxos; entretanto, os níveis mais elevados dos dois primeiros caracteres apresentam-se baixos relativamente aos encontrados em germoplasma de soja.
- ii)* A variabilidade genética desse germoplasma para caracteres relacionados aos teores de ácidos graxos no óleo (palmítico, esteárico, oleico, linoleico e linolênico) é ampla; porém, há limitação de acessos com elevados teores dos ácidos oleico, esteárico e linolênico.
- iii)* Teores elevados de ácido oleico e reduzidos de ácidos poli-insaturados, fundamentais à estabilidade oxidativa do óleo de soja, são características restritas a poucos acessos da subcoleção avaliada; sobretudo quando se consideram, simultaneamente, ambas as características num mesmo acesso.
- iv)* O germoplasma apresenta variabilidade genética satisfatória em teor de ácido linoleico (ômega-6), porém, baixo potencial para uso no melhoramento visando elevados teores de ácido linolênico (ômega-3). A cultivar UFV-3 (Bag 586), devido ao seu alto teor de ácidos graxos poli-insaturados, sobretudo o ácido linoleico, representa um acesso com potencial de uso direto ou para o melhoramento visando a obtenção de produtos com melhor perfil nutricional.
- v)* No que tange aos caracteres avaliados (teores de óleo, proteína e dos principais ácidos graxos no óleo de soja), a subcoleção de germoplasma mantida pela Seagro-GO pode ser estratificada em cinco grupos de similaridade: A (426 acessos), B (92 acessos), C (seis acessos), D (dois acessos) e E (um acesso).
- vi)* Os acessos do grupo C apresentam teores de ácidos graxos com melhor perfil para a qualidade industrial do óleo (estabilidade oxidativa), representando, portanto, genitores potenciais para programas de melhoramento que visem atender a esse tipo de demanda, incluindo-se a produção de biodiesel.

- vii)* Entre os acessos com óleo de melhor perfil de estabilidade oxidativa, é possível estabelecer grupos de similaridade com diferenças significativas entre si nos teores de ácidos graxos.
- viii)* Correlações genéticas e fenotípicas negativas foram observadas entre teores dos ácidos oleico e linoleico, e entre os ácidos oleico e linolênico. Isto representa um elemento facilitador para o melhoramento da soja visando maior estabilidade oxidativa do óleo (elevados níveis de ácido oleico e reduzidos níveis de ácidos graxos poli-insaturados).
- ix)* Entre os caracteres avaliados, os teores de ácidos graxos foram mais importantes na estruturação da divergência genética desse germoplasma, tendo se destacado os ácidos linolênico e palmítico (nesta ordem).

6 REFERÊNCIAS

- ALBERTAZZI, P.; PANSINI, F.; BONACCORSI G.; ZANOTTI, L.; FORINI, E.; ALOYSIO, D. The effect of dietary soy supplementation on hot flushes. **Obstetrics and Gynecology**, Washington, v. 91, n. 1, jan. 1998.
- ALBRECHT, L. P.; BRACCINI, A. L.; ÁVILA, M. R.; SUZUKI, L. S.; SCAPIM, C. A.; BARBOSA, M. C. Teores de óleo, proteínas e produtividade de soja em função da antecipação da semeadura na região oeste do Paraná. **Bragantia**, Campinas, v. 67, n. 4, p. 865-873, out./dez. 2008.
- ALBUQUERQUE, T. L.; LIMA, M. A.; OLIVEIRA, V. S.; COELHO, R. M.; RODRIGUES, M. C. P.; Processamento e aceitação sensorial e produto do tipo hambúrguer à base de soja (*Glycine max*) e atum (*Thunnus spp.*). **Centro de Pesquisa de Processamento de Alimentos**, Curitiba, v. 27, n. 2, p. 191-198, jul. 2009.
- ALDENDERFER, M.S.; BLASHFIELD, R. K. **Cluster Analysis**. 3.ed. Beverly Hills: SAGE Publications, 1984. 88 p.
- ALT, J. L.; FEHR, W. R, WELKE, G. A.; SANDHU, D. Phenotypic and Molecular Analysis of Oleate Content in the Mutant Soybean Line M23. **Crop Science**, Madison, v. 45, n. 5, p. 1997-2000, set./out. 2005.
- AMADOU, I.; YONG-HUI, S.; SUN, J.; GUO-WEI, L. Fermented soybean products: some methods, antioxidants compound extraction and their scavenging activity. **Asian Journal of Biochemistry**, New York, v. 4, n. 3, p. 68-76, dez. 2009.
- ANAI, T.; YAMADA, T.; HIDESHIMA, R.; KINOSHITA, T.; RAHMAN, S. M.; TAKAGI, Y. Two high-oleic-acid soybean mutants, M23 and KK21, have disrupted microsomal omega-6 fatty acid desaturase, encoded by *GmFAD2-1a*. **Breeding Science**, Tokyo, v. 58, n. 4, 447-452, abr. 2008.
- ANDERSON, J. W.; JOHNSTONE, B. M.; COOK-NEWELL, M. E. Meta-analysis of the effects of soy protein intake on serum lipids. **New England Journal of Medicine**: Waltham, v. 333, n.5, p. 276-82, ago. 1995.
- ANTHONY, M. S. Soy and cardiovascular disease: cholesterol lowering and beyond. **The Journal of Nutrition**, Washington, v. 130, n. 3, p. 662-663, mar. 2000.
- ANTI, A. B. Caracterização de germoplasma de soja e de feijão através de eletroforese de isoenzimas da semente. **Bragantia**, Campinas, v. 59, n. 2, p. 139-142, abr./jun. 2000.
- AZAM, M. M.; WARIS, A.; NAHAR, N. M. Prospects and potential of fatty acid methyl esters of some non-traditional seed oils for use as biodiesel in India. **Biomass and Bioenergy**, Oxford, v. 29, n. 4, p. 293-302, out. 2005.

AZIADEKEY, M.; SCHAPAUGH, W. T.; HERALD, T. J. Genotype by environment interaction for soymilk and tofu quality characteristics. **Journal of Food Quality**, Ames, v. 25, n. 3, p. 243-259, jul. 2002.

BACHLAVA, E.; BURTON, J. W.; BROWNIE, C.; WANG, S.; SANBAO, J.; CARDINAL, A. J. Heritability of Oleic Acid Content in Soybean Seed Oil and Its Genetic Correlation with Fatty Acid and Agronomic Traits. **Crop Science**, Madison, v. 48, n. 5, p. 1764-1772, set./out. 2008.

BAKER, K. M.; UTTERBACK, P. L.; PARSONS, C. M.; STEIN, H. H. Nutritional value of soybean meal produced from conventional, high-protein, or low-oligosaccharide varieties of soybeans and fed to broiler chicks. **Poultry Science**, Champaign, v. 90, n. 2, p. 390-395, fev. 2011.

BANDEGAN, A.; KIARIE, E.; PAYNE, R. L.; CROW, G. H.; GUENTER, W.; NYACHOTI, C. M. Standardized ileal amino acid digestibility in dry-extruded expeller soybean meal, extruded canola seed-pea, feather meal, and poultry by-product meal for broiler chickens. **Poultry Science**, Champaign, v. 89, n. 12, p. 2626-2633, dez. 2010.

BARRETO, S. L. T.; MOURA, W. C. O.; REIS, R. S.; HOSODA, L. R.; MAIA, G. V. C.; PENA, G. M. Soja integral processada em dietas para codornas japonesas em postura. **Revista Brasileira de Zootecnia**, Viçosa, v.39, n.9, p.1978-1983, set. 2010.

BARROS, L. M. **Caracterização morfológica e isoenzimática do cajueiro (*Anacardium occidentale* L.), tipos comum e anão precoce, por meio de técnicas multivariadas.** 1991. 256 f. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 1991.

BEKELE, F. L.; BEKELE, I.; BUTLER, D. R.; BIDAISEE, G. G. Patterns of Morphological Variation in a Sample of Cacao (*Theobroma Cacao* L.) Germplasm from the International Cocoa Genebank, Trinidad. **Genetic Resources and Crop Evolution**, Dordrecht, v. 53, n. 5, p. 933-948, ago. 2006.

BENIN, G.; CARVALHO, F. I. F.; OLIVEIRA, A. C.; MARCHIORO, V. S.; LORENCETTI, C.; KUREK, A. J.; SILVA, J. A. G.; CRUZ, P. J.; HARTWIG, I.; SCHMIDT, D. A. M. Comparações entre medidas de dissimilaridade e estatísticas multivariadas como critérios no direcionamento de hibridações em aveia. **Ciência Rural**, Santa Maria, v. 33, n. 4, ago. 2003.

BERED, F.; NETO, J. F. B.; ROCHA, B. M.; PEGORARO, D. G.; VACARO, E.; CARVALHO, F. I. F. Caracterização de germoplasma de trigo por meio dos caracteres adaptativos ciclo e estatura. **Pesquisa Agropecuária Brasileira**, Brasília, v. 37, n.2, p. 145-150, 2002.

BERGSMA, G.; KAMPMAN, B.; CROEZEN, H.; SEVENSTER, M. **Biofuels and their global influence on land availability for agriculture and nature.** Delft: CE, 2006. 86 p. Disponível em: <http://www.cedelft.eu/publicatie/%09biofuels_and_their_global_influence_on_land_availability_for_agriculture_and_nature/505>. Acesso em: 10 dez. 2010.

BERTAN, I.; CARVALHO, F. I. F.; OLIVEIRA, A. C.; VIEIRA, E. A.; HARTWIG, I.; SILVA, J. A. G.; SHIMIDT, D. A. M.; VALÉRIO, I. P.; BUSATO, C. C.; RIBEIRO, G.

Comparação de métodos de agrupamento na representação da distância morfológica entre genótipos de trigo. **Revista Brasileira de Agrociência**, Pelotas, v.12, n. 3, p.279-286, jul./set. 2006.

BHATIA, C. R.; NICHTERLEIN, K.; MALUSZYNSKI, M. Oilseed cultivars developed from induced mutations and mutations altering fatty acid composition. **Mutation Breeding Review**, Vienna, n. 11, p. 1-36, 1999.

BHATIA, J.; GREER, F. Use of Soy Protein-Based Formulas in Infant Feeding. **Pediatrics**, Elk Grove Village, v. 121, n. 5, p. 1062-1068, mai. 2008.

BONACINA, M. S. **Desenvolvimento e Caracterização de Empanado a partir de Corvina (*Micropogonias furnieri*)**. 2006. 115 f. Dissertação (Mestrado em Engenharia e Ciência dos Alimentos) - Fundação Universidade Federal do Rio Grande, Rio Grande, 2006.

BONATO, E. R.; BERTAGNOLLI, P. F.; LANGE, C. E.; RUBIN, S. A. L. Teor de óleo e de proteína em genótipos de soja desenvolvidos após 1990. **Pesquisa Agropecuária Brasileira**, Brasília, v. 35, n. 12, p.2391-2398, dez. 2000.

BONATO, E. R.; DALLAGNOL, A. Soybean in Brazil: Production and Research. In: World Soybean Research Conference, 3., 1985, Iowa/USA. **Proceedings...** Ames: Soyatech, 1985. p.1248-1256, 1985.

BONETTI, L. P. Cultivares e seu melhoramento genético. In: VERNETTI, F.J. (Ed.). **Soja: Genética e Melhoramento**. Campinas: Fundação Cargill, 1983. v. 2, cap. 9, p. 741-803.

BONETTI, L. P. Resultados de cinco anos de testes e principais características de variedades de soja recomendadas para o Rio Grande do Sul. **Boletim Técnico Fecotrigo**, Porto Alegre, v. 53, p. 2-28. 1981.

BORBA, T. C. O. **Diversidade genética e mapeamento associativo para produção e qualidade de grão dos acessos melhorados da coleção nuclear de arroz da Embrapa**. 2007. 110 f. Tese (Doutorado em Agronomia: genética e melhoramento de plantas) - Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2007.

BRANDÃO, A. S. P.; REZENDE, G. C.; MARQUES, R. W. C. Crescimento agrícola no período 1999/2004: a explosão da soja e da pecuária bovina e seu impacto sobre o meio ambiente. **Economia Aplicada**, Ribeirão Preto, v. 10, n. 2, p. 249-266, abr. 2006.

BRASIL. Agência Nacional de Vigilância Sanitária, Anvisa. **Lista das alegações de propriedade funcional aprovadas: Alimentos com Alegações de Propriedades Funcionais e ou de Saúde, Novos Alimentos/Ingredientes, Substâncias Bioativas e Probióticos**. 2005. Disponível em: < <http://www.anvisa.gov.br/alimentos/comissoes/tecno.htm> >. Acesso em: 20 dez. 2010.

BRASIL. Agência Nacional de Vigilância Sanitária, Anvisa. **Resolução nº 19, de 30 de abril de 1999**. Aprova o Regulamento Técnico de Procedimentos para Registro de Alimento com Alegação de Propriedades Funcionais e ou de Saúde em sua Rotulagem.

Diário Oficial da República Federativa do Brasil, Brasília, 10 dez. 1999. Disponível em: <http://www.anvisa.gov.br/legis/resol/19_99.htm>. Acesso em: 30 ago. 2010.

BRASIL. Ministério da Agricultura Pecuária e Abastecimento, MAPA. **Portaria nº 795 de 15 de dezembro de 1993**. Norma de identidade, qualidade, embalagem, marcação e apresentação do farelo de soja. 1993.

BRIM, C. A.; BURTON, J. W. Recurrent selection in soybeans. II. Selection for increased percent protein in seeds. **Crop Science**, v.19, n. 4, p.494-498. jul./ago. 1979

BROGIN, R. L.; GODINHO, V. P. C.; UTUMI, M. M.; ALVES, J. C.; RIBEIRO, I. A.; REIS, F. L. F.; DUARTE, J. P.; FARIA, L. C.; ROLIM, R. B.; MONTEIRO, P. M. F. O.; NUNES JÚNIOR, J.; KIIHL, R. A. S.; ALMEIDA, L. A.; SOUZA, P. I. M.; ARANTES, N. E.; NEIVA, L. C. S.; GUIMARÃES, L. B.; SILVA, L. O. BRSGO 204 [Goiânia]: Extensão de Indicação para o Estado de Rondônia. In: Reunião de Pesquisa de Soja da Região Central do Brasil, 31., 2010, Brasília. **Resumos...**Brasília: Embrapa, 2010. P. 367-368.

BURTON, J. W. Recent developments in breeding soybeans for improved oil quality. **European Journal of Lipid Science and Technology**, Weinheim, v. 93, n. 4, p. 121- 128, abr. 1991.

BURTON, J. W.; WILSON, R. F.; BRIM, C. A. Recurrent selection in soybean . IV. Selection for increased oleic acid percentage in seed oil. **Crop Science**, Madison, v. 23, n. 4, p. 744-747, jul./ago. 1983.

BURTON, J. W.; WILSON, R. F.; BRIM, C. A. Registration of N79-2077-12 and N87-2122-4, two soybean germplasm lines with reduced palmitic acid in seed oil. **Crop Science**, Madison, v. 34, n. 1, p. 301, jan./fev. 1994.

BURTON, J. W.; WILSON, R. F.; REBETZKE, G.J. Registration of N98-4445A mid-oleic soybean germplasm line. **Crop Science**, Madison, v.46 ,n. 2, p. 1010-1012, mar./abr. 2006.

BUSO, G. S. C.; RANGEL, P. H.; FERREIRA, M. E. Analysis of genetic variability of South American wild rice populations (*Oryza glumaepatula*) with isozymes and RAPD markers. **Molecular Ecology**, Columbia, v. 7, n. 1, p. 107-117, jan. 1998.

CAFÉ, M. B.; SAKOMURA, N. K.; JUNQUEIRA, O. M.; CARVALHO, M. R. B.; DEL BIANCHI, M. Determinação do Valor Nutricional das Sojas Integrais Processadas para Aves. **Revista Brasileira de Ciência Avícola**, Campinas, v. 2, n. 1, p. 67-74, jan./abr. 2000.

CALINSKI, R. B.; HARABASZ, J. A dendrite method for cluster analysis. **Communications in Statistics: Theory and Methods**, London, v. 3, n. 1, p. 1-27, jan. 1974.

CASSUCE, F. C. C.; SILVA, F. M. Análise Econômica do Cultivo da Soja em Grão: 1990-2003. **Revista Unileste**, Coronel Fabriciano, v. 2, n. 1, p. 1-16, jan. 2006. Disponível em: <http://www.unilestemg.br/revistaonline/volumes/02/downloads/artigo_21.pdf>. Acesso em: 18 abr. 2008.

- CHANDRA, S.; HUAMAN, Z.; KRISHNA S. H.; ORTIZ, R. Optimal sampling strategy and core collection size of Andean tetraploid potato based on isozyme data - a simulation study. **Theoretical and Applied Genetics**, Berlin, v. 104, n. 8, p. 1325-1334, jun. 2002.
- CHATTI, K.; SADDOUD, O.; SALHI-HANNACHI, A.; MARS, M.; MARRAKCHI, M.; TRIFIL, M. Analysis of Genetic Diversity and Relationships in a Tunisian Fig (*Ficus carica*) Germplasm Collection by Random Amplified Microsatellite Polymorphisms. **Journal of Integrative Plant Biology**, Malden, v. 49, n. 3, p. 386-391, mar. 2007.
- CHEESBROUGH, T. M. Changes in the Enzymes for Fatty Acid Synthesis and Desaturation during Acclimation of Developing Soybean Seeds to Altered Growth Temperature. **Plant Physiology**, Rockville, v. 90, n. 2, p. 760-764, jun. 1989.
- CHENG, Y.; HU, Q.; LI, L.; SAITO, M.; YIN, L. Production of Sufu, a Traditional Chinese Fermented Soybean Food, by Fermentation with *Mucor flavus* at Low Temperature. **Food Science and Technology Research**, Seoul, v. 15, n. 4, p. 347-352, mar. 2009.
- CHERRAK, C. M.; PANTALONE, V. R.; MEYER, E. J.; ELLIS, D. L.; MELTON, S. L.; WEST, D. R.; MOUNT, J. R. Low-palmitic, low-linolenic soybean development. **Journal of the American Oil Chemists' Society**, Urbana, v. 80, n. 6, p. 539-543, jun. 2003.
- CHUNG, J.; BABKA, H. L.; GRAEF, G. L.; STASWICK, P. E.; LEE, D. J.; CREGAN, P. B.; SHOEMAKER, R. C.; SPECHT, J. E. The Seed Protein, Oil, and Yield QTL on Soybean Linkage Group I. **Crop Science**, Madison, v. 43, n. 3, p. 1053-1067, mai./jun. 2003.
- CIABOTTI, S.; BARCELOS, M. F. P.; PINHEIRO, A. C. M.; CLEMENTE, P. R.; LIMA, M. A. C. Características sensoriais e físicas de extratos e tofus de soja comum processada termicamente e livre de lipoxigenase. **Ciência e Tecnologia de Alimentos**, Campinas, v. 27, n. 3, p. 643-648, jul./set. 2007.
- CLEMENTE, T. E.; CAHOON, E. B. Soybean Oil: Genetic Approaches for Modification of Functionality and Total Content. **Plant Physiology**, Rockville, v. 151, n. 3, p. 1030-1040, nov. 2009.
- COBER, E. R.; VOLDENG, H. D. Developing high-protein, high-yield soybean populations and lines **Crop Science**, Madison, v. 40, n. 1, p. 39-42, jan. 2000.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da Safra Brasileira: Grãos 2009/2010 - Décimo Segundo Levantamento (setembro/2010)**. Brasília: Conab, 2010b. Disponível em: <<http://www.conab.gov.br>>. Acesso em: 20 out. 2010.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento de safra brasileira: grãos - Quarto levantamento(janeiro/ 2011)**. Brasília: Conab, 2011. Disponível em: <<http://www.conab.gov.br>>. Acesso em: 10 jan. 2011.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **Indicadores da Agropecuária: Exportação complexo soja e trigo**. Brasília: Conab, 2010a. Disponível em: <<http://www.conab.gov.br>>. Acesso em: 20 out. 2010.

- CORDEIRO, M. T. C.; ABADIE, T. Coleções nucleares. In: NASS, L.L. (Ed.). **Recursos genéticos vegetais**. Brasília, DF: Embrapa Recursos Genéticos e Biotecnologia: Brasília, 2007. cap. 17, p. 575-602.
- CORONEL, D. A.; MACHADO, J. A. D.; CARVALHO, F. M. A. Fontes de crescimento das exportações do complexo soja brasileiro. **Análise**, Porto Alegre, v. 19, n. 2, p. 62-77, jul./dez. 2008.
- COSTA NETO, P. R.; ROSSI, L. F. S.; ZAGONEL, G.F.; RAMOS, L. P. Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em frituras. **Química Nova**, v. 23, n. 4, p. 531-537, jul./ago. 2000.
- CRUZ, C. D. **Programa Genes - Análise multivariada e simulação**. Viçosa: Editora UFV, 2006. 175 p.
- CRUZ, C. D. Programa Genes: Aplicativo Computacional em Estatística Aplicada à Genética (Genes - Software for Experimental Statistics in Genetics). **Genetics and Molecular Biology**, São Paulo, v. 21, n. 1, mar. 1998.
- CRUZ, C. D.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2003. 585 p.
- CRUZ, C. D.; REGAZZI, A. D. **Modelos Biométricos Aplicados ao Melhoramento Genético**. 2 ed. Viçosa: Editora UFV, 2001. 390 p.
- DECKELBAUM, R. J. n-6 and n-3 Fatty Acids and Atherosclerosis: Ratios or Amounts? **Arteriosclerosis, Thrombosis, and Vascular Biology**, Dallas, v. 30, n. 12, p. 2325-2326. dez. 2010.
- DEMIRBAS, A. Importance of biodiesel as transportation fuel. **Energy Policy**, Oxford, v. 5, n. 9, p. 4661-4670, set. 2007.
- DEMIRBAS, A. Progress and recent trends in biodiesel fuels. **Energy Conversion and Management**, Oxford, v. 50, n. 1, p. 14-34, jan. 2009.
- DEUTSCH, M. J. **AOAC official methods of analysis**. Vitamins and other nutrients. Hardbound: AOAC, 1995. cap. 45, p. 65
- DIAS, W. P.; SILVA, J. F. V.; CARNEIRO, G. E. S.; GARCIA, A.; ARIAS, C. A. A. Nematóide de Cisto da Soja: Biologia e Manejo Pelo Uso da Resistência Genética. **Nematologia Brasileira**, Piracicaba, v. 33, n. 1, p. 1-16, jan. 2009.
- DORNBOS, D. L. JR.; MULLEN, R. E. Soybean seed protein and oil contents and fatty acid composition adjustments by drought and temperature. **Journal of the American Oil Chemists' Society**, Urbana, v. 69, n. p. 228-231, mar.1992.
- DUARTE, J. B. **Princípios e utilização de técnicas multivariadas no melhoramento de plantas**. Piracicaba: Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo. Monografia. 1997. 69 p.

- DUARTE, J. B. **Sobre o emprego e a análise estatística do delineamento em blocos aumentados no melhoramento genético vegetal**. 2000. 293 f. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2000.
- DUARTE, J. M. **Conversão de linhagens elites em milho de alta qualidade protéica (QPM)**. 2003. 129 f. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas)-Universidade Federal de Lavras, Lavras, 2003.
- DUDA, R. O.; HART, P. E. **Pattern classification and scene analysis**. New York: Wiley, 1973. 482 p.
- DUNN, R. O.; SHOCKLEY, M. W.; BAGBY, M. O. Improving the low-temperature properties of alternative diesel fuels: Vegetable oil-derived methyl esters. **Journal of the American Oil Chemists' Society**, Urbana, v. 79, n. 7, p. 709-715, dec. 2002.
- EDWARDS, A.W. F.; CAVALLI-SFORZA, L. L. A method for cluster analysis. **Biometrics**, Washington, v. 21, n. 2, p. 362-75, jun. 1965.
- ERKKILA A.; MELLO, V. D. F.; RISERUS, U.; LAAKSONEN, D. E. Dietary fatty acids and cardiovascular disease: An epidemiological approach. **Progress in Lipid Research**, Oxford, v. 47, n. 3, p. 172-187, mai. 2008.
- ESBENSEN, K. H. **Multivariate Data Analysis in Practice: An Introduction to Multivariate Data Analysis and Experimental Design**. 5 ed. Oslo: CAMO Process, 2002. 598 p.
- EVANS, D. E.; TSUKAMOTO, C.; NIELSEN, N. C. A Small Scale Method for the Production of Soymilk and Silken Tofu. **Crop Science**, Madison, v. 37, n. 5, p. 1463-1471, set./out. 1997.
- FARRIS, J. S. On the Cophenetic Correlation Coefficient. **Systematic Biology**, Oxford, v.18, n. 3, p. 279-285, set. 1969.
- FARZANEH-FAR, R.; LIN, J.; EPEL, E. S.; HARRIS, W. S.; BLACKBURN, E. HWHOOLEY, M. A. Association of Marine Omega-3 Fatty Acid Levels With Telomeric Aging in Patients With Coronary Heart Disease. **Journal of American Medical Association**, Chicago, v. 303, n. 3, p. 250-257, jan. 2010.
- FEARNSIDE, P. M. Soybean cultivation as a threat to the environment in Brazil. **Environmental Conservation**, Cambridge, v. 28, n. 1, p. 23-38, mai. 2001.
- FEHR, W. R. Breeding for Modified Fatty Acid Composition in Soybean. **Crop Science**, Madison, v. 47. supl. 3, p. S72-S87, dez. 2007.
- FEHR, W. R.; WELKE, G. A.; HAMMOND, E. G.; DUVICK, D. N.; CIANZIO, S. R. Inheritance of reduced linolenic acid content in soybean genotypes A16 and A17. **Crop Science**, Madison, v. 32, n. 4, p. 903-906, jul./ago. 1992.
- FEHR, W. R.; HAMMOND, E. G. Soybeans having low linolenic acid content and method of production. **United States Patent**, Washington, patent 5534425, jul, 1996.

FEHR, W.R.; WELKE, G.A.; HAMMOND, E.G.; DUVICK, D.N.; CIANZIO, S.R. Inheritance of reduced palmitic acid content in seed oil of soybean. **Crop Science**, Madison, v. 31, n. 1, p.88-89, jan./fev. 1991.

FENG, L.; BURTON, J. W.; CARTER JR, T. E.; PANTALONE, V. R. Recurrent Half-Sib Selection with Testcross Evaluation for Increased Oil Content in Soybean. **Crop Science**, Madison, v. 44, n. 1, p. 63-69, jan./fev. 2004.

FERREIRA, D. F. **Estatística multivariada**. Lavras: Editora UFLA, 2008. 662 p.

FERREIRA, M. A. J. F. **Utilização das técnicas de marcadores moleculares na genética de populações, na genética quantitativa e no melhoramento de plantas**. Boa Vista: Embrapa Roraima (Embrapa Roraima. Documentos, 1), 2003. 63 p.

FONSECA, J. R.; SILVA, H. T. Identificação de duplicidades de acessos de feijão por meio de técnicas multivariadas. **Pesquisa Agropecuária Brasileira**, Brasília, v.34, n.3, p. 409-414, mar. 1999.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **Corporate Repository. Crop Prospects and Food Situation**, Roma, n. 2, 2008. Disponível em: <<http://www.fao.org/documents>>. Acesso em: 20 abr. 2008.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **FAOSTAT- Food and Agricultural commodities production**. Disponível em: <<http://faostat.fao.org/site/339/default.aspx>>. Acesso em: 07 nov. 2010.

FRANKEL, O.H. Genetic perspectives of germplasm conservation. In: ARBER, W.; LLIMENSEE, K.; PEACOCK, W. J.; STARLINGER, P. (Ed.). **Genetic Manipulation: Impact on Man and Society**. Cambridge: Cambridge University Press, 1984. p. 161-170.

FREEMANTLE, E.; VANDAL, M.; TREMBLAY-MERCIER, J.; TREMBLAY, S.; BLACHÈRE, J.; BÉGIN, M. E.; BRENNAN, J. T.; WINDUST, A.; CUNNANE, S. C. Omega-3 fatty acids, energy substrates, and brain function during aging. **Prostaglandins, Leukotrienes and Essential Fatty Acids**, London, v. 75, n. 3, p. 213-220, set. 2006

FREITAS, D. G. C.; MORETTI, R. H. Caracterização e avaliação sensorial de barras de cereais funcional de alto teor protéico e vitamínico. **Ciência e Tecnologia de Alimentos**, Campinas, v. 26, n. 2, p. 318-324, 2006.

GARCIA, J. R.; ROMEIRO, A. R. Governança da cadeia produtiva do biodiesel brasileiro. **Revista de Política Agrícola**, Brasília, v. 18, n. 1, p. 60-79, jan./mar. 2009.

GIZLICE, Z.; CARTER, T. E.; BURTON, J. W. Genetic diversity in North American soybean. I: Multivariate analysis of founding stock and relation to coefficient of parentage. **Crop Science**, Madison, v. 33, n. 3, p. 614-620, mai./jun. 1993.

GOEDERT, C. O. Histórico e avanços em recursos genéticos no Brasil. In: NASS, L.L. (Ed.). **Recursos genéticos vegetais**. Brasília, DF: Embrapa Recursos Genéticos e Biotecnologia: Brasília, 2007. cap. 1, p. 23-60.

GOLBITZ, P. Traditional soyfoods: processing and products. **Journal of Nutrition**, Bethesda, v. 125, Supl. 3, p. 570S-572S. mar. 1995.

- GONÇALVES, C. A.; SOARES, N. S.; BOLINA, C. O.; BARROS, E. G. Influência da Temperatura no Acúmulo de Proteínas em Sementes de Soja. **Revista Brasileira de Biociências**, Porto Alegre, v. 5, supl. 2, p. 1038-1040, jul. 2007.
- GRAEF, G.; LAVALLEE, B. J.; TENOPIR, P.; TAT, M.; SCHWEIGER, B.; KINNEY, A. J.; GERPEN, J. H. V.; CLEMENTE, T. E. A high-oleic-acid and low-palmitic-acid soybean: agronomic performance and evaluation as a feedstock for biodiesel. **Plant Biotechnology Journal**, Bethesda, v. 7, n. 5, p. 411-421, jun. 2009.
- GRESSLER, L. A. **Introdução à pesquisa: projetos e relatórios**. 2. ed. São Paulo: Loyola, 2004. 295p.
- GUI, M.M.; LEE, K.T.; BHATIA, S. Feasibility of edible oil vs. non-edible oil vs. waste edible oil as biodiesel feedstock. **Energy**, v. 33, n. 11, p. 1646-1653, nov. 2008.
- GUIMARÃES, P. E. O.; PACHECO, C. A. P.; PAES, M. C. D.; SANTOS, M. X.; PARENTONI, S. N.; GAMA, E. E. G.; MEIRELLES, W. F.; RIBEIRO, P. H. E.; MONTEIRO, M. A. R. **BR 473: variedade de milho amarelo com qualidade protéica melhorada (QPM)**. Sete Lagoas: Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA Milho e Sorgo, 2004. 4 p. (Comunicado Técnico, 105).
- GUODONG, Z.; JINLING, W. Genetic variability of main agronomic characters and protein percentage in interspecific crosses of soybeans. **Soybean Science**, Beijing, v. 8, n. 1, p. 1-10, 1989.
- HA, B.; BENNETT, J. B.; HUSSEY, R. S.; FINNERTY, S. L.; BOERMA, H. R. Pedigree Analysis of a Major QTL Conditioning Soybean Resistance to Southern Root-Knot Nematode. **Crop Science**, Madison, v. 44, n. 3, p. 758-763, mai. 2004.
- HALKIDI, M.; BATISTAKIS, Y.; VAZIRGIANNIS, M. Clustering validity checking methods: part II. **Newsletter ACM SIGMOD Record**, New York, v.31 n.3, set. 2002
- HALKIDI, M.; BATISTAKIS, Y.; VAZIRGIANNIS, M. On Clustering Validation Techniques. **Journal of Intelligent Information Systems**, New York, v.7, n. 2-3, dez. 2001.
- HALKIDI, M.; VAZIRGIANNIS, M. Clustering Validity Assessment: Finding the Optimal Partitioning of a Data Set. In: International Conference on Data Mining, 2001, California. **Proceedings...**California: IEEE, 2001, p. 187-194.
- HAMMOND, E.G.; FEHR, W.R. Registration of A5 germplasm line of soybean (Reg. No. GP44). **Crop Science**, Madison, v. 23, n. 1, p. 192, jan./fev. 1983.
- HARRIS, W. S.; MILLER, M.; TIGHE, A. P.; DAVIDSON, M. H.; SCHAEFER, E. J. Omega-3 fatty acids and coronary heart disease risk: Clinical and mechanistic perspectives. **Atherosclerosis**, London, v. 197, n. 1, p. 12-24, mar. 2008
- HEPPARD, E. P.; KINNEY, A. J.; STECCA, K. L.; MIAO, G. H. Developmental and Growth Temperature Regulation of Two Different Microsomal [omega]-6 Desaturase Genes in Soybeans. **Plant Physiology**, Rockville, v. 110, n. 1, p. 311-319, jan. 1996.

- HERNÁNDEZ, M. D.; MARTÍNEZ, F. J.; JOVERC, M.; GARCÍA, B. Effects of partial replacement of fish meal by soybean meal in sharpnose seabream (*Diplodus puntazzo*) diet. **Aquaculture**, Oxford, v. 263, n. 1-4, p. 159-167, mar. 2007.
- HIROMOTO, D. M.; VELLO, N. A. The genetic base of Brazilian soybean (*Glycine max* (L.) Merrill) cultivars. **Revista Brasileira de Genética**, Ribeirão Preto, v. 9, n. 2, p. 295-306, 1986.
- HOLBROOK, C. C.; BURTON, J. W.; CARTER, T. E. Evaluation of recurrent restricted index selection for increasing yield while holding seed protein constant in soybean. **Crop Science**, Madison, v. 29, n. 2, p. 324-329, mar./abr. 1989.
- HOLGERSSON, M. The limited value of cophenetic correlation as a clustering criterion. **Pattern Recognition**, Oxford, v. 10, n. 4, p. 287-295, 1978.
- HOREJSI, T. F.; FEHR, W. R.; WELKE, G. A.; DUVICK, D. N.; HAMMOND, E. G.; CIANZIO, S. R. Genetic control of reduced palmitate content in soybean. **Crop Science**, Madison, v. 34, n. 2, p. 331-334, mar./abr. 1994.
- HOU, G.; ABLETT, G. R.; PAULS, K. P.; RAJCAN, I. Environmental effects on fatty acid levels in soybean seed oil. **Journal of the American Oil Chemists' Society**, Urbana, v. 83, n. 9, p. 759-763, set. 2006.
- HUNGRIA, M.; CAMPO, R. J.; MENDES, I. C. **Fixação biológica do nitrogênio na cultura da soja**. Londrina: Embrapa Soja, 2001. 48 p. (Circular Técnica n. 35).
- HYMOWITZ, T.; COLLINI, F. I.; PANAZNER, J.; WALKER, W. M. Relationship between the content of oil, protein and sugar in soybean seed. **Agronomy Journal**, Madison, v. 64, n. 5, p. 613-616, 1972.
- INSEL, P.; TURNER, R. E.; ROSS, D. **Discovering Nutrition**. 3. ed. Sudbury: Jones and Bartlett Publishers, 2010. p. 169-207.
- INSTITUTO ADOLFO LUTZ. **Normas analíticas: métodos químicos e físicos para análises de alimentos**. 4. ed. São Paulo, 2008. 1020 p.
- IQBAL, Z.; ARSHAD, M.; ASHRAF, M.; NAEEM, R.; MALIK, M. F.; WAHEED, A. Genetic Divergence and Correlation Studies of Soybean [*Glycine Max* (L.) Merrill.] Genotypes. **Pakistan Journal of Botany**, Karachi, v. 42, n. 2, p. 971-976, abr. 2010.
- IQBAL, Z.; ARSHAD, M.; ASHRAF, M.; TARIQ, M.; WAHEED, A. Evaluation of soybean [*Glycine max* (L.) Merrill] Germplasm for some important morphological traits using multivariate analysis. **Pakistan Journal of Botany**, Karachi, v. 40, n. 6, p. 2323-2328, dez. 2008.
- JI, S.; HARDY, R. W.; WOOD, P. A. Transgenic expression of n-3 fatty acid desaturase (fat-1) in C57/BL6 mice: Effects on glucose homeostasis and body weight. **Journal of Cellular Biochemistry**, Malden, v. 107, n. 4, p. 809-17, jul. 2009.
- JOHNSON, H. W.; ROBINSON, H. F.; COMSTOCK, R. E. Estimates of genetic and environmental variability in soybeans. **Agronomy Journal**, Madison, v. 47, n. 11, p. 314-318. 1955.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 3. ed. New Jersey: Prentice-Hall, 1992. 642 p.

JOLLIFFE, I. T. **Principal component analysis**. 2. ed. New York: Springer, 2002. 487 p

KAIMOWITZ, D.; SMITH, J. Soybean technology and the loss of natural vegetation in Brazil and Bolivia. In: ANGELSEN, A; KAIMOWITZ, D. (Ed.). **Agricultural technologies and tropical deforestation**. Wallingford: CABI Publishing, 2001. cap. 11, p. 195-211.

KARIA, C. T. **Caracterização genética e morfoagronômica de germoplasma de *Stylosanthes guianenses* (Aubl.) SW**. 2008. 138 f. Tese (Doutorado em Agronomia: genética e melhoramento de plantas) - Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2008.

KATAN, M. B. Omega-6 polyunsaturated fatty acids and coronary heart disease. **Journal of Clinical Nutrition**, Bethesda, v. 89, n. 5, p. 1283-1284, mai. 2009.

KEMP, W. H. **Biodiesel Basics and Beyond: A Comprehensive Guide to Production and Use for the Home and Farm**. Ontario: Aztext Press, 2006. 300 p.

KENDALL-TACKETT, K. Long-Chain Omega-3 Fatty Acids and Women's Mental Health in the Perinatal Period and Beyond. **Journal of Midwifery & Women's Health**, London, v. 55, n. 6, p. 561-567, nov. 2010

KHATIB, K. A.; ARAMOUNI, F. M.; HERALD, T. J.; BOYER, J. E. Physicochemical characteristics of soft tofu formulated from selected soybean varieties. **Journal of Food Quality**, Ames, v. 25, n. 4, p. 289-303, out. 2002.

KNOTHE, G. Biodiesel Derived from a Model Oil Enriched in Palmitoleic Acid, Macadamia Nut Oil. **Energy & Fuels**, Washington, v. 24, n. 3, p. 2098-2103, fev. 2010.

KNOTHE, G. Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. **Fuel Processing Technology**, Oxford, v. 86, n. 10, p. 1059-1070, jun. 2005.

KWON, S. H.; TORRIE, J. H. Heritability and interrelationship among traits of two soybean populations. **Crop Science**, Madison, v. 4, n. 2, p. 194-198, mar./abr. 1964.

LACERDA, A. L. S.; LAZARINI, E.; SA, M. E.; VALERIO FILHO, W. V. Armazenamento de sementes de soja dessecadas e avaliação da qualidade fisiológica, bioquímica e sanitária. **Revista Brasileira de Sementes**, Pelotas, v. 25, n. 2, p. 97-105, dez. 2003.

LAM, M. K.; LEE, K. T.; MOHAMED, A. R. Life cycle assessment for the production of biodiesel: A case study in Malaysia for palm oil versus jatropha oil. **Biofuels, Bioproducts and Biorefining**, v. 3, n. 6, p. 601-612, nov./dez. 2009.

LANNAL, A. C.; JOSÉ, I. C.; OLIVEIRA, M. G. A.; BARROS, E. G.; MOREIRA, M. A. Effect of temperature on polyunsaturated fatty acid accumulation in soybean seeds. **Brazilian Journal of Plant Physiology**, Londrina, v. 17, n. 2, p. 213-222, abr./jun. 2005.

- LEE, S.; SHU, X.; LI, H.; YANG, G.; CAI, H.; WEN, W.; JI, B.; GAO, J.; GAO, Y.; ZHENG, W. Adolescent and adult soy food intake and breast cancer risk: results from the Shanghai Women's Health Study American. **Journal of Clinical Nutrition**, Bethesda, v. 89, n. 6, p. 1920-1926, jun. 2009.
- LEGENDRE, P.; LEGENDRE, L. **Numerical ecology**. 2 ed. Amsterdam: Elsevier, 1998. 853 p.
- LI, Z.; NELSON, R. L. Genetic diversity among soybean accessions from three countries measured by RAPDs. **Crop Science**, Madison, v. 41, n. 4, p. 1337-1347. jul./ago. 2001.
- LIMA, G. J. M. M. Importância da qualidade nutricional da soja e de seus subprodutos no mercado de rações: situação atual e perspectivas futuras. In: CONGRESSO BRASILEIRO DE SOJA. 1., **Anais...** Londrina: Embrapa Soja. 1999. p 165.
- LOGOZZO, G.; DONNOLI, R.; MACALUSO, L.; PAPA, R.; KNÜPFER, H.; ZEULI, P. S. Analysis of the contribution of Mesoamerican and Andean gene pools to European common bean (*Phaseolus vulgaris* L.) germplasm and strategies to establish a core collection. **Genetic Resources and Crop Evolution**, Dordrecht, v. 54, n. 8, p. 1763-1779, dez. 2007.
- LUSAS, E. W.; RIAZ, M. N. Soy protein products: processing and use. **Journal of Nutrition**, Bethesda, v. 125, supl. 3, p. 573S-580S, mar.1995.
- MAESTRI, D. M.; LABUCKAS, D. O.; MERILES, J. M.; LAMARQUE, A. L.; ZYGADLO, J. A.; GUZMÁN, C. A. Seed composition of soybean cultivars evaluated in different environmental regions. **Journal of the Science of Food and Agriculture**, Malden, v. 77, n. 4, p. 494-498, ago. 1998.
- MAHMOUD, A. A.; NATARAJAN, S. S.; BENNETT, J. O.; MAWHINNEY, T. P.; WIEBOLD, W. J.; KRISHNAN, H. B. Effect of Six Decades of Selective Breeding on Soybean Protein Composition and Quality: A Biochemical and Molecular Analysis. **Journal of Agricultural and Food Chemistry**, Washington, v. 54, n. 11, p. 3916-3922, mai. 2006.
- MALHOTRA, N. K. **Pesquisa de Marketing: uma orientação aplicada**. 4. ed. edição. Porto Alegre: Bookman, 2004.
- MALOSETTI, M.; ABADIE, T. Sampling strategy to develop a core collection of Uruguayan maize landraces based on morphological traits. **Genetic Resources and Crop Evolution**, Dordrecht, v. 48, n. 4, p. 381-390, ago. 2001.
- MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. **Multivariate Analysis**. London: Academic Press. 1979. 518 p.
- MAREGA FILHO, M.; DESTRO, D.; MIRANDA, L. A. Relationships among oil content, protein content and seed size in soybeans. **Brazilian Archives of Biology and Technology**, Curitiba, v. 44, n. 1, p. 23-32, mar. 2001.
- MARIM, B. G.; SILVA, D. J. H.; CARNEIRO, P. C. S.; MIRANDA, G. V.; MATTEDI, A. P.; CALIMAN, F. R. B. Variabilidade genética e importância relativa de caracteres em acessos de germoplasma de tomateiro. **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, n. 10, p. 1283-1290, out. 2009.

- MARIOT, M. P.; BARBIERI, R. L. Divergência genética entre acessos de espinheira-santa (*Maytenus ilicifolia* Mart. ex Reissek e *M. aquifolium* Mart.) com base em caracteres morfológicos e fisiológicos. **Revista Brasileira de Plantas Mediciniais**, Botucatu, v. 12, n. 3, p. 243-249, jul./set. 2010.
- MASCARENHAS, H. A. A.; TANAKA, R. T.; GALLO, P. B.; PEREIRA, J. C. V. N. A.; AMBROSANO, G. M. B.; CARMELLO, Q. A. C. Efeito da calagem sobre a produtividade de grãos, óleo e proteína em cultivares precoces de soja. **Scientia Agricola**, Piracicaba, v. 53, n. 1, p. 164-172, jan./abr. 1996.
- MASIERO, G.; LOPES, H. Etanol e biodiesel como recursos energéticos alternativos: perspectivas da América Latina e da Ásia. **Revista brasileira de política internacional**, Brasília v. 51, n. 2, jul/dez. 2008.
- MAYER, J. E.; PFEIFFER, W. H.; BEYER, P. Biofortified crops to alleviate micronutrient malnutrition. **Current Opinion in Plant Biology**, v. 11, p. 166-170, abr. 2008.
- MAZZA, M.; POMPONI, M.; JANIRI, L.; BRIA, P.; MAZZA, S. Omega-3 fatty acids and antioxidants in neurological and psychiatric diseases: An overview. **Progress in Neuro-Psychopharmacology and Biological Psychiatry**, London, v. 31, n. 1, p. 12-26, jan. 2007.
- McGARIGAL, K.; CUSHMAN, S.; STAFFORD, S. **Multivariate statistics for wildlife and ecology research**. New York: Springer/Verlag, 2000. 283 p.
- McQUITTY, L. L. Similarity Analysis by Reciprocal Pairs for Discrete and Continuous Data. **Educational and Psychological Measurement**, v. 26, n. 4, p. 825-831, 1966.
- MELLO FILHO, O. L.; SEDIYAMA, C. S.; MOREIRA, M. A.; REIS, M. S.; MASSONI, G. A.; PIOVESAN, N. D. Grain yield and seed quality of soybean selected for high protein content. **Pesquisa agropecuária brasileira**, Brasília, v. 39, n. 5, p. 445-450, mai. 2004.
- MILLER, J. E.; FEHR, W. R. Direct and indirect recurrent selection for protein in soybeans. **Crop Science**, v. 4, n. 2, p. 196-198, mar./abr. 1979.
- MILLIGAN, G. W.; COOPER, M. C. An Examination of Procedures for Determining the Number of Clusters in a Data Set. **Psychometrika**, Madison, v. 50, n. 159-179, dez. 1985.
- MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: UFMG, 2007. 295 p
- MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA E COMÉRCIO EXTERIOR. **O futuro da indústria: biodiesel**. Brasília: MDIC-STI/IEL, 2006. 145 p. Disponível em: <<http://www.biodiesel.gov.br>>. Acesso em: 10 jan. 2011.
- MINUZZI, A.; RANGEL, M. A. S.; LUCCA E BRACCINI, A.; SCAPIM, C. A.; MORA, F.; ROBAINA, A. D. Rendimento, teores de óleo e proteínas de quatro cultivares de soja, produzidas em dois locais no estado do Mato Grosso do Sul. **Ciência e Agrotecnologia**, Lavras, v. 33, n. 4, jul./ago. 2009.
- MOCELLIN, S.; ROSSI, C. R. Principles of Gene Microarray Data Analysis. In: MOCELLIN, S. (Ed.). **Microarray Technology and Cancer Gene Profiling**. New York: Landes Bioscience and Springer Science. 2006. cap. 3, p. 19-29.

MONTEIRO, J. B. R.; COSTA, N. M. B.; ESTEVES, E. A.; MILAGRES, K. H. Avaliação da qualidade protéica de dois formulados em pó, à base de soja enriquecidos com zinco, selênio e magnésio para utilização em nutrição enteral. **Ciência e Tecnologia de Alimentos, Campinas**, v. 24, n. 1, p. 6-10, jan./mar. 2004.

MORAES, R. M. A.; JOSÉ, I. C.; RAMOS, F. G.; BARROS, E. G.; MOREIRA, M. A. Caracterização bioquímica de linhagens de soja com alto teor de proteína. **Pesquisa Agropecuária Brasileira**, Brasília, v. 41, n. 5, mai. 2006.

MOREIRA, R. M. P.; FERREIRA, J. M.; TAKAHASHI, L. S. A.; VANCONCELOS, M. E. C.; GEUS, L. C.; BOTTI, L. Potencial agrônômico e divergência genética entre genótipos de feijão-vagem de crescimento determinado. **Ciências Agrárias**, Londrina, v. 30, supl. 1, p. 1051-1060, 2009.

MOURA, M. C. C. L. Algoritmo de Gower na estimativa da divergência genética em germoplasma de pimenta. **Horticultura Brasileira**, Brasília, v. 28, n. 2, p. 155-161, abr./jun. 2010.

MUFTI, G. B.; BERTRAND, P.; EL MOUBARKI, L. Determining the Number of Groups from Measures of Cluster Validity. In: International Conference on Applied Stochastic Models and Data Analysis, 2005, Brest. **Proceedings...** Brest, France: ASMDA, 2005. p. 404-414.

MUJOO, R.; TRINH, D. T.; NG, P. K. W. Evaluation of soybean varieties for soymilk and tofu production potential using laboratory-developed procedures. **Food Science and Biotechnology**, Seoul, v. 11, p. 470-476, 2002.

NAPIER, J. A. The Production of Unusual Fatty Acids in Transgenic Plants. **Annual Review of Plant Biology**, Palo Alto, v. 58, p. 295-319, jun. 2007.

NASS, L.L.; PELLICANO, I.J.; VALOIS, A.C.C. Utilization of genetic resources for maize and soybean breeding in Brazil. **Brazilian Journal of Genetics**, Ribeirão Preto, v.16, n. 4, p. 983-988, 1993.

NDZANA, X.; FEHR, W. R.; WELKE, G. A.; HAMMOND, E. G.; DUVICK, D. N.; CIANZIO, S. R. Influence of reduced palmitate content on agronomic and seed traits of soybean. **Crop Science**, Madison, v. 34, n. 3, p. 646-649, mai./jun. 1994.

NEGREIROS, J. R. S.; ALEXANDRE, R. S.; ÁLVARES, V. S.; BRUCKNER, C. H.; CRUZ, C. D. Divergência genética entre progênies de maracujazeiro- amarelo com base em características das plântulas. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 30, n. 1, mar. 2008.

NEPSTAD, D. C.; STICKLER, C. M.; ALMEIDA, O. T. Globalization of the Amazon Soy and Beef Industries: Opportunities for Conservation. **Conservation Biology**, Washington, v. 20, n. 6, p. 1595-1603, dez. 2006.

NESTEL, P.; BOUIS, H. E.; MEENAKSHI, J. V.; PFEIFFER, W. Biofortification of staple food crops. **Journal of Nutrition**, Bethesda, v. 136, n. 4, p. 1064-1067, abr. 2006.

NETO, F. V. B.; LEAL, N. R.; GONÇALVES, L. S. A.; RÊGO FILHO, L. M.; AMARAL JÚNIOR, A. T. Descritores quantitativos na estimativa da divergência genética entre genótipos de mamoneira utilizando análises multivariadas. **Revista Ciência Agronômica**, Fortaleza, v. 41, n. 2, p. 294-299, 2010.

NIELSEN, S. S. **Food analysis**. 2. ed. New York: Plenum, 2003. 292 p.

OHARA, M.; SHIMAMOTO, Y. Importance of genetic characterization and conservation of plant genetic resources: The breeding system and genetic diversity of wild soybean (*Glycine soja*). **Plant Species Biology**, Malden, v. 17, n.1, p. 51-58, abr. 2002.

OLIVEIRA, J. E. D. Uso da soja como feijão. In: MIYSAKA, S.; MEDINA, J.C. **A soja no Brasil**. Campinas: ITAL, 1981. cap. 4, p. 847-849.

OLIVEIRA, J. P. **Avaliação da qualidade nutricional do grão em populações de milho de alta qualidade protéica e seus cruzamentos**. 2003. Dissertação (Doutorado em Agronomia) - Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, 2003.

OLIVEIRA, M. F. **Avaliação de cinco estratégias de amostragem para a obtenção da coleção nuclear de soja (*Glycine max* (L.) Merrill)**. 2007. 143 f. Tese (Doutorado em Agronomia: genética e melhoramento de plantas) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2007.

OLIVEIRA, M. F.; NELSON, R. L.; GERALDI, I. O.; CRUZ, C. D.; TOLEDO J. F. F. Establishing a soybean germplasm core collection. **Field Crops Research**, New York, v. 119, n. 2-3, p. 277-289, nov./dez. 2010.

OLIVEIRA, M. S. P. Caracterização molecular e morfo-agronômica de germoplasma de Açaizeiro. 2005. 171 f. Tese (Doutorado em Agronomia: genética e melhoramento de plantas)-Universidade Federal de Lavras, Lavras, 2005.

OLIVEIRA, M. S. P.; FERREIRA, D. F.; SANTOS, J. B. Divergência genética entre acessos de açaizeiro fundamentada em descritores morfoagronômicos. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, n. 4, p. 501-506, abr. 2007.

ORTIZ, R.; RUIZ-TAPIA, E. N; MUJICA-SANCHEZ, A. Sampling strategy for a core collection of Peruvian quinoa germplasm. **Theoretical and Applied Genetics**, Berlin, v. 96, n. 3-4, p. 475-483, mar. 1998.

PAETAU, I.; CHEN, C.; JANE, J. Biodegradable Plastic Made from Soybean Products: Effect of Preparation and Processing on Mechanical Properties and Water Absorption. **Industrial and Engineering Chemistry Research**, Washington, v. 33, n. 7, p. 1821-1827, jul. 1994.

PANTHEE, D. R.; PANTALONE, V. R.; WEST, D. R.; SAXTON, A. M.; SAMS, C. E. Quantitative Trait Loci for Seed Protein and Oil Concentration, and Seed Size in Soybean. **Crop Science**, Madison, v. 45, n. 5, p. 2015-2022, set./out. 2005.

PEQUENO, S. A. **Parentesco e diversidade genética molecular em cultivares de soja recomendados para o estado de Goiás**. 2001. 98 f. Dissertação (Mestrado em Agronomia: genética e melhoramento de plantas) - Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2001.

PÉREZ, A. A.; DRAGO, S. R.; CARRARA, C. R.; DE GREEF, D. M.; TORRES, R. L. GONZÁLEZ, R. J. Extrusion cooking of a maize/soybean mixture: Factors affecting expanded product characteristics and flour dispersion viscosity. **Journal of Food Engineering**, New York, v. 87, n. 3, p. 333-340, ago. 2008.

PFEIFFER, W. H.; MCCLAFFERTY, B. HarvestPlus: Breeding Crops for Better Nutrition. **Crop Science**, Madison, v. 47, supl. 3, p. 88-105, dez. 2007.

PINTO, A. C.; GUARIEIRO, L. L. N.; REZENDE, M. J. C; RIBEIRO, N. M.; TORRES, E. A.; LOPES, W. A; PEREIRA, P. A. P.; ANDRADE, J.B. Biodiesel: An Overview. **Journal of the Brazilian Chemical Society**, São Paulo, v.16, n.6, nov./dez. 2005.

PINZI, S.; GARCIA, I. L.; LOPEZ-GIMENEZ, F. J.; LUQUE DE CASTRO, M. D.; DORADO, G.; DORADO, M. P. The Ideal Vegetable Oil-based Biodiesel Composition: A Review of Social, Economical and Technical Implications. **Energy & Fuels**, Washington, v. 23, n. 5, p. 2325-2341, abr. 2009.

PÍPOLO, A. E. **Influência da temperatura sobre as concentrações de proteína e óleo em sementes de soja (*Glycine max* (L.) Merrill)**. 2002. 128 f. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2002.

PONTI, L.; GUTIERREZ, A. P. Overview on biofuels from a European perspective. **Bulletin of Science Technology Society**, Thousand Oaks, v. 29, n. 6, p. 493-504, dez. 2009.

POTTER, S. M. Soy protein and serum lipids. **Current Opinion in Lipidology**, London, v. 7, n. 4, p. 260-264, ago. 1996.

PRIOLLI, R. H. G.; MENDES JUNIOR, C. T.; SOUSA, S. M. B.; SOUSA, N. E. A; CONTEL, E. P. B. Diversidade genética da soja entre períodos e entre programas de melhoramento no Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, v.39, n.10, out. 2004.

QUEIROZ, F. A. Impacts of the international trade of soybeans on the biodiversity of Cerrado Savannah. **Sociedade e Natureza**, Uberlândia, v. 21, n. 2, ago. 2009. Disponível em: <http://www.scielo.br/scielo.php?pid=S1982-45132009000200013&script=sci_arttext> Acesso em: 10 dez 2009.

QUEIROZ, M. A.; LOPES, M. A. Importância dos recursos genéticos para o agronegócio. In: NASS, L. L. (Ed.). **Recursos genéticos vegetais**. Brasília: Embrapa Recursos Genéticos e Biotecnologia. 2007. cap. 2, p. 281-305.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2010. Disponível em: <<http://www.R-project.org/>> Acesso em: 10 dez. 2010.

RAHMAN, S. M.; KINOSHITA, T.; ANAI, T.; TAKAGI, Y. Combining ability in loci for high oleic and low linolenic acids in soybean. **Crop Science**, Madison, v. 41, n. 1, p. 26-29, jan./fev. 2001.

RAHMAN, S. M.; TAKAGI, Y.; KUBOTA, K.; MIYAMOTO, K. High oleic acid mutant in soybean induced by x-ray irradiation. **Bioscience, biotechnology, and biochemistry**, Tokyo, v. 58, n. 6, p. 1070-1072, jul. 1994.

RAHMAN, S. M.; TAKAGI, Y.; KUMAMARU, T. Low linolenate sources at the fan locus in soybean lines M-5 and IL-8. **Breeding Science**, Tokyo, v. 46, p. 155-158. mar. 1996.

RAJAGOPAL, D.; ZILBERMAN, D. Review of Environmental, Economic and Policy Aspects of Biofuels. **Policy Research Working Papers Series**, Washington, n. 4341, set. 2007. Disponível em: < http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2007/09/04/000158349_20070904162607/Rendered/PDF/wps4341.pdf>. Acesso em: 12 ago. 2010.

RAMOS, L. P. Conversão de óleos vegetais em biocombustível alternativo ao diesel convencional. In: Congresso Brasileiro de Soja, Londrina. **Resumos...** Londrina: Embrapa, 1999. p. 233.

RAMOS, L.P.; DOMINGOS, A. K.; KUCEK, K. T.; WILHELM, H. M. Biodiesel: um projeto de sustentabilidade econômica e sócio-ambiental para o Brasil. **Biociência, Ciência e Desenvolvimento**, Brasília, v. 31, p. 28-37, jul./dez. 2003.

RANGEL, M. A. S.; CAVALHEIRO, L. R.; CAVICHIOLO, D.; CARDOSO, P. C. **Efeito do genótipo e do ambiente sobre os teores de óleo e proteína nos grãos de soja, em quatro ambientes da Região Sul de Mato Grosso do Sul, safra 2002/ 2003**. Dourados: Embrapa Agropecuária Oeste, 2004. 20 p. (Boletim de pesquisa e desenvolvimento, 17).

RAO, R. V. Characterization and evaluation of plant genetic resources. In: ZANGH, Z.; ZHOU, M.; RAO, R. V. (Ed.). **Plant genetic resources network in East Asia**. Ulaanbaatar : Proceedings of the meeting for the regional network for conservation and use of plant genetic resources in East Asia, 2001. p. 89-99.

REBETZKE, G. J.; BURTON, J. W.; CARTER, T. E.; WILSON, R. F. Changes in agronomic and seed characteristics with selection for reduced palmitic acid content in soybean. **Crop Science**, Madison, v. 38, n. 2, p. 297-302, mar./abr. 1998.

RIBEIRO, F. E.; SOARES, A. R.; RAMALHO, M. A. P. Divergência genética entre populações de coqueiro-gigante-do-Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, v. 34, n. 9, p. 1615-1622, set. 1999.

RIBEIRO, K. O. **Heterose entre linhagens de soja quanto aos teores de proteína e de óleo no grão**. 2006. 68 f. Dissertação (Mestrado em Agronomia: genética e melhoramento de plantas) - Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2006.

RIBEIRO, R. M. A. Glossário de termos de coleta e conservação de recursos genéticos. **Ciência da Informação**, Brasília: IBICT, v. 24, n. 3, 7 p. 1995. Disponível em: < revista.ibict.br/index.php/ciinf/article/download/499/453. Acesso em: 12 maio 2011.

ROMESBURG, H. C. **Cluster Analysis for Researchers**. Morrisville: Lulu Press, 2004. 334 p.

RONGWEN, J.; AKKAYA, M.S.; BHAGWAT, A.A.; LAVI, U.; CREGAN, P.B. The use of microsatellite DNA markers for soybean genotype identification. **Theoretical and Applied Genetics**, Berlin, v.90, n.1, 1995.

ROSA, A. M.; CLAVISO, J.; PASSOS, L. M. L.; AGUIAR, C. L. Alimentos fermentados à base de soja (*Glycine max* (Merrill) L.): importância econômica, impacto na saúde e efeitos associados às isoflavonas e seus açúcares. **Revista Brasileira de Biociências**, Porto Alegre, v. 7, n. 4, out./dez. 2009.

SANIBAL, E. A. A.; MANCINI-FILHO, J. Perfil de ácidos graxos trans de óleo e gordura hidrogenada de soja no processo de fritura. *Ciência e Tecnologia de Alimentos*, Campinas, v. 24, n. 1, p. 27- 31, jan./mar. 2004.

SANTOS, A. L.; SAKOMURA, N. K.; FREITAS, E. R.; FORTES, C. M. L. S.; CARRILHO, E. N. V. M.; FERNANDES, J. B. K. Estudo do crescimento, desempenho, rendimento de carcaça e qualidade de carne de três linhagens de frango de corte. **Revista Brasileira de Zootecnia**, Viçosa, v. 34, n. 5, out. 2005.

SAS INSTITUTE INC. **SAS/STAT® 9.22 User's guide**. Cary: SAS Institute Inc., 2010. 8444 p. Disponível em < <http://support.sas.com/documentation/onlinedoc/stat/#stat922>>. Acesso em: 01 jan. 2010.

SCHMUTZ, J.; CANNON, S. B.; SCHLUETER, J.; MA, J.; MITROS, T.; NELSON, W.; HYTEN, D. L.; SONG, Q.; THELEN, J. J.; CHENG, J.; XU, D.; HELLSTEN, U.; MAY, G. D.; YU, Y.; SAKURAI, T.; UMEZAWA, T.; BHATTACHARYYA, M. K.; SANDHU, D.; VALLIYODAN, B.; LINDQUIST, E.; PETO, M.; GRANT, D.; SHU, S.; GOODSTEIN, D.; BARRY, K.; FUTRELL-GRIGGS, M.; ABERNATHY, B.; DU, J.; TIAN, Z.; ZHU, L.; GILL, N.; JOSHI, T.; LIBAULT, M.; SETHURAMAN, A.; ZHANG, X.; SHINOZAKI, K.; NGUYEN, H. T.; WING, R. A.; CREGAN, P.; SPECHT, J.; GRIMWOOD, J.; ROKHSAR, D.; STACEY, G.; SHOEMAKER, R. C.; JACKSON, S. A. Genome sequence of the palaeopolyploid soybean. **Nature**, London, v. 463, p.178-183, jan. 2010.

SCOTT, R. A.; KEPHART, K. D. Selection for yield, protein, and oil in soybean crosses between adapted and introduced parents. **Field Crops Research**, New York, v. 49, n. 2-3, p. 177-185, fev. 1997.

SEDIYAMA, T.; ATHOW, K. L.; SEDIYAMA, C. S.; REIS, M. S.; ARANTES, N. E. UFV-3, variedade de soja para o Norte de Minas Gerais. **Revista Ceres**, Viçosa, v. 27, n. 149, p. 91-95, fev. 1980.

SEDIYAMA, T.; TEIXEIRA, R. C.; REIS, M. S. Melhoramento da soja. In: BORÉM, A. (Ed.). **Melhoramento de Espécies Cultivadas**. 2 ed. Viçosa: Universidade Federal de Viçosa, 2005. cap. 14, p. 553-604.

SESLI, M.; YEGENOGLU, E. D. Compare various combinations of similarity coefficients and clustering methods for *Olea europaea sativa*. **Scientific Research and Essays**, Lagos, v. 5, n. 16, p. 2318-2326, ago. 2010.

SERRAJ, R.; SINCLAIR, T. R.; PURCELL, L. C. Symbiotic N₂ fixation response to drought. **Journal of Experimental Botany**, v. 50, n. 331, p. 143-155, set. 1999.

- SHORTER, R.; BYTH, D.E.; MUNGOMERY, V. E. Estimates of selection parameters associated with protein and oil content of soybean seeds (*Glycine max* (L.) Merr.). **Australian Journal of Agricultural Research**, New Jersey, v. 28, n. 2, p. 211-222, fev. 1976.
- SILVA, D. J. H.; MOURA, M. C. C. L.; CASALI, V. W. D. Genetic resources of the vegetable germplasm bank at the UFV, Brazil: historical background and assessment. **Horticultura Brasileira**, Brasília, v.19, n.2, jul. 2001.
- SILVA, H. T.; COSTA, A. O. **Caracterização botânica de espécies silvestres do gênero *Phaseolus* L. (Leguminosae)**. Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2003. 40 p. (Documentos, 156).
- SILVA, J A.; SALOMÃO, A.N. **Banco germoplasma de espécies florestais nativas do Campo Experimental Sucupira - Ipê Amarelo (*Tabebuia áurea* (Silva Manso) Benth. & Hook. f. ex. S. Moore) Bignoniaceae**. Brasília: Embrapa Recursos Genéticos e biotecnologia, 2006. 22 p. (Documentos, 181).
- SILVA, M. H. L. **Desenvolvimento e caracterização de um isolado protéico de soja modificado com perfil de solubilidade da caseína do leite humano**. 2007. Tese (Doutorado em Ciência e Tecnologia em Alimentos)–Universidade Federal de Viçosa, Viçosa. 2007
- SILVA, T. M.; MUNHOZ, R. P.; ALVAREZ, C.; NALIWAIKO, K.; KISS, A.; ANDREATINI, R.; FERRAZ, A. C. Depression in Parkinson's disease: A double-blind, randomized, placebo-controlled pilot study of omega-3 fatty-acid supplementation. **Journal of Affective Disorders**, London, v. 111,n. 2-3, p. 351-359, dez. 2008.
- SILVA, W. C. J. **Escolha de descritores mínimos e estabelecimento de coleções nucleares em *Capsicum* spp.** 2008. 138 f. Tese (Doutorado em Agronomia: genética e melhoramento de plantas) - Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2008.
- SIMOPOULOS, A. P. Importance of the Omega-6/Omega-3 Essential Fatty Acids: Evolutionary Aspects. In: SIMOPOULOS, A.P.; CLELAND, L. G. (Ed.). **Omega-6/omega-3 essential fatty acid ratio: the scientific evidence**. Suíça: World Review of Nutrition and Dietetics, 2003. v. 92, cap. 1, p. 1-23.
- SIMOPOULOS, A.P. The Importance of the Omega-6/Omega-3 Fatty Acid Ratio in Cardiovascular Disease and Other Chronic Diseases. **Experimental Biology and Medicine**, London, v. 233, n. 6, p. 674-688, abr. 2008.
- SIMPSON JUNIOR, A. M.; WILCOX, J. R. Genetic and phenotypic associations of agronomic characteristics in four high protein soybean populations. **Crop Science**, Madison, v. 23, n.6, p. 1077-1081, nov./dez. 1983.
- SINGH, P.; KUMAR, R.; SABAPATHY, S. N.; BAWA, A. S. Functional and Edible Uses of Soy Protein Product. **Comprehensive Reviews in Food Science and Food Safety**, Malden, v. 7, n. 1, p. 14-28, jan. 2008.
- SIZER, F. S.; WHITNEY, E. **Nutrition: Concepts and Controversies**. 11. ed. Belmont: Thomson Wadsworth, 2008. 574 p.

- SMITH, R. R.; WEBER, C. R. Mass Selection by Specific Gravity for Protein and Oil in Soybean **Crop Science**, Madison, v. 8, n. 3, p. 373-377, mai./jun.1968.
- SOKAL, R. R.; MICHENER, C. D. A Statistical Method for Evaluating Systematic Relationships. **University of Kansas Science Bulletin**, Lawrence, v. 38, p. 1409-1438, 1958.
- SOKAL, R. R.; ROHLF, F. J. The comparison of dendrograms by objective methods. *Taxon*, Utrecht, v. 11, n. 2, p. 30-40, fev. 1962.
- SOUZA, P. J. O. P.; RIBEIRO, A.; ROCHA, E. J. P.; LOUREIRO, R. S.; BISPO, C. J. C.; SOUSA, A. M. L. Albedo of soybean crop in the frontier region of the Amazon. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 14, n. 1, Jan. 2010.
- STOJSIN, D.; ABLETT, G. R.; LUZZI, B. M.; TANNER, J. W. Use of gene substitution values to quantify partial dominance in low palmitic acid soybean. **Crop Science**, Madison, v. 38, n. 6, p. 1437-1441, nov./dez. 1998.
- SUDRÉ, C.P.; RODRIGUES, R.; RIVA, E.M.; KARASAWA, M.; AMARAL JÚNIOR, A.T. Divergência genética entre acessos de pimenta e pimentão utilizando técnicas multivariadas. **Horticultura Brasileira**, Brasília, v. 23, n. 1, p. 22-27, jan/mar. 2005.
- TAKAGI, Y.; RAHAMAN, S. M.; JOO, H.; KAWAKITA, T. Reduced and elevated palmitic acid mutants in soybean developed by X-ray irradiation. **Bioscience, biotechnology, and biochemistry**, Tokyo. v. 59, n. 9, p. 1788-1797, 1995.
- TAKAGI, Y.; RAHMAN, S. M. Inheritance of high oleic acid content in the seed oil of soybean mutant M23. **Theoretical and Applied Genetics**, Berlin, v. 92, n. 2, fev. 1996.
- TANGO, J.S.; TURATTI, J.M.; SILVA, M.T.C.; SHIROSE, I.; YOTSUYANAGI, K. Influência do local de plantio no teor e características do óleo de algumas cultivares de soja. **Boletim do Instituto de Tecnologia de Alimentos**, Campinas, v.20, n.4, p.355-374, 1983.
- TASHIMA, E. H.; CARDELLO, H. M. A. B. Perfil sensorial de extrato hidrossolúvel de soja (*Glycine max* L. Merrill) comercial adoçado com sacarose e com sucralose. **Boletim do Centro de Pesquisa de Processamento de Alimentos**, Curitiba, v. 21, n 2, fev 2003.
- TERRAZAS, M.; CIVERA, R.; IBARRA, L.; GOYTORTÚA, E. Apparent digestion coefficients for dry matter, protein and essential amino acids in terrestrial ingredients for Pacific shrimp *Litopenaeus vannamei* (Decapoda: Penaeidae). **Revista de Biología Tropical**, San José, v. 58, n. 4, p. 1561-1576, dez. 2010.
- THOENES, P. **Biofuels and Commodity Markets – Palm Oil Focus**. Rome: FAO Commodities and Trade Division, 2006. 15 p.
- THORNE, J. C.; FEHR, W. R. Incorporation of High-Protein, Exotic Germplasm into Soybean Populations by 2- and 3-way Crosses. **Crop Science**, Madison, v. 10, n. 6, p. 652-655, set./out. 1970.
- TILMAN, D. Global environmental impacts of agricultural expansion: The need for sustainable and efficient practices. **Proceedings of the National Academy of Sciences**, Washington, v. 96, n. 11, p. 5995-6000, mai. 1999.

- TOPFER, R.; MARTINI, N.; SCHELL, J. Modification of plant lipid synthesis. **Science**, Washington, v. 268, n. 5211, p. 681-686, mai. 1995.
- TYSON, K. S.; McCORMICK, R. L. **Biodiesel Handling and Use Guidelines**. 3. ed. Golden: National Renewable Energy Laboratory, 2006. Disponível em: <<http://www.osti.gov/bridge>>. Acesso em: 12 de dezembro de 2010.
- UNITED STATES DEPARTMENT OF AGRICULTURE. **Commodity Data and Statistics- Soybean Oil: World Supply and Distribution**. Washington: USDA, 2010. Disponível em: <<http://www.fas.usda.gov/psdonline/psdHome.aspx> >. Acesso em: 20 jan. 2010.
- VALLS, J. F. M. Caracterização de recursos genéticos vegetais. In: NASS, L. L. (Ed.). **Recursos genéticos vegetais**. Brasília: Embrapa Recursos Genéticos e Biotecnologia, 2007. cap. 8, p. 281-342.
- VAN HINTUM, TH. J. L. V.; BROWN, A. H. D.; SPILLANE, C.; HODGKIN, T. **Core collection of plant genetic resources**. Rome: IPGRI, 2000. 48 p. (IPGRI Technical Bulletin, 3).
- VAN SLOTEN, D. H. The role of curators, breeders and others users of germplasm in characterization and evaluation of crop genetic resources. **IBPGR/SEAN**, Roma, Special Issue, p. 3-8, 1987.
- VASCONCELOS, E. S.; CRUZ, C. D.; BHERING, L. L.; FERREIRA, A. Estratégias de amostragem e estabelecimento de coleções nucleares. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, n.4, p. 507-514, abr. 2007.
- VELLO, N.A.; HIROMOTO, D.M.; AZEVEDO FILHO, A.J.B.V. Coefficient of parentage and breeding of Brazilian soybean germplasm. **Revista Brasileira de Genética**, Ribeirão Preto, v. 11, n. 3, p. 679-697, jul. 1988.
- VENCOVSKY, R.; BARRIGA, P. **Modelos biométricos no fitomelhoramento**. Sociedade Brasileira de Genética: Ribeirão Preto, 1992. 496p.
- VIDIGAL, M. C. G.; VIDIGAL FILHO, P. S.; AMARAL JUNIOR, A. T.; LUCCA E BRACCINI, A. Divergência genética entre cultivares de mandioca por meio de estatística multivariada. **Bragantia**, Campinas, v. 56, n. 2, p. 263-271, mar. 1997.
- VIEIRA, C.R.; CABRAL, L.C.; PAULA, A.C.O. Composição centesimal e conteúdo de aminoácidos, ácidos graxos e minerais de seis cultivares de soja destinadas à alimentação humana. **Pesquisa Agropecuária Brasileira**, Brasília, v.34, n.7, p. 1277-1283, jul. 1999.
- VISENTAINER, J.V.; FRANCO, M.R.B. **Ácidos graxos em óleos e gorduras: identificação e quantificação**. São Paulo: Varela, 2006. 120 p.
- VOLDENG, H. D.; COBER, E. R.; HUME, D.J. ; GILLARD, C.; MORRISON, M. J. Fifty-eight years of genetic improvement of short-season soybean cultivars in Canada. **Crop Science**, Madison, v.37, n. 2, p.428-431, mar./abr. 1997.

WANG, L.; GUAN, Y.; GUAN, R.; LI, Y.; MA, Y.; DONG, Z.; LIU, X.; ZHANG, H.; ZHANG, Y.; LIU, Z.; CHANG, R.; XU, H.; LI, L.; LIN, F.; LUAN, W.; YAN, Z.; NING, X.; ZHU, L.; CUI, Y.; PIAO, R.; LIU, Y.; CHEN, P.; QIU, L. Establishment of Chinese soybean *Glycine max* core collections with agronomic traits and SSR markers. **Euphytica**, Wageningen, v. 151, n. 2, p. 215-223. set. 2006.

WEHRMANN, V. K.; FEHR, W. R.; CIANZIO, S.R.; CAVINS, J.F. Transfer of high seed protein to high-yielding soybean cultivars. **Crop Science**, Madison, v. 27, n. 5, p. 927-931, set./out. 1987.

WENG, T. M.; CHEN, M. T. Changes of protein in natto (a fermented soybean food) affected by fermenting time. **Food Science and Technology Research**, Seoul, v. 16, n. 6, p. 537-542, 2010.

WETZEL, M. M. V. S.; FERREIRA, F.R. Sistema de curadorias de germoplasma. In: NASS, L. L. **Recursos Genéticos Vegetais**. Brasília: Embrapa, 2007. cap. 3, p. 121-141.

WILCOX, J. R.; CAVINS, J. F. Backcrossing high seed protein to a soybean cultivar. **Crop Science**, Madison, v. 35, n. 4, p. 1036-1041, jul./ago. 1995.

WILCOX, J. R.; CAVINS, J. F. Inheritance of low linolenic acid content of the seed oil of a mutant in *Glycine max*. **Theoretical and Applied Genetics**, Berlin, v. 71, n. 1, p. 74-78, nov. 1985.

WILCOX, J. R.; CAVINS, J. F. Registration of C1726 and C1727 soybean germplasm with altered levels of palmitic acid. **Crop Science**, Madison, v. 30, n. 1, p. 240, jan./fev. 1990.

WILCOX, J.R. Increasing seed protein in soybean with eight cycles of recurrent selection. **Crop Science**, Madison, v. 38, n. 6, p. 1536-1540, nov./dez. 1998.

WILCOX, J.R.; GUODONG, Z. Relationship between seed yield and seed protein in determinate and indeterminate soybean populations. **Crop Science**, Madison, v. 37, n. 2, p. 361-364, mar./abr. 1997.

WILSON, R. F. Outlook for soybeans and soybean products in 21st century markets. **Lipid Technology**, Weinheim, v. 22, n. 9, p. 199-202, set. 2010.

WILSON, R. F. Seed composition. In: BOERMA, H. R.; SPECHT, J. E. (Ed.). **Soybeans: Improvement, Production, and Uses**. 3. ed. Madison: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, 2004. cap. 13, p. 621-677.

WYSMIERSKI, P. T. **Contribuição genética dos ancestrais da soja às cultivares brasileiras**. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2010.

XU, H.; WILCOX, J. R. Recurrent selection for maturity and percent seed protein in *Glycine max* based on S₀ plant evolutions. **Euphytica**, Wageningen, v. 62, n. 1, p. 52-57, jan. 1992.

YAKLICH, R. W.; DAUN, J. K.; DeCLERCQ, D. R. Historical trends in the composition of soybean, canola-rape, and flax seeds grown in the United States and Canada. **Journal of the American Oil Chemists' Society**, Urbana, v. 79, n. 2, p. 207-208, fev. 2002.

YAO, Q.; FAN, P.; ZOU, S. Constructing a Core Collection for Maize (*Zea mays* L.) Landrace from Wuling Mountain Region in China. **Agricultural Sciences in China**, Beijing, v. 7, n. 12, p.1423-1432 dez. 2008.

YU, L.; YUNSU, S.; YONGSHENG, C.; WANG, T. Establishment of a core collection for maize germplasm preserved in Chinese National Genebank using geographic distribution and characterization data. **Genetic Resources and Crop Evolution**, Dordrecht, v. 51, n. 8, p. 845-852, jan. 2004.

ZARKADAS, C. G.; GAGNON, C.; GLEDDIE, S.; KHANIZADEH, S.; COBER, E. R.; GUILLEMETTE, R. J. D. Assessment of the protein quality of fourteen soybean [*Glycine max* (L.) Merr.] cultivars using amino acid analysis and two-dimensional electrophoresis. **Food Research International**, Ontario, v. 40, n. 1, p.129-146, jan. 2007.

ZEVEN, A.C.; WANINGE, J.; VAN HINTUM, TH.; SINGH, S.P. Phenotypic variation in a core collection of common bean (*Phaseolus vulgaris* L.) in the Netherlands. **Euphytica**, Wageningen, v. 109, n. 2, p. 93-106, set. 1999.

ZHANG, D.; PIRTLE, I. L.; PARK, S. J.; NAMPAISANSUK, M.; NEOGI, P.; WANJIE, S. W.; PIRTLE, R.M.; CHAPMAN, K. D. Identification and expression of a new delta-12 fatty acid desaturase (*FAD2-4*) gene in upland cotton and its functional expression in yeast and *Arabidopsis thaliana* plants. **Plant Physiology and Biochemistry**, Oxford, v. 47, n. 6, p. 462-471, jun. 2009.

ZHAO, L.; DONG, Y.; LIU, B.; HAO, S.; WANG, K.; LI, X. Establishment of a Core Collection for the Chinese annual wild soybean. **Chinese Science Bulletin**, Beijing, v. 50, n. 10, p. 989-996, mai. 2005.

APÊNDICES

- Apêndice A.** Relação de 527 acessos de soja pertencentes à coleção mantida pela Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), com informações¹ de genealogia, ciclo de maturação (CM) e teores médios (g/100 g) de óleo (OL), proteína (PROT) e dos ácidos graxos palmítico (PAL), esteárico (EST), oleico (OLC), linoleico (LIN) e linolênico (LIL), com a respectiva classificação em grupos (G) e subgrupos (SG), segundo análise de agrupamento pelo método hierárquico aglomerativo com critério de ligação UPGMA..... 125
- Apêndice B.** Análise de ácidos graxos em óleo de soja, em cromatógrafo gasoso (Focus GC Master), com identificação dos ácidos palmítico, esteárico, oleico, linoleico e linolênico e os respectivos valores de porcentagens e do tempo de retenção de cada ácido (éster metílico) pelo capilar..... 142
- Apêndice C.** Rotina computacional para implementação das análises estatísticas no aplicativo R, relativas ao estudo de caracterização de 527 acessos do banco de germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), acerca de sete atributos relacionados à qualidade do grão (teor de óleo – OL, teor de proteína – PROT, teores dos ácidos graxos palmítico – PAL, esteárico – EST, oleico – OLC, linoleico – LIN e linolênico – LIL, e teores totais dos ácidos saturados – SAT e poli-insaturados – POLI) 143
- Apêndice D.** Identificação dos grupos (A, B, C, D e E) e subgrupos de similaridade de 527 acessos do germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), estabelecidos pelo agrupamento via UPGMA em relação à sete variáveis relacionadas à qualidade do grão..... 146
- Apêndice E.** Detalhamento da similaridade ou divergência entre os acessos de soja do grupo A (426 acessos), com subdivisão nos subgrupos A1, A2 ... A8, estabelecidos pelo agrupamento via UPGMA da subcoleção de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO) em relação à sete variáveis relacionadas à qualidade do grão..... 147
- Apêndice F.** Detalhamento da similaridade ou divergência entre os acessos de soja do grupo B (92 acessos), com subdivisão nos subgrupos B1, B2 ... B5, estabelecidos pelo agrupamento via UPGMA da subcoleção de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO) em relação à sete variáveis relacionadas à qualidade do grão..... 150
- Apêndice G.** Detalhamento da similaridade ou divergência entre os acessos de soja do grupo C (6 acessos), com subdivisão nos subgrupos C1 e C2; e dos grupos D e E, que não tiveram subdivisão, estabelecidos pelo agrupamento via UPGMA da subcoleção de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO) em relação à sete variáveis relacionadas à qualidade do grão..... 151

Apêndice A. Relação de 527 acessos de soja pertencentes à coleção mantida pela Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), com informações¹ de genealogia, ciclo de maturação (CM) e teores médios (g/100 g) de óleo (OL), proteína (PROT) e dos ácidos graxos palmítico (PAL), esteárico (EST), oleico (OLC), linoleico (LIN) e linolênico (LIL), com a respectiva classificação em grupos (G) e subgrupos (SG), segundo análise de agrupamento pelo método hierárquico aglomerativo com critério de ligação UPGMA.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
1	EMGOPA-316	EMGOPA-302 x FT-79-2564	P	38,2	17,0	11,4	2,8	26,8	49,7	6,4	B	B1
2	EMGOPA-302	Parana x Mandarinim	P	38,6	16,7	11,7	1,9	24,1	53,9	7,0	A	A2
3	BRS GO-204 (GOIÂNIA)	[BR-13 (3) x (BR-16 x Ss-1)] x Braxton	M	39,0	16,3	11,9	2,5	22,2	54,6	7,1	A	A2
4	EMGOPA-315	Dourados (2) x (Amambai (2) x Ss-1)	M	38,7	16,3	12,5	1,9	26,0	52,4	6,0	A	A1
5	MGBR-46 CONQUISTA	LO 76-4484 x Numbaira	M	40,0	17,0	12,1	2,1	25,3	52,4	6,5	A	A2
6	EMGOPA-314 (GARÇA BRANCA)	EMGOPA-301 (6) x EMBRAPA-20 (Doko RC)	T	39,0	16,6	12,1	1,9	22,5	55,3	6,9	A	A1
7	LA SUPREMA	[Hardee x (Hill x PI-274454)] x Cristalina	P	37,4	17,6	11,1	1,9	27,5	52,1	5,9	A	A1
8	ST93-2051	-	M	37,5	18,2	10,9	2,0	26,3	54,3	5,9	A	A1
10	EMGOPA-308 BELA VISTA	EMGOPA-308 (6) x BR92-31914	T	40,0	16,5	12,2	2,0	26,0	52,1	6,2	A	A1
11	EMGOPA-313 JATAÍ	[IAC-7 x (Santa Rosa x GO79-3068)] x BR92-31910	T	37,8	16,4	12,2	2,2	20,4	57,7	6,7	A	A1
12	NOVA IAC-7	Selecao na IAC-7 (GO83-2031)	T	36,9	17,3	11,2	2,2	27,9	52,2	5,7	A	A1
13	EMGOPA-306 (CATALÃO)	(EMGOPA-308 (6) x BR92-31914) x (E-306(6) x BR92-31910)	T	38,4	16,1	12,5	2,0	19,9	56,8	7,1	A	A1
14	FT-2000	-	M	39,5	17,2	12,2	2,4	24,1	53,4	6,4	A	A1
15	CSM-1	-	M	40,5	16,7	10,5	1,8	44,5	36,2	5,3	C	C1
17	FT-109	Bulk de 18 cruzamentos	P	39,0	16,6	12,4	2,1	25,2	52,9	5,7	A	A1
19	CRISTALINA	(FT-Cristalina (5) x Doko RCH)	T	38,4	16,5	11,5	1,9	23,1	55,3	6,8	A	A1
20	BRS-CARLA	BR-16 x BR83-147	M	39,3	16,6	12,4	2,2	26,9	50,0	6,3	A	A2
21	BRS-MILENA	FT-Ayara x BR83-147	M	38,6	17,6	11,7	2,1	26,1	53,1	5,8	A	A1
22	PINTADO	Sharkey x {Hartwig x [BR87-567 (3) x FT-Estrela]}	T	39,4	17,1	11,6	1,7	31,6	47,8	6,3	B	B1
23	BRS-CELESTE	Bossier x BR-1 T	T	39,2	16,7	12,3	3,2	20,5	57,2	6,3	A	A3
24	IAC-15	-	P	40,1	15,6	12,9	1,6	21,6	55,3	7,1	A	A1
25	GOBR95-121119	(BR-16 (5) x SS-1) x BR-37-55	M	38,3	17,0	12,6	2,2	20,7	56,3	6,8	A	A1
26	GOBR94-122243	LEFLORE x {Hartwig x [BR87-567 (3) x FT-Estrela]}	M	39,6	16,7	12,1	2,2	24,7	52,7	6,6	A	A2
27	GOBR97-055009	Stone Wall x BR93-32041	M	37,5	16,0	11,3	2,2	23,6	55,3	6,4	A	A1
29	GOBR95-17538	FT-2 x (BR80-6989 x IAC-12) x Ayara	T	38,7	16,6	12,3	2,0	21,9	54,7	7,1	A	A2
30	GOBR95-12218	OCEPAR-8 x FT-Estrela	T	40,0	16,9	13,2	2,3	23,1	53,5	7,2	A	A2

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
31	GOBR97-058163	Sharkey *2 x BR93-32043	T	39,7	16,8	11,7	1,8	31,1	48,2	6,1	B	B1
32	GOBR97-64258	MTBR94-20117 x Sharkey	T	39,1	16,1	12,0	1,9	28,8	50,9	5,8	A	A1
33	GOBR97-065029	Hartwig x MTBR94-18190	M	38,0	16,9	12,2	2,2	24,1	53,7	6,8	A	A2
34	GOBR97-067033	MTBR94-20514 x Sharkey	M	39,4	16,7	11,5	1,9	20,7	51,1	6,3	A	A2
35	GOBR97-067004	MTBR94-20514 x Sharkey	P	40,5	15,4	11,8	1,9	23,8	54,2	6,5	A	A1
36	GOBR95-3412	OCEPAR-8 x FT-5	P	38,6	17,0	12,9	2,2	23,3	53,3	6,3	A	A1
37	GOBR95-6018	BR-38 x FT-5 RNC	P	38,0	16,6	12,2	2,2	26,0	51,6	5,9	A	A1
41	GOBR95-2019	MGBR22 x EMGOPA-304	P	40,6	16,3	11,7	2,3	25,7	52,2	6,2	A	A1
43	BR95-15305	FT-Guaira x FT-Estrela	T	39,3	16,9	11,9	2,0	35,1	42,4	6,8	B	B3
44	BR92-249	CPAC	T	38,5	16,9	11,2	2,1	30,7	49,1	6,1	B	B1
47	GOBR95-2852	CPAC	T	37,6	16,7	12,0	2,4	21,3	55,7	7,0	A	A2
48	GOBR95-69943	(OC-8 x BR-13) x (Doko x PI-229358)	T	38,9	16,5	11,9	2,2	22,8	55,1	6,6	A	A1
50	GOBR95-3734	Cristalina x FT-Estrela	T	39,3	17,3	12,9	3,6	22,7	52,7	6,1	A	A3
51	GOBR94-09543-00.02	FT-Manaca x BR83-147-B	P	38,7	16,0	17,4	2,1	21,6	50,3	7,2	D	D
52	GOBR94-11191	FT-Ayara x BR83-147-B	P	39,1	16,1	12,6	2,7	20,1	55,4	7,0	A	A1
53	GOBR95-13848	BR83-9547 x (BR87-567 (3) x BR-9 x Bossier)	T	38,9	16,9	12,0	2,2	23,5	54,7	6,6	A	A1
54	GOBR93-1483 (LUZLÂNIA)	Braxton x (FT-5 x Dourados-1 (5) x Ss-1)	M	39,7	16,8	11,0	2,0	33,1	47,2	5,9	B	B1
56	GOBR95-57007	T89-46 R (Bulk Percevejo)	T	37,1	16,7	11,9	2,0	25,1	53,5	6,6	A	A2
59	GOBR95-3622	OCEPAR-3 x Doko RC	M	37,8	17,3	12,2	2,6	27,6	50,4	6,1	A	A2
60	GOBR91-11779	BR-16 x BR83-147	P	36,9	17,4	12,1	2,7	21,3	55,5	6,3	A	A1
61	GOBR95-54105	Doko x IAC74-2832	M	39,4	16,5	12,2	2,0	20,4	57,3	7,3	A	A1
62	SANTA CRUZ	FT-14 x [Deus-2 (2) x Ss-1]	M	38,4	16,8	11,8	2,0	26,6	52,5	6,2	A	A1
63	MGBR95-19125	-	T	38,6	16,9	11,7	2,0	24,5	55,2	6,0	A	A1
65	CSST-1	-	M	39,1	16,3	12,0	2,4	21,2	56,2	6,7	A	A1
66	GOBR93-8220	IAC-13 x BR-16	M	39,3	15,9	12,8	2,1	21,2	56,2	6,8	A	A1
67	GOBR95-15956	BR-37 x 555 x BR90-7013	P	38,1	17,1	11,4	2,2	28,2	50,5	6,2	A	A2
68	GOBR95-12250	OCEPAR-8 x FT-Estrela	T	38,4	16,9	12,2	3,3	20,4	54,7	5,9	A	A3
69	GOBR97-056184	LEFLORE *4 x BR90-7057	T	38,0	16,9	11,5	1,8	28,8	50,9	6,0	A	A2
70	GOBR97-057130	Sharkey* (3) x BR92-31861	M	39,8	17,0	11,3	1,8	30,3	49,7	6,2	A	A2
71	MGBR95-964	Braxton x [BR13*3 x (BR16 x SS-1)]	M	38,4	16,9	11,8	2,2	25,5	52,3	7,2	B	B1
73	BRSMG-LIDERANÇA	Centennial x [(Parana x Bossier) x Davis-1]	P	41,0	16,1	10,9	2,1	25,6	53,4	6,0	A	A1
75	EMGOPA-306	Cruzamento com EMGOPA-301	T	38,9	16,4	12,4	2,2	19,1	57,7	7,6	A	A1
76	GOBR95-12895	FT-2 x (BR80-6989 x Tracy-M)	M	38,3	15,8	12,1	3,1	23,9	51,1	6,2	A	A2
77	GOBR00-622015	Hartwig x MTBR95-121319	M	39,8	16,3	12,6	2,1	23,0	54,4	6,5	A	A1

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
78	EMGOPA-311	Parana Goiana x Cristalina	T	38,5	16,6	12,2	1,9	23,7	55,4	6,4	A	A1
79	SAMBAIBA	FT-5 x [Dourados (4) x Ocepar-9 SS-1]	T	38,2	17,0	12,1	2,0	24,0	54,5	6,4	A	A1
80	MONSOY-6101	-	P	37,4	17,9	11,1	2,6	22,2	55,5	7,2	A	A2
83	MONSOY-2002	-	P	39,3	16,7	12,1	2,3	23,4	54,4	6,4	A	A1
86	GOBR94-1704	Cristalina x FT-Estrela	M	41,1	16,6	12,4	2,0	25,3	52,1	6,8	A	A2
88	EMGOPA-309	Bulk Londrina retrocruzamento 4 vezes	P	37,0	16,4	12,3	2,3	28,7	47,8	6,6	B	B1
92	GOBR99-772034	-	T	36,8	16,4	12,4	2,4	19,3	57,4	7,2	A	A1
94	OCEPAR-19	-	P	37,5	17,0	12,5	2,7	23,4	51,6	6,8	A	A2
95	OCEPAR-207	-	P	36,8	17,9	12,4	1,8	29,0	48,7	6,4	B	B1
96	OCEPAR-10	-	P	36,7	17,4	12,0	1,9	21,6	55,4	7,2	A	A2
97	OCEPAR-17	-	P	38,1	17,0	11,2	2,2	29,0	49,9	5,6	A	A2
99	MATRINCHA	-	M	40,3	17,1	12,2	2,2	26,5	50,8	6,7	B	B1
100	BRGO99-2850	FT-18 x Dourados-B	T	38,6	16,8	11,8	1,9	29,0	49,3	6,3	B	B1
101	OCEPAR-3	-	P	39,2	16,8	11,9	2,0	23,8	53,7	6,6	A	A1
102	OCEPAR-8	-	P	39,1	15,8	12,2	2,4	20,5	55,4	7,0	A	A1
103	DOURADOS	-	P	38,7	17,0	11,9	1,9	26,8	52,3	5,9	A	A1
107	EMGOPA-307	GO79-3090 x Parana Goiana	M	39,5	16,5	12,8	2,4	21,9	54,6	6,6	A	A1
108	EMGOPA-303	IAC-73-2736 x IAC-6	T	39,5	16,0	12,1	2,2	20,4	57,4	6,7	A	A1
109	EMGOPA-304	Parana x Mandarin	P	39,1	15,9	13,6	2,3	21,9	53,5	6,6	A	A1
112	EMGOPA-312	Paraná Goiana x Cristalina	T	39,4	16,4	12,0	3,2	20,1	55,1	6,4	A	A3
116	MONSOY-8800	-	T	38,7	15,6	12,4	2,4	19,6	57,4	7,0	A	A1
117	IAC-14	-	T	38,8	16,1	12,6	2,1	23,3	54,7	6,2	A	A1
118	DM-NOBRE	Doko x BR-15 (Mato Grosso)	T	39,5	17,0	11,1	2,2	23,5	55,4	6,3	A	A1
119	DM-VITORIA	IAC-8 x MFV-9 (Sucupira)	T	37,1	16,5	10,6	1,8	29,7	51,4	5,8	A	A1
123	DM-339	Doko x BR83-6288 (Kge)	T	39,5	16,3	12,0	2,2	20,9	56,6	7,1	A	A1
124	EMGOPA-305 GOIATUBA	-	T	38,5	16,8	12,5	2,4	20,8	56,0	6,9	A	A1
127	FORREST	-	P	39,1	17,1	11,3	2,0	26,9	51,8	5,9	A	A1
129	TRACAJAS	-	T	38,6	16,9	11,6	2,3	27,5	52,1	5,7	A	A1
130	BRS JUSSARA	-	T	39,2	16,5	11,8	2,1	21,5	56,3	6,7	A	A1
131	GOBR96-013519	Sharkey x BR90-4630	P	41,0	15,6	11,7	2,1	31,4	45,8	5,9	B	B1
137	GOBR96-04303	E-313 (3) x BR90- 6979	T	36,8	16,7	12,0	2,2	20,5	58,0	6,6	A	A1
138	GOBR96-005547	BR-9 *2 x BR97-555	T	38,7	16,9	11,6	2,1	26,6	53,4	6,0	A	A1
139	GOBR97-056191 (IPAMERI)	LEFLORE *4 x BR90- 7057	T	40,1	16,3	11,6	1,9	30,8	48,5	6,0	B	B1
140	GOBR95-9108	BR 27 RC x (BR 27 RC x BR90-30604)	T	39,0	16,5	11,7	2,1	25,1	53,7	6,5	A	A1
141	GOBR96-005122	(EMGOPA-308 *2 x BR90-1983) x EMGOPA-313	T	38,6	16,7	12,3	2,3	23,8	53,0	6,6	A	A2
142	GOBR96-057216	Sharkey x BR92- 31861	T	39,4	16,8	11,8	1,9	22,9	55,7	6,4	A	A1

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
143	GOBR96-000913	BR 97 RC x (IAC-12 x FT-Cristalina)	T	40,4	16,7	11,5	2,3	21,7	56,9	6,6	A	A1
144	HARTWIG	-	P	38,4	17,0	12,2	2,2	23,6	53,7	6,5	A	A1
145	CENTENNIAL	-	P	39,7	16,4	10,3	1,8	36,3	43,8	5,8	B	B2
146	PI-88788	-	P	42,6	15,8	12,1	1,8	26,6	50,3	7,4	B	B1
147	PI-90763	-	P	40,0	15,6	11,7	1,5	31,1	45,7	8,0	B	B3
148	PICKET	-	P	41,2	17,1	11,0	1,5	35,6	43,5	6,4	B	B2
149	PEKING	-	P	40,1	15,9	12,2	2,2	24,1	51,6	7,9	B	B1
150	LEE	-	P	39,8	17,3	10,8	1,6	31,9	48,0	6,2	B	B1
151	MT/BR-53 TUCANO	-	T	37,9	16,8	12,4	2,2	21,3	56,6	6,8	A	A1
152	ARARA AZUL	-	T	37,8	17,0	11,9	2,0	25,7	53,8	5,9	A	A1
153	FMT-TUCUNARE	-	M	38,7	17,0	12,1	2,2	23,4	53,2	6,8	A	A2
154	MONSOY-5942	-	P	38,4	16,8	12,2	2,2	24,6	51,7	6,7	A	A2
156	BRGO00-14467-01-01	IAC-100 x (FT-5 x OC-9)	T	40,7	15,6	13,0	2,3	19,4	56,6	7,2	A	A1
157	DWIGHT	-	P	37,9	16,9	13,5	3,0	24,1	50,8	6,5	A	A7
158	INA	-	P	38,5	16,5	10,8	1,8	34,0	44,6	6,2	B	B2
159	CACHARA	-	M	40,5	16,4	11,9	2,0	29,6	49,2	6,0	A	A2
160	BRSMG-RENASCENÇA	-	P	39,4	16,3	12,2	2,3	28,1	49,5	6,0	A	A2
161	JACK	-	T	39,3	16,7	11,6	3,3	28,5	48,2	5,2	A	A2
164	GOBR99-4095	Dekalb-480 x FT-2000 e Don Mario 43 x FT-2000 e Pioneer 9942 x FT-2000	P	38,0	16,9	11,0	2,3	23,7	54,2	6,5	A	A2
166	GOBR99-4102	Don Mario 48 x Suprema	P	38,8	16,6	10,9	2,1	36,3	43,0	5,7	B	B2
168	BR-16 FRNG	-	P	38,1	16,3	12,0	2,5	20,3	55,9	6,9	A	A1
169	BR-36	-	P	38,9	17,0	12,1	2,1	23,0	55,1	6,0	A	A1
171	EMBRAPA-48	(Davis x Parana) x IAS-4 x BR-5	P	37,1	16,8	12,0	2,3	23,3	53,9	6,5	A	A1
173	EMBRAPA-59	-	P	37,7	16,7	12,8	2,5	23,5	52,2	6,2	A	A1
174	EMBRAPA-60	FT-Ayara x BR83-147	M	39,0	16,1	12,5	2,3	23,7	52,8	6,7	A	A2
187	EMGOPA-308.00.07	EMGOPA-308 (6) x BR 92-31914	T	38,4	17,2	11,7	2,2	22,6	56,8	6,3	A	A1
189	GOBR99-103043	CRIXAS x (EMGOPA-308(5) x BR92-31917)	T	38,4	16,3	11,6	2,2	23,4	55,3	6,5	A	A1
190	GOBR99-72070	FT-SERIEMA(2) x FT-ESTRELA	T	39,4	16,0	12,4	2,6	21,3	56,4	6,6	A	A1
191	GOBR99-419016	FT-45263 x (EMGOPA-306(5) x BR92-31910)	T	38,0	16,1	11,5	2,4	20,4	58,2	6,5	A	A1
193	GOBR99-366033	BR88-4903 x BR83-8072	T	38,9	16,6	12,2	2,1	21,6	55,9	6,8	A	A1
194	GOBR99-772025	BR82-1166 x E95-243	T	37,9	16,6	12,9	1,9	20,6	56,5	7,3	A	A1
195	GOBR99-884070	FT-106 x MTBR92-5291	T	38,9	16,8	12,1	2,1	23,2	55,8	6,1	A	A1
196	GO 00-366-01-01	GOBR95-69443 x EMGOPA-302	M	40,0	17,1	10,7	2,2	32,6	47,4	5,5	B	B1
202	IAC-100	-	P	41,3	15,4	13,0	2,5	19,8	54,2	7,7	A	A2

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
204	GOBR97-061004 (Chapadões)	Hartwig x (BR90-7063 x BR90-7213)	T	38,5	16,3	11,9	2,1	21,9	55,8	7,1	A	A1
207	GOBR98-008062	BR90-4784M x BR92- 15266	M	39,6	17,2	11,6	2,0	24,7	54,0	7,0	A	A2
210	GOBR98-17191	BRAGG x BR94- 23315	P	38,7	16,8	11,5	2,2	26,5	52,3	5,9	A	A1
215	GOBR98-078009	CENTENNIAL x BR94-23373	M	40,1	16,6	11,1	2,1	24,7	53,9	6,7	A	A1
217	GOBR98-10489	BRAGG x BR93- 32021	M	38,6	17,2	11,9	2,1	25,7	52,8	6,1	A	A1
221	GOBR98-057078 (EDEIA)	THOMAS x BR95- 22829	T	39,2	17,0	11,9	2,2	23,4	54,6	6,4	A	A1
222	GOBR98-015075	FORREST x BR95- 22829	M	38,5	16,6	11,6	2,3	28,1	51,2	6,0	A	A2
223	GOBR98-127058	CENTENNIAL x BR94-23386	T	39,8	16,8	10,4	1,6	31,8	48,8	6,1	B	B1
225	GOBR98-034035	MTBR92-25542 x E95-445	T	40,4	16,8	11,0	2,0	26,0	53,4	6,2	A	A1
226	GOBR98-053084	EMGOPA-308 x BR96-40916	T	39,1	16,4	12,0	2,4	21,5	56,2	6,5	A	A1
227	GOBR98-070075	CENTENNIAL x BR95-22780	T	39,6	16,3	12,0	2,2	20,4	57,4	6,7	A	A1
231	GOBR98-084078	CENTENNIAL*5 x BR94-23386	T	38,8	16,9	11,1	2,0	24,2	54,8	6,4	A	A1
239	GOBR98-072041	HARTWIG x BR90- 7063	T	39,4	16,4	11,9	2,2	22,6	55,0	7,0	A	A1
240	GOBR98-010092	MS/BR-39 x BR94- 23360	T	39,2	17,1	11,4	1,9	29,2	50,3	6,1	A	A2
245	Caiapônia	-	P	37,6	16,1	12,6	2,0	23,1	53,6	7,2	A	A2
257	GO 99-01.023	EMGOPA-302 x CRIXAS	P	38,6	16,5	12,2	2,6	22,5	53,6	6,9	A	A2
260	GO 99-03.009	CRIXAS x EMGOPA- 315	P	39,2	15,5	11,9	2,3	20,2	56,5	6,9	A	A1
265	GO 99-05.029	CRIXAS x GOBR95- 6224	P	37,8	16,2	12,1	2,4	23,8	52,7	6,7	A	A2
268	GO 99-06.013	GOBR95-1337 x EMGOPA-302	P	36,1	16,7	10,8	2,2	23,7	55,0	6,8	A	A2
271	GO 99-07.006	EMGOPA-316 x GOBR95-40469	P	38,3	16,9	13,0	2,5	24,3	51,0	6,3	A	A2
274	GO 99-08.069	GOBR95-2019 x EMGOPA-302	P	39,1	16,1	12,0	2,4	23,2	53,4	6,6	A	A1
278	BRGO99-4095-05	DEKALB-480 x FT- 2000 e DON MARIO 43 x FT-2000 e PIONEER 9442 x FT- 2000	P	38,0	17,5	16,6	2,9	32,4	35,5	3,6	E	E
285	BRGO99-4105-14	MYCOSOY-45 x SUPREMA	P	36,4	17,6	11,2	2,2	23,6	55,0	6,3	A	A1
289	BRGO99-4109-17	TIJERETA-42 x SUPREMA	P	38,2	17,4	10,8	2,2	27,6	51,7	6,0	A	A2
293	BRGO99-4186-23	MTBR95-123246 x MTBR95-123650	M	38,4	16,7	12,6	2,2	21,1	55,6	7,1	A	A1
296	BRGO99-4104-29	EUREKA-51 x SUPREMA e HARDEE x SUPREMA e JOKETA-46 x SUPREMA	P	37,7	17,4	11,1	1,8	29,3	50,3	5,6	A	A1
297	BR00-10072	BR97-21797 x UIRAPURU	P	38,7	16,9	11,8	1,9	22,0	54,9	8,1	B	B1
298	BR00-9163	MG/BR-42(8) x BR86-4009	P	38,3	17,4	10,7	2,1	25,7	53,8	5,7	A	A1
299	BR00-13246	PEQUI(2) x (TROPICAL(2) x EMBRAPA-20)	P	41,3	16,1	12,8	2,5	23,0	52,8	6,8	A	A2
304	BR00-11971	LN89-5699 x EMGOPA-308 RCH	P	39,3	17,2	12,6	2,1	22,9	54,4	6,5	A	A1

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
306	GOBR99-4102.30	Don Mario 48 x Suprema	M	37,4	17,0	11,0	2,3	26,7	52,8	6,2	A	A2
307	BR 00-12163	Pequi x BR94-23325	P	40,5	15,9	12,9	2,7	19,2	55,6	7,5	A	A1
308	PI-219653	-	M	44,0	14,8	12,1	2,1	27,3	48,8	7,4	B	B1
310	PI-171451	-	P	40,7	15,6	13,2	1,9	28,3	48,4	6,1	A	A5
312	PI-200490	-	P	40,1	16,9	9,6	1,8	39,6	40,4	6,2	C	C1
313	PI-240667	-	T	42,5	16,0	11,7	1,9	31,4	47,8	6,2	B	B1
314	PI-94159 FRF	-	M	40,1	16,1	10,1	1,6	34,7	45,6	6,8	B	B2
318	Santa Rosa	-	P	41,4	17,3	11,6	2,2	25,8	53,1	5,4	A	A1
322	IAC-8	-	M	39,9	17,1	11,7	2,2	26,6	51,1	6,8	B	B1
324	IAC-9	-	T	37,7	16,1	12,4	2,1	20,9	56,7	6,5	A	A1
328	IAC-13	-	P	38,3	16,5	11,9	2,3	21,8	55,3	6,5	A	A1
330	IAC-15-2	-	M	39,4	16,5	11,5	2,1	29,2	50,3	5,8	A	A2
331	IAC-16	-	P	38,3	17,5	12,1	1,8	19,6	58,5	7,0	A	A1
332	IAC-17	-	P	41,0	16,7	13,1	2,4	23,7	51,6	6,7	A	A2
333	IAC-18	-	P	40,1	16,2	12,3	2,5	22,3	53,7	7,0	A	A2
334	IAC-19	-	P	40,6	15,8	11,5	2,6	21,3	54,7	7,4	A	A2
336	IAC-2	-	M	39,6	16,4	11,2	2,0	24,2	55,7	6,0	A	A1
337	IAC-4	-	M	40,4	16,5	11,6	2,0	21,6	57,4	6,3	A	A6
339	IAC-22	-	P	39,4	16,2	13,2	2,2	18,5	56,4	7,4	A	A1
340	IAC-23	-	M	40,0	16,5	11,7	2,2	25,5	52,7	6,6	A	A2
345	FT-2	-	M	39,4	16,8	11,7	1,9	26,7	51,2	6,6	A	A2
348	PARANAGOIANI A	-	T	39,2	16,4	11,9	2,2	21,1	57,2	7,1	A	A1
349	MONSOY 9350	-	T	38,1	16,9	12,3	2,3	20,8	57,5	6,5	A	A1
350	MONSOY 8211	-	M	40,3	16,8	11,9	2,1	27,1	51,6	6,3	A	A2
351	MONSOY 9010	-	T	38,9	16,1	12,4	2,3	18,5	57,5	7,6	A	A1
352	MONSOY 8914	-	T	37,8	17,1	12,0	2,3	23,4	55,0	6,5	A	A1
353	MONSOY 8866	-	T	38,4	16,9	10,8	2,1	27,4	53,0	6,0	A	A1
355	MONSOY 8998	-	T	39,2	16,5	11,8	2,2	21,7	57,3	6,6	A	A1
356	MONSOY 8222	-	M	39,5	16,5	11,9	2,2	23,8	53,5	7,0	A	A2
357	M 98594	-	P	38,9	16,0	12,1	1,9	21,2	55,1	7,6	A	A2
358	MONSOY 7900	-	M	37,6	17,3	11,2	2,5	22,7	54,8	7,2	A	A2
359	TAINUNG 4 (PI 368039)	-	P	41,2	16,6	11,3	1,8	36,4	42,4	6,1	B	B2
360	ANDREWS	-	M	40,0	16,3	12,1	2,1	21,9	56,5	6,4	A	A1
361	BOSSIER	-	P	38,9	16,4	11,8	2,2	28,4	49,0	6,3	B	B1
362	IAS 3 (DELTA)	-	P	40,7	15,8	11,9	2,5	24,2	53,3	6,0	A	A1
367	CEP 12 (CAMBARÁ)	-	P	37,6	17,0	11,5	2,2	22,4	56,3	6,0	A	A1
370	BRSMÁ PARNAÍBA	-	T	39,6	16,6	11,9	2,2	23,5	55,0	6,3	A	A1
390	GOBR99-10134	BR86-11864(6) x BR37-555	P	37,2	17,9	12,3	2,2	28,0	49,8	5,8	A	A2

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
392	BRS JIRIPOCA	-	M	39,2	17,0	11,7	2,2	21,4	56,2	7,1	A	A1
394	GOBR97-14346 (ARAÇU)	SHARKEY x [FORREST*3 x (LANCER x BR80- 6989)]	P	38,0	16,2	10,9	2,0	31,8	46,5	6,0	B	B1
397	GOBR97-14723	MGBR-22*7 x DOURADOS (TOLMIN)	P	39,7	17,2	10,4	1,7	39,7	40,4	5,3	B	B2
398	GOBR98-063077	THOMAS x BR95- 22767	M	38,8	16,2	12,1	2,4	24,6	52,5	6,4	A	A2
400	GOBR98-008097	BR90-4784M x BR92- 15266	M	38,6	16,3	12,0	1,8	22,0	55,5	7,4	A	A2
403	GOBR98-065071	HARTWIG x MTBR94-18190	T	39,9	16,5	11,9	2,1	21,6	56,0	7,1	A	A1
404	GOBR97-08728	BR93-32121 x LAMAR	T	40,3	16,8	10,6	1,9	30,9	49,0	5,8	B	B1
405	GOBR98-085064	E93-402*3 x E94-64	T	40,0	16,7	10,9	1,8	36,5	43,9	5,5	B	B1
409	GOBR99-523040	FT-2000 x (IAC-31 x FOSCARIU)	M	38,7	17,3	13,1	2,5	23,1	54,1	6,3	A	A1
410	GOBR99-143010	SHARKEY x MTBR95-12243	M	42,1	16,7	11,7	2,1	31,5	46,5	6,3	B	B1
413	GOBR99-971004	BR94-23373 x FORREST	M	40,2	16,7	10,5	1,9	39,8	40,7	5,7	B	B2
415	GOBR99-12548	IAC-8(8) x CRISTALINA	M	40,7	16,9	11,8	1,9	27,3	50,7	6,6	A	A2
416	GOBR95-1337- 00-08	FT-COMETA x FT-5 RNCS	M	37,5	17,0	11,4	2,2	27,3	51,7	5,9	A	A1
422	GOBR95-322-00- 01	BR85-6356 x OCEPAR-3	M	38,1	16,5	12,5	2,3	24,8	52,0	6,6	A	A2
423	GOBR99-961001	FORREST x BR95- 22882	M	39,0	16,2	11,8	2,0	32,9	45,5	6,0	B	B1
426	GOBR99-945001	SHARKEY(4) x PADRE	T	39,5	17,2	10,8	1,8	32,3	47,6	6,2	B	B1
427	GOBR99-671019	FT-5 SDSR x (FT- JATOBÁ x E 94-189c)	T	39,6	17,0	11,4	1,8	30,5	49,8	5,5	A	A2
429	GOBR99-316019	(EMGOPA-306(5) x BR92-31910) x FT- JATOBÁ SDSR	T	38,8	16,9	12,2	2,0	25,0	53,8	6,3	A	A1
430	GOBR99-148011	LEFLORE x MT95- 121461	M	38,3	16,6	12,1	2,1	24,2	53,9	6,4	A	A1
433	GOBR99-627004	CARVER x MTBR95- 121430	T	40,7	16,4	11,7	2,0	21,9	55,8	7,0	A	A1
439	GOBR99-671061	FT-5 SDSR x (FT- JATOBÁ x E 94-189c)	T	40,1	17,0	11,2	1,8	33,9	47,0	5,4	B	B1
442	GOBR99-577035	FT-109 x CONQUISTA	T	39,7	16,4	11,9	1,7	25,3	53,4	6,5	A	A1
449	GOBR99-610052	HARTWIG x MT95- 122360	T	39,2	17,3	13,9	3,5	28,6	45,3	5,2	A	A7
450	GOBR99-123039	(CURIÓ x EMGOPA- 313 (4) x BR92-31910	T	38,4	16,3	12,1	2,2	20,9	57,3	6,8	A	A1
464	GOBR99-888020	FT-108 x MTBR91- 50789	M	38,5	15,6	12,0	2,3	22,1	55,7	6,7	A	A1
470	GOBR99-773014- 03-01	BR82-1166 x E 95-466	M	37,9	16,5	12,6	2,3	20,8	56,9	6,5	A	A1
489	PI-224270 FRF	(HONGYO R4)	P	40,8	16,9	11,5	1,5	24,1	55,8	6,5	A	A1
491	PI-262180	SAN RUO	P	41,5	17,4	11,0	2,2	31,8	48,1	5,2	A	A2
494	PI-200455	ASO Nº 1	P	41,2	16,9	11,5	2,1	25,9	52,6	6,1	A	A1
495	PI-200492	KOMATA Rpp 1	P	39,8	16,6	11,5	2,0	31,4	47,2	5,9	B	B1
496	PI-200474	ARAGE DAIZU	P	39,2	16,6	11,7	2,6	26,4	51,8	5,6	A	A1
497	PI-200466	GARA BUN	P	40,3	16,8	12,3	1,8	28,2	50,6	5,7	A	A1
500	BRS APIAKÁS	-	M	38,7	16,0	11,9	2,1	23,8	55,1	6,6	A	A1
501	IAC 23	-	P	41,5	16,5	12,1	2,5	22,4	54,0	6,8	A	A1

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
502	BRMS SAVÁ	-	P	40,2	15,9	12,4	2,1	20,1	55,4	8,0	A	A2
504	BRMS ACARÁ	-	P	37,9	17,1	11,1	2,2	29,4	49,8	5,7	A	A2
506	BRS ROSA	-	P	40,5	16,6	11,1	1,7	36,6	42,4	5,9	B	B2
507	BRS BARREIRAS	-	T	38,5	16,1	11,6	2,1	21,0	56,6	7,2	A	A1
511	UFV 18 (PATOS DE MINAS)	-	T	37,7	16,8	11,3	2,0	20,6	57,9	7,3	A	A4
512	BRS 206	-	P	38,0	16,4	12,6	1,9	23,6	53,6	6,6	A	A1
513	FT-LIDER	-	P	38,8	17,1	11,7	2,1	30,5	47,5	6,1	B	B1
514	BRS 216	-	P	41,0	15,7	13,2	2,7	17,6	55,1	8,1	A	A2
515	BRS ANHUMAS	-	T	38,8	17,1	11,9	2,0	33,3	45,5	5,5	B	B1
516	PI-371609	-	P	40,0	16,3	11,2	2,1	23,0	55,5	6,2	A	A1
517	PI-90406	-	P	39,0	17,0	9,6	1,6	45,4	36,0	5,3	C	C1
518	BRS 219 (BOA VISTA)	-	P	40,9	15,5	11,8	2,4	21,5	55,8	6,7	A	A1
519	BRMS TAQUARI	-	T	38,9	17,3	11,3	1,8	30,8	49,2	5,9	A	A2
520	HOOD 75	-	P	39,4	16,5	12,8	2,5	19,5	55,7	7,0	A	A1
522	RB 501	-	P	39,0	16,5	12,4	2,2	25,2	51,2	6,6	A	A2
523	FT-107	-	P	37,2	16,4	14,0	2,3	23,5	52,1	6,0	A	A5
524	RB 604	-	P	37,9	18,5	11,9	1,6	27,9	51,0	5,8	A	A1
525	CD 207	-	M	39,3	17,5	10,7	2,0	35,5	44,9	6,0	B	B2
526	BRI-98-641	-	P	39,7	16,1	13,5	2,4	21,4	53,8	6,8	A	A1
527	BRS 217 (FLORA)	-	M	40,7	17,3	11,6	2,1	25,3	53,2	6,3	A	A1
528	CD 205	-	T	39,4	16,7	11,7	1,9	24,0	55,0	6,7	A	A1
529	CD 206	-	P	38,6	16,8	10,3	2,0	32,1	48,0	5,5	B	B1
530	BRS 181	-	P	38,0	16,3	13,7	2,5	21,2	53,6	6,7	A	A1
531	UFVS 2002	-	T	38,0	16,2	12,6	2,1	18,8	58,4	7,5	A	A1
532	S 0203	-	P	40,2	16,4	12,3	1,4	23,1	54,4	7,7	A	A2
533	RJ 9819485	-	P	38,3	17,0	13,2	2,2	20,3	55,5	7,0	A	A1
534	BRQ 94-7277	-	P	41,1	16,2	12,6	2,5	30,5	46,5	6,1	B	B1
535	BRQ 95-1873	-	T	39,0	16,2	12,4	1,9	34,1	44,1	6,4	B	B4
536	BRQ 96-1257	-	T	39,4	16,5	12,1	1,9	23,1	54,2	7,1	A	A2
537	BRQ 96-3065	-	T	41,0	15,7	12,9	2,3	22,2	53,8	7,3	A	A2
538	PI-595099	-	P	39,8	16,3	12,4	2,1	23,4	51,8	7,8	B	B1
539	PI-230977	-	P	38,3	17,3	11,8	1,5	27,2	50,9	6,3	A	A2
543	BRQ 95-1159	-	P	39,7	16,1	11,8	1,4	27,6	49,3	7,4	B	B1
545	BRQ 94-7898	-	P	39,0	17,0	12,3	1,7	29,0	48,5	6,3	A	A2
546	S 0205	-	M	38,9	17,4	12,1	1,8	25,2	53,6	6,2	A	A1
547	BED FORD	-	P	38,0	16,6	10,6	2,1	30,8	48,6	5,8	B	B1
548	S 0206	-	P	39,1	16,0	12,1	1,6	18,9	56,0	6,8	A	A1
549	BRQ 94-2914	-	P	39,2	16,5	12,5	2,1	25,1	51,2	6,8	A	A2
550	BRQ 95-799	-	P	38,9	16,0	14,3	2,2	23,9	50,1	6,6	A	A5
552	MONSOY 9001	-	T	37,7	17,1	13,0	2,4	22,1	54,9	6,6	A	A1

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
553	SHARKEY	-	P	41,1	16,6	10,3	1,4	48,9	31,3	5,2	C	C2
554	RB 605	-	P	38,9	16,3	12,4	1,8	21,9	55,3	7,4	A	A2
558	UFVS 2001	-	P	39,2	16,3	11,2	1,7	22,3	55,3	7,1	A	A1
559	MONSOY 7901	-	P	37,8	17,4	11,9	2,2	22,1	55,8	6,4	A	A1
561	MONSOY 7501	-	M	38,2	16,4	12,5	2,1	21,8	55,1	7,2	A	A2
562	BRQ 95-2313	-	M	37,3	17,0	11,6	2,1	22,2	55,9	7,0	A	A1
568	BRS PÉTALA	-	T	38,9	17,0	10,9	2,1	23,9	55,0	6,9	A	A2
569	MONSOY 8550	-	M	38,4	15,9	12,1	2,2	20,5	57,1	7,0	A	A1
570	FT 100	-	M	39,3	16,6	11,0	2,0	32,0	48,3	6,2	B	B1
572	CEP 26 (UMBU)	-	P	38,1	17,4	11,9	1,9	27,1	51,0	6,3	A	A2
573	OCEPAR 18	-	P	38,8	17,1	12,0	1,8	25,4	52,3	6,2	A	A1
574	MONSOY 7201	-	P	38,9	16,7	11,9	2,3	21,1	56,8	6,3	A	A1
576	BRS 214	-	M	39,4	16,9	12,5	1,9	25,2	52,7	6,3	A	A1
577	DM RAINHA	-	M	39,1	15,9	11,3	1,8	33,3	47,2	5,8	B	B1
578	BRS 185	-	P	39,3	16,5	12,3	2,3	21,3	55,3	6,7	A	A1
579	BRSMS PIAPARA	-	P	38,5	17,7	12,1	1,7	21,7	55,9	6,8	A	A1
581	EMBRAPA 47	-	P	37,9	16,5	12,6	2,2	21,5	55,4	6,4	A	A1
582	FT 104	-	M	38,0	16,0	12,6	2,2	21,7	52,6	7,0	A	A2
583	FT MARACAJÚ	-	M	38,9	16,8	12,5	2,2	22,0	54,8	6,9	A	A1
585	SANT' ANA	-	M	38,5	16,5	12,1	2,2	23,7	54,1	7,0	A	A2
586	UFV 3	-	M	40,4	17,0	10,0	2,1	19,2	62,1	6,2	A	A8
587	FT EUREKA	-	P	38,9	16,6	12,2	2,2	31,8	45,9	5,7	B	B1
588	FT MANACÁ	-	P	38,2	16,3	16,6	2,3	19,2	52,1	7,0	D	D
589	MT/BR 50 (PARECIS)	-	M	40,3	16,6	11,6	2,0	21,8	55,8	7,3	A	A1
590	BRS 213	-	P	39,4	16,1	12,4	1,9	29,8	46,1	6,7	B	B1
592	BRS 212	-	P	40,5	16,5	14,0	2,2	24,7	50,5	6,9	A	A5
593	CD 202	-	P	37,0	16,9	11,4	2,0	34,0	44,5	5,7	B	B1
594	FEPAGRO-RS 10	-	P	41,3	15,7	12,6	2,3	19,0	56,1	7,2	A	A1
595	BRSMG GARANTIA	-	T	39,7	17,3	11,6	2,0	24,9	53,9	6,5	A	A1
598	CEP 10	-	P	38,5	17,0	11,8	2,1	24,0	54,3	6,3	A	A1
599	FT SARAY	-	P	38,0	17,5	13,0	2,2	23,4	52,4	7,0	A	A2
600	FT 102	-	T	38,0	17,2	12,5	2,3	25,4	51,7	6,3	A	A2
605	PLANALTO	-	P	39,1	17,2	12,9	2,5	26,4	50,1	6,0	A	A2
607	UFV 9 (SUCUPIRA)	-	P	38,8	16,7	12,4	3,4	21,8	52,4	5,2	A	A3
608	TIARAJU	-	P	39,3	16,7	12,2	1,6	20,3	57,2	7,2	A	A1
610	FT 20 (JAÚ)	-	P	40,5	16,8	12,4	2,1	27,6	50,7	5,9	A	A1
612	INVICTA	-	P	38,8	17,2	12,9	2,3	23,8	52,7	6,2	A	A1
614	IPAGRO 20	-	M	39,5	16,2	12,4	2,5	19,5	57,7	7,2	A	A1
615	CAC 1	-	M	40,0	16,4	11,7	2,4	21,4	55,6	7,1	A	A1
616	MISSÕES	-	M	39,4	17,0	11,6	2,1	24,3	55,0	6,2	A	A1

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
617	UFV 4	-	M	39,2	17,2	11,9	2,1	23,5	54,9	6,4	A	A1
619	EMBRAPA 5	-	P	37,6	17,4	12,4	2,4	21,4	55,0	6,4	A	A1
621	SPS 1	-	P	37,8	16,9	11,7	1,6	26,6	50,9	6,8	A	A2
623	COKER 136	-	P	38,9	16,4	12,6	2,6	22,6	53,1	6,6	A	A2
624	DÉCADA	-	P	38,6	16,9	11,1	2,0	25,8	51,8	7,2	B	B1
625	BR 4 RC	-	M	39,2	16,7	11,7	1,8	30,2	48,9	6,1	B	B1
626	PAMPEIRA	-	P	38,3	17,0	11,6	1,6	23,1	54,9	7,5	A	A2
627	UNIÃO	-	P	39,5	16,3	12,6	2,1	25,3	50,7	6,9	A	A2
628	UFV ARAGUAIA	-	T	39,8	16,8	11,7	2,3	20,6	57,3	6,8	A	A1
630	FT 14 (IRACEMA)	-	T	37,1	16,9	12,1	2,2	20,5	57,6	6,7	A	A1
632	SP/BR 41	-	P	38,7	16,9	11,6	2,1	28,5	49,7	6,2	A	A2
636	EMBRAPA 64 (PONTA PORÃ)	-	P	38,2	16,8	12,6	2,2	21,6	53,9	7,0	A	A2
637	CAJEME	-	P	39,5	16,8	11,6	1,8	31,5	47,4	5,9	B	B1
639	HARDEE	-	M	38,7	16,6	12,0	1,9	21,6	56,6	6,7	A	A1
645	MS/BR 17 (SÃO GABRIEL)	-	M	35,7	17,0	10,8	2,0	24,5	55,7	6,0	A	A4
647	BR 24	-	P	37,5	17,2	11,8	2,0	24,1	53,1	6,3	A	A1
648	MS/BR 20 (IPE)	-	P	39,7	17,1	11,4	2,0	22,8	55,9	6,5	A	A1
650	UFV 14	-	T	39,7	16,9	10,7	1,9	32,6	48,1	5,9	B	B1
651	UFV 2	-	M	39,3	17,1	12,2	2,0	20,2	58,6	6,9	A	A1
652	HILL	-	P	39,4	16,6	11,9	1,7	29,7	47,2	6,2	B	B1
663	FT 11 (ALVORADA)	-	T	39,0	16,6	11,9	2,2	22,4	55,6	6,6	A	A1
664	FT 8	-	M	37,9	16,3	11,7	2,1	24,1	54,6	6,3	A	A1
672	CORDELL	-	P	39,4	16,5	11,1	2,0	35,0	44,4	5,4	B	B1
677	BRQ 95-1136	-	M	38,7	16,2	12,3	2,1	34,4	42,5	6,1	B	B4
679	BRQ 94-1311	-	M	39,4	16,8	12,4	2,0	28,3	49,2	6,8	B	B1
685	MONSOY 7603	-	M	38,9	16,4	11,7	1,9	30,4	49,2	5,8	A	A2
691	BRS BABAÇU	-	T	38,5	16,5	11,8	2,5	20,5	56,6	7,2	A	A1
693	MGBR-48 (GARIMPO)	-	P	38,5	16,8	12,6	2,0	26,8	50,4	6,2	A	A2
697	EMBRAPA 25	-	P	38,9	17,2	10,7	2,1	25,7	53,8	5,7	A	A1
709	IVORÁ	-	P	37,7	16,9	11,7	2,0	20,8	57,0	6,6	A	A1
715	UFV 15 (UBERLANDIA)	-	M	40,0	16,6	12,7	2,3	24,2	53,0	6,3	A	A1
716	J 200	-	M	39,9	16,8	11,6	1,9	33,2	46,3	5,8	B	B1
719	SERTANEJA	-	P	38,6	16,5	12,6	2,2	22,4	53,9	6,7	A	A1
720	FT 101	-	T	40,0	16,9	11,6	1,9	27,6	51,7	6,1	A	A2
722	PIRAPO 78	-	P	38,1	17,4	11,9	2,4	20,6	55,5	7,0	A	A1
726	IPAGRO 21	-	P	39,2	16,3	11,8	2,5	29,5	47,0	6,0	B	B1
728	RS 7 (JACÚ)	-	P	39,4	17,0	12,7	2,4	26,4	49,7	6,6	B	B1
735	UFV 8 (MONTE RICO)	-	T	40,3	16,6	12,0	2,3	20,1	57,6	6,8	A	A1

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
753	INDUSTRIAL	-	M	40,0	17,0	12,0	2,4	18,1	59,3	6,1	A	A6
754	BRS 138	-	M	39,0	16,9	11,8	2,4	26,0	52,8	6,2	A	A1
760	MSBR 34 (EMPAER 10)	-	M	38,4	16,6	11,9	2,2	23,5	54,5	6,9	A	A2
762	EMGOPA 313	-	T	37,0	17,1	12,6	2,3	21,4	56,5	6,1	A	A1
763	BRQ 95-2412	-	M	40,2	17,0	12,2	2,3	20,5	56,6	7,2	A	A1
764	BRS 211	-	M	38,9	17,0	12,4	2,2	23,8	54,0	6,5	A	A1
769	CD 209	-	M	39,5	15,8	12,9	2,3	21,5	54,8	6,9	A	A1
771	EMGOPA 315 (R. VERMELHO)	-	M	40,1	17,2	13,2	2,3	22,2	55,0	6,7	A	A1
772	PRATA	-	M	40,5	16,5	11,4	2,0	32,5	46,7	5,8	B	B1
773	EMBRAPA 34 (TEREZINA)	-	M	40,5	16,8	11,5	2,0	33,4	45,9	5,5	B	B1
775	DM SOBERANA	-	M	40,6	16,1	11,4	2,0	27,8	51,6	6,2	A	A2
776	FT MORENA	-	M	36,6	16,1	12,9	1,9	21,0	55,6	7,3	A	A1
790	MONSOY 7202	-	M	39,3	16,2	12,3	1,9	29,0	50,1	6,4	A	A2
791	PI-200490	-	T	41,0	17,0	11,7	1,9	32,6	46,4	5,9	B	B1
815	BRSMS MANDI	-	M	37,3	16,9	12,2	2,0	24,7	54,2	6,4	A	A1
826	BRS 153	-	P	38,9	17,3	11,7	2,1	25,9	53,1	5,9	A	A1
827	BRS 154	-	T	38,6	16,1	12,1	2,0	26,4	51,3	6,9	A	A2
829	BRSMG VIRTUOSA	-	P	37,4	16,8	11,8	2,4	24,5	53,9	5,7	A	A1
832	GOBR00-0466- 013	GOBR97-055009 NC x GOBR95-1337	M	39,1	16,9	10,3	1,9	36,4	45,1	5,5	B	B1
833	GOBR00-0464- 048	PINTADO x GOBR93-1483	P	39,1	17,4	10,5	1,4	34,4	46,0	6,0	B	B1
834	GOBR00-09-072	MGBR95-19125 x EMGOPA-316	M	38,5	17,3	11,3	2,2	25,3	53,9	6,3	A	A1
835	TAGHV- STWART-1-03-01	-	M	38,5	17,5	11,3	2,2	27,0	51,8	5,8	A	A1
837	GOBR00-028-057	EMGOPA-316 x MGBR9-964	T	39,7	17,1	10,9	2,5	25,5	53,4	6,4	A	A2
839	GOBR00-0396- 060	GOBR97-057130 NC x EMGOPA-306 RCH	P	38,7	16,7	12,8	2,2	24,0	52,8	6,4	A	A1
843	GOBR00-288-041	BR95-1709 x CSM-1	T	37,7	17,2	12,1	2,1	26,9	51,6	6,4	A	A2
846	GOBR00-28-035	EMGOPA-316 x MGBR95-964 BKCX-458 x	P	39,3	16,0	11,8	2,7	20,0	55,9	7,2	A	A1
847	PI-30-18	SUPREMA e DEKALB-480 x SUPREMA	P	37,4	16,9	10,9	2,5	26,3	52,2	6,1	A	A2
852	GOBR00-181-015	EMGOPA-306 RCH x GOBR98-1483	M	39,2	15,7	12,6	1,9	22,7	54,9	6,6	A	A1
855	BRGO02-12206	PINTADO x MONSOY 8400	T	39,0	16,7	11,8	2,0	24,0	55,0	6,6	A	A1
857	BRGO02-11627	PARANÁ SDSR x PINTADO	M	39,0	17,3	11,7	2,1	23,9	54,5	6,7	A	A1
864	BRB02-685	BRS-136 x (BRS- 133*6 x PROG.96- 246)	P	41,7	16,6	11,8	2,3	24,1	53,0	6,2	A	A1
869	BRB01-207	BRMS SAUÁ x BRMS CANADÁ x EMBRAPA-62	P	38,4	16,0	12,0	2,4	25,2	51,1	6,3	A	A2
886	GOBR99-773014- 03-02	BR82-1166 x E95-466	T	38,4	16,1	11,7	2,2	20,8	57,4	6,6	A	A1
887	BRGO99-4313-03	BR82-1166*2 x E96- 1274	T	38,3	16,5	11,0	1,8	32,0	48,1	6,2	B	B1
888	BRGO99-4366-16	E96-78417 x BR92- 1166	T	37,5	16,7	12,8	2,2	21,1	56,8	6,6	A	A1

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
890	GOBR00-345-028	EMGOPA-315 x LUZIANIA	T	39,8	16,4	11,7	1,8	30,9	48,9	5,4	A	A1
893	GOBR00-0545-03-01	SANTA CRUZ x EMGOPA-313 RCH	T	37,6	17,1	12,3	2,0	23,5	54,9	6,3	A	A1
894	GOBR00-0544-03-01	LUZIANIA x JATAÍ	T	39,2	17,0	11,8	1,9	29,6	49,2	6,2	B	B1
895	BRGO01-02373-03-01	BRQ-95-2193 x VENCEDORA	M	40,1	16,3	12,9	2,6	21,3	54,6	7,0	A	A1
900	BRGO01-00017-03-01	BR83-9396 RCH x E96-782	M	38,5	17,3	12,1	2,1	23,4	55,2	6,5	A	A1
907	BRGO01-00848-03-01	UFV-5*3 x E96-1274	M	38,6	16,5	11,6	2,0	26,9	52,0	6,5	A	A2
908	BRGO01-03182-03-01	BR82-1166*3 x (BR83-954 x EMGOPA-308 RCH)	M	37,9	16,8	12,0	2,4	19,6	58,0	6,8	A	A1
909	BRGO01-01081-03-01	EMGOPA-308 RCH x E96-781	T	37,8	17,0	13,0	3,8	20,3	53,6	5,9	A	A3
910	BRGO01-00460-03-01	E96-778*3 x EMBRAPA-20	T	38,3	16,7	11,9	2,3	20,7	57,1	7,1	A	A1
915	BRGO01-01381-03-01	EMGOPA-308 RCH*2 x E96-527/34	T	38,2	16,6	11,8	2,0	22,8	55,3	6,6	A	A1
916	BRGO01-00777-03-01	UIRAPURU x E96-781	T	38,4	16,8	11,6	2,0	23,4	55,9	6,5	A	A1
917	BRGO01-004542-03-01	E96-778*3 x BR28*3 x EMBRAPA-20	T	39,2	16,9	10,9	2,1	28,5	52,0	5,7	A	A1
921	BRGO99-43-1425-03-01	BR82-1162 x (EMGOPA-313 RCH, EMGOPA-308 RCH, EMGOPA-308 RCH)	T	37,6	16,7	12,3	2,3	21,6	55,8	6,8	A	A1
922	BR00-2334-03-01	BR27 x BR93-32061	T	40,0	16,9	12,2	2,3	21,5	56,4	6,4	A	A1
923	BR00-14381-03-01	DOURADOS x OCEPAR-9 x FT-5	M	39,7	16,8	11,8	2,1	22,6	56,3	6,3	A	A1
924	BRGO99-431328-03-01	BR82-1166*2 x E96-1274	T	37,1	17,1	10,3	1,9	20,9	59,5	6,9	A	A4
925	BR00-8-03-01	EMGOPA-306*6 x BR92-31910	T	39,0	16,2	11,4	2,5	22,4	55,9	6,5	A	A1
927	BRGO99-409659-03-01	A-4004 x SUPREMA e A-4656 x SUPREMA e A-4702 x SUPREMA	T	37,7	16,6	10,7	2,3	32,5	47,9	5,6	B	B1
928	BR00-8113-03-01	FT-JATOBÁ x BR89-11899-1 D	T	37,8	17,3	11,6	2,0	27,5	52,1	5,5	A	A1
929	BRGO99-4101-51-03-01	DON MARIO 43 x SUPREMA	T	39,4	17,5	9,9	1,8	35,1	47,3	5,6	B	B1
931	BRGO99-4175-13-03-01	CARVER x [LETORE*5 x (BR83-9396 x BR86-4009)]	M	38,3	16,4	11,9	2,4	24,8	53,8	6,0	A	A1
932	BR00-12-482-03-01	SHARK*3 x [WARTW x [BR16 (FOREST*3 x PRGO)]]	M	39,7	16,9	11,9	1,9	28,1	50,8	6,0	A	A2
934	GOBR01-0978	OCEPAR-207 x EMGOPA-302	P	39,0	16,2	11,8	2,2	31,4	47,3	5,8	B	B1
936	GO 02-2118	((EMGOPA-316 x GOBR95-40469) x (GOBR95-2019 x EMGOPA-302))	M	40,8	16,6	11,5	2,0	33,2	46,1	5,9	B	B1
938	GOBR00-280-032	BR89-10744 x FT-109	P	38,3	16,7	12,1	2,5	26,0	51,2	6,1	A	A2
946	BRB02-789-03-01	EMBRAPA-48*7 x EMBRAPA-62	P	38,4	17,2	12,1	2,1	22,9	54,4	6,5	A	A1
950	BRB02-788-03-01	(G 93-9223 x BR93-11995) x (CARVER x BR93-11995)	P	38,8	16,5	11,6	2,2	27,7	50,9	6,4	A	A2
954	BRB02-742-03-01	Bedford x BR98-16128	P	38,9	16,5	11,6	1,9	24,7	54,1	6,5	A	A1
1076	GOBR02-2215-18	EMGOPA-315 x MT/BR-53 TUCANO	M	38,0	17,0	12,2	1,9	23,7	54,5	6,5	A	A1

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
1077	GOBR02-2255-25	BRGO99-4186-23 x GOBR93-9960	M	38,5	17,1	12,6	2,2	22,1	54,7	6,9	A	A1
1079	GOBR02-2200-10	ST93-2051 x GOBR98-002055	M	38,4	17,4	10,6	2,0	28,4	51,8	6,0	A	A2
1080	GOBR03-2578-10	BRS-134 x MONSOY-8914	M	39,2	15,6	12,4	2,2	18,6	57,8	7,6	A	A1
1081	GOBR03-2550-11	FT-COMETA x EMGOPA-308.00.07 RCH	M	39,3	16,9	11,7	2,2	20,7	57,4	6,9	A	A1
1150	GOBR03-3160-9	GO02-2217 x BRGO99-4109-17	P	40,4	16,9	11,7	2,1	28,7	49,7	6,0	A	A2
1151	GOBR03-2893-3	BRS-133 x BRS- APIAKAS	P	39,2	17,3	11,4	2,0	28,6	50,6	5,9	A	A2
1152	GOBR03-2866-30	BRS-133 x BRGO99- 4102-08	P	38,3	17,6	10,8	2,1	31,9	46,5	6,0	B	B1
1187	BRGO02-1798- 03-01	GOBR99-35610 x GOBR99-4102	M	38,5	16,5	12,2	2,2	21,2	56,1	6,8	A	A1
1189	BRGO01-00453- 03-01	E96-778*3 x (BR- 28*3 x EMBRAPA- 20)	T	37,6	16,7	12,2	2,3	20,6	57,3	6,3	A	A1
1190	BRGO01-01020- 03-02	EMGOPA-308 RCH x E96-781	T	37,7	16,8	12,4	2,1	23,3	54,8	6,4	A	A1
1234	BRB02-816-03-01	BR98-25488 (46,3% Prot) x BR-36	M	41,3	16,6	12,3	2,2	21,7	52,6	7,0	A	A2
1288	BRB02-895-03-01	BR95-29133-1 x G 93- 9223	M	39,3	16,8	10,6	2,0	27,9	51,8	6,4	A	A2
1313	BRB02-854-03-01	BR96-25619 x BR97- 11260 (Res NCS)	M	40,3	16,2	10,1	1,7	33,2	47,0	6,5	B	B1
1316	GOBR00-118-093	antiga 584 do bag	M	38,6	16,9	12,0	2,6	26,7	50,5	6,2	A	A2
1344	BRB02-974-03-01	BR96-25917 (Res NCS, MI e MJ) x BR93-11995	M	38,8	17,3	11,4	2,1	24,7	54,8	6,0	A	A1
1349	BRB02-986-03-01	BR98-25978 (45,7% Prot) x BR93-11995	M	41,0	15,9	11,8	2,2	24,6	53,8	6,5	A	A1
1351	BRB02-989-03-01	BR97-11984 x BEDFORD	M	40,6	16,6	11,3	2,0	28,4	50,6	6,2	A	A2
1362	BRB02-977-03-01	BR98-25710 (43,8% Prot) x BR-36	M	41,1	16,6	11,5	2,2	34,5	43,8	6,2	B	B2
1363	BRB02-1377-03- 01	PROG.99-1265 x MONSOY-8757	M	38,1	16,9	12,3	2,2	20,8	56,2	7,0	A	A1
1451	BRB02-1030-03- 01	BR96-25917 x GORDON	M	39,7	16,9	11,4	2,3	23,1	55,4	6,4	A	A1
1452	BRB02-1054-03- 01	SÃO GABRIEL*7 x EMBRAPA-64	M	39,0	17,2	12,0	1,8	28,2	50,7	6,2	A	A2
1462	BRB02-1003-03- 01	BR96-25619 x BR96- 25917 (Res NCS, MI e MJ)	M	39,8	16,8	11,4	2,1	24,1	54,6	6,5	A	A1
1472	BRB02-1000-03- 01	CD 205 x BR96-9990 (Res NCS)	M	38,2	16,2	12,1	2,3	24,3	53,3	6,4	A	A1
1475	BRB02-1108-03- 01	MTBR99-39982 (Res M. A) x PINTADO	M	40,5	14,9	13,2	2,7	20,2	54,7	7,1	A	A1
1476	BRB02-120/121- 03-01	BRSMG- LIDERANÇA*2 x EMBRAPA-62/ (EMBRAPA-20 x BRS89-87)	M	38,7	16,9	12,4	2,6	20,3	55,5	7,6	A	A2
1493	BRB02-1064-03- 01	BR98-25854 (42,4% Prot) x E00- 2781/84/87/90/93/96 (11,12,13)	T	39,4	16,7	11,1	2,1	20,2	57,5	7,4	A	A4
1495	BRB02-1061-03- 01	CONQUISTA x MTBR99-39982 (Res M. A) obs.:	T	39,6	17,0	12,3	2,3	26,2	50,9	6,6	A	A2
1496	BRB02-1051-03- 01	MTBR95-133800 obs.:	T	37,2	16,3	12,0	2,3	21,1	56,0	7,3	A	A2

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
1497	BRB02-1083-03-01	CARVER*4 x [STONEWALL*2 x (HARTWIG*2 BR92-31380)]	M	38,6	17,0	11,3	2,2	23,4	55,4	6,7	A	A1
1498	BRB01-131-03-01	BRSMT PINTADO*2 x (EMBRAPA-20 x BRS89-81)	M	39,2	17,0	12,2	2,2	23,0	53,9	7,0	A	A2
1501	BRB01-212-03-01	BRS BORORO x [BRSMS CARANDÁ x (EMBRAPA-20 x BRS89-81)]	M	38,3	16,1	11,2	2,0	26,7	50,8	5,8	A	A1
1507	BRB01-129-03-01	BR 9 (SAVANA)*2 x (EMBRAPA-20 x BRS89-117)	M	39,2	16,7	12,2	1,8	26,6	52,0	6,0	A	A1
1509	BRB01-210-03-02	BRSMT CARANDÁ x (EMBRAPA-20 x BRS89-81)	M	37,5	16,9	11,8	2,4	22,8	55,6	6,4	A	A1
1515	BRB01-124-03-01	BRSMT CRIXÁS*2 x EMBRAPA-62	M	38,2	16,8	11,7	2,1	22,9	54,7	7,1	A	A2
1516	BRB01-132-03-01	BRSMA PATI*2 x FT 106	M	37,6	16,3	12,4	2,1	21,6	55,5	6,9	A	A1
1517	BRB01-124-03-02	BRSMT CRIXÁS*2 x EMBRAPA-62	M	40,4	16,5	12,6	2,1	22,0	54,5	7,4	A	A2
1518	BRB01-122/123-03-01	BRSMT CRIXÁS*2 x EMBRAPA-62/ (EMBRAPA-20 x BRS89-87)	M	39,6	16,8	10,9	1,7	34,1	46,3	5,6	B	B1
1805	BRGO01-04629	PROG.MBR-30 x {PROG.MG-315 x [CONQUISTA*4 x (PROG.96-246 x VIR)]}	M	37,6	16,6	12,2	2,1	22,8	55,0	7,0	A	A2
1806	BRGO01-03110	BR96-9508 x {BR94-493*2 x [PROG.MBR-61*3 x (PROG.96-246 x BRS-133)]}	M	38,7	16,3	11,6	2,4	20,1	57,5	6,9	A	A1
1807	BRGO99-48-19-16-03-01	{[(BR-16 RCH*6 x OC-9) x PROG.96-246] x PROG.MG-315} x BR95-1709	T	38,9	16,8	11,9	2,1	21,9	56,0	6,6	A	A1
1812	BRGO02-14409	HTBR91-55292 x MONSOY-8411	M	39,8	16,8	12,3	2,0	27,3	51,6	6,1	A	A1
1813	BRGO02-14876	CRIXÁS x E99-189	M	38,9	16,9	12,5	2,1	25,0	52,4	6,5	A	A2
1814	BRGO02-1403	UIRAPURU x MONSOY-8411	M	38,3	16,6	12,4	2,3	20,7	56,3	6,9	A	A1
1815	GOBR99-01-006	EMGOPA-302 x CRIXÁS	P	37,6	16,4	13,9	2,3	23,4	52,3	6,5	A	A5
1831	GOBR99-08-018	GOBR95-2919 x EMGOPA-302	P	40,2	16,3	11,7	2,3	30,9	47,2	5,7	B	B1
1841	GOBR99-05-189	CRIXÁS x GOBR95-6224	M	38,6	15,8	11,8	2,0	22,0	56,0	7,0	A	A1
1842	GOBR99-02-189	EMGOPA-302 x CRIXÁS	P	39,3	16,8	10,3	1,6	29,4	51,1	6,5	B	B1
1850	BRGO01-00281	DKCX-458 x SUPREMA e DEKALB-480 x	M	39,0	17,1	10,7	2,2	26,4	52,7	6,4	A	A2
1851	BRGO01-01876	SUPREMA e BRQ94-1311 x CONQUISTA	M	39,2	16,3	11,8	2,3	26,8	51,5	6,4	A	A2
1852	BRGO01-00231	D. EDUARDO-411 x SUPREMA e DELIA x SUPREMA	M	39,5	17,4	12,2	2,2	25,1	53,0	6,6	A	A2
1853	BRGO01-02931	MTBR91-52292 x LA SUPREMA	M	36,8	17,0	11,8	2,1	28,3	50,1	6,1	A	A2

(Continua)

Apêndice A. Continuação.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
1855	BRGO01-01833	BRQ94-7951 x CONQUISTA	T	38,7	16,8	11,5	2,1	32,8	46,0	5,9	B	B1
1869	BRGO01-03094	BR82-1166*3 x (BR83-9547 x BR92-31969)	T	38,6	16,3	12,2	2,3	20,4	57,1	6,9	A	A1
1880	BRGO01-04949	BR-16*4 x OCEPAR-9	T	37,8	16,5	11,9	2,1	23,7	54,9	6,4	A	A1
1884	BRGO01-01020	EMGOPA-308 RCH x E96-781	T	38,0	16,3	12,3	2,0	21,9	56,4	6,7	A	A1
1885	BRGO01-01270	EMGOPA-313 RCH*2 x E96-527134	T	37,3	16,8	12,5	2,6	21,6	55,9	6,2	A	A1
1887	BRGO01-00100	(FT BAHIA x TRACY M) x E96-781	T	39,3	16,3	12,3	2,3	20,2	57,2	6,9	A	A1
1888	BRGO01-00782	UIRAPURU x E96-781	T	38,0	16,4	12,5	2,4	19,8	56,8	7,1	A	A1
1893	BR 00-923	FT-Jatoba x (OCEPAR-9 x Dourados)	M	39,4	16,5	11,9	1,8	32,4	43,3	7,3	B	B3
1895	BRGO215-2	BRQ94-1311 x VENCEDORA	T	38,8	16,1	11,3	1,9	24,2	54,1	7,2	A	A2
1899	GO 02-2052-03-33	(TRACAJAS x SAMBAIBA) x GOBR99-356010	T	40,7	17,1	11,0	2,0	32,2	48,0	5,6	B	B1
1970	BR00-12141	-	T	39,3	16,5	11,6	2,0	27,9	51,2	6,1	A	A2
1971	MGBR00-54417	-	T	40,2	16,7	11,8	2,2	21,9	56,6	6,6	A	A1
1972	MGBR00-5068	-	T	40,4	16,8	12,4	2,1	21,4	56,3	6,8	A	A1
2013	GOBR00-B2995.18.04.01	PINTADO NC x CONQUISTA	M	41,3	16,6	12,3	2,0	22,2	55,2	6,7	A	A1
2015	GOBR00-0712.133.04.01	LA SUPREMA x IAC-15	M	38,8	16,8	11,7	2,3	25,1	53,7	6,2	A	A1
2018	GOBR00-665.14.04.01	FT-109 x EMGOPA-313 RCH	M	39,9	17,1	10,7	2,2	37,4	42,7	5,6	B	B2
2019	GOBR01-889.73.04.01	GOBR94-122243 NC x EMGOPA-311	M	40,2	17,0	11,8	2,1	27,1	51,4	6,4	A	A2
2020	GOBR01-1026.38.04.01	(EMGOPA-311 x SANTA CRUZ) x OCEPAR-17	M	41,0	16,8	11,6	2,3	24,9	53,5	6,2	A	A1
2023	GOBR01-966.04.01	GOBR95-2019 x DM-339	M	40,3	16,3	13,5	2,2	22,2	53,8	6,7	A	A1
2025	GOBR01-1033.24.04.01	GOBR95-3622 x MONSOY-2002	M	37,7	17,2	12,1	2,2	26,0	52,3	6,1	A	A1
2026	GOBR00-640.42.04.01	SAMBAIBA x EMGOPA-302	M	38,6	16,7	11,4	2,2	21,9	56,3	6,9	A	A1
2114	BRGO02-2166-13	(MS/BR-39*7 x EMBR-20)*6 x (E96-246 x UIR)]	M	38,1	16,7	12,4	2,0	24,2	53,8	6,6	A	A1
2122	GO 00-0682-56	BRS-MILENA x LA SUPREMA	T	36,8	17,4	12,8	2,2	20,8	57,1	6,5	A	A1
2192	GOBR02-2234-9	TK-5 x CONQUISTA	M	40,6	17,1	12,0	2,1	28,9	49,1	6,3	B	B1
2194	GOBR02-2203-1	PI-159322 x EMGOPA-315	M	42,0	16,7	10,7	1,7	38,3	42,2	6,1	B	B2
2339	BRGO03-215.02.04.01	BRQ94-1311 x VENCEDORA	T	39,0	16,0	12,5	2,1	24,9	52,2	7,3	A	A2
2340	GOBR00-0083-103	EMGOPA-315 x BRSMG-68 (VENCEDORA)	T	40,9	16,6	11,6	1,8	21,0	58,6	6,6	A	A6
2372	GOBR97-085064	E93-402*3 x E94-64 x (FORREST*4 x PRGO) x {SHARKEY x [BR-16 x FORREST x PRGO]}	T	40,9	16,7	11,8	2,3	22,6	55,8	6,5	A	A1
2373	GOBR99-701006	(HARTWIG*4 x BR92-31820) x MTBR95-122068	T	38,2	17,0	12,3	2,0	21,3	55,7	7,4	A	A2

(Continua)

Apêndice A. Continuação.

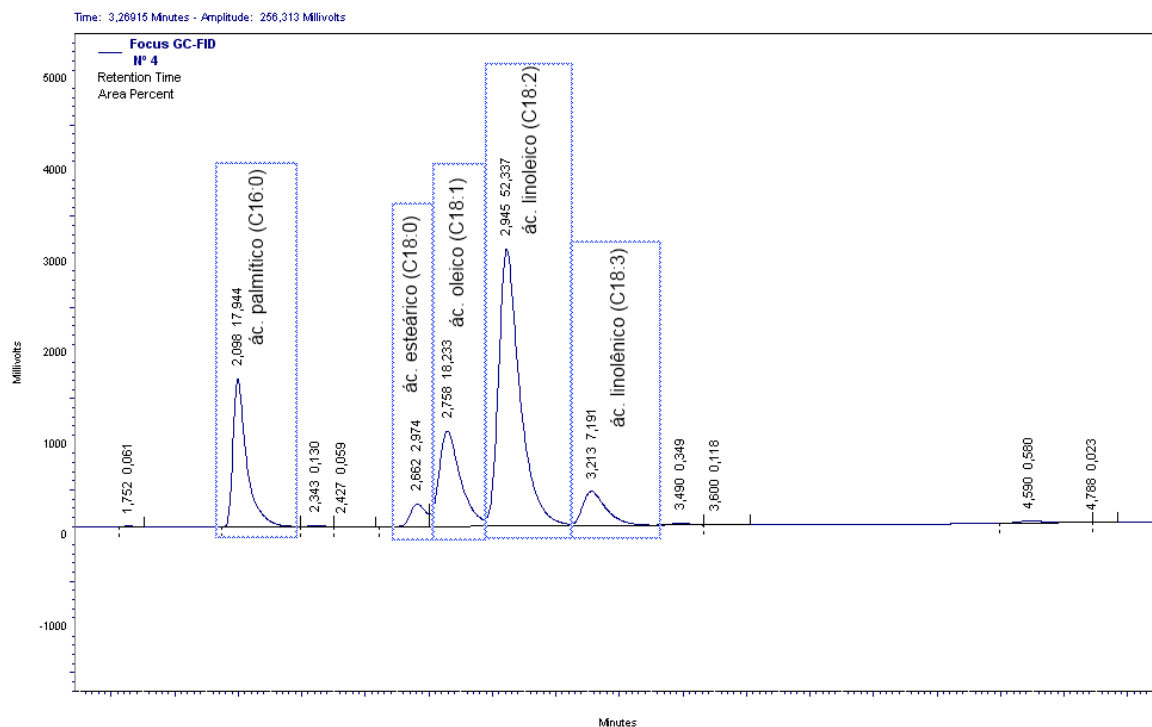
Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
2375	BRGO01-02725	BRQ94-7951 x VENCEDORA (resistente a ferrugem)	T	38,8	16,4	12,2	2,4	21,4	55,2	7,2	A	A2
2376	BRGO01-02849	MTBR91-55292 x DM-VITÓRIA	T	38,5	16,9	12,4	2,1	20,7	56,9	7,0	A	A1
2404	GOBR00-0197.74	FT-109 x PINTADO NC (peso 1240)	M	39,2	16,4	12,0	1,9	26,5	51,7	6,5	A	A2
2410	GOBR00-0457.26	GOBR95-13843 x CRISTALINA RCH (peso 1430)	M	38,5	17,4	12,2	2,2	23,4	54,1	6,5	A	A1
2413	GOBR01-0983.27	GOBR97-067033 NC x GOBR94-11191 x OCEPAR-207	T	37,5	16,4	12,7	2,1	19,5	58,0	7,0	A	A1
2414	GOBR01-0812.88	EMGOPA-314 x GOBR94-122243	T	37,6	17,1	12,8	3,3	20,4	52,7	6,0	A	A3
2415	GOBR01-0940.20	ST93-2051 x GOBR96-021931	T	38,8	17,2	12,2	2,7	22,0	54,8	6,7	A	A1
2416	GOBR01-0906.32	GOBR97-065029 NC x LA SUPREMA	T	38,7	16,6	11,3	2,0	27,1	52,9	6,0	A	A1
2419	GOBR00-0100.59	CONQUISTA x GOBR95-69943	T	39,0	16,8	12,0	1,8	26,9	52,6	5,8	A	A1
2420	GOBR00- 0479.120	EMGOPA-302 x GOBR95-2852	T	38,4	15,8	10,9	2,2	20,5	57,9	7,1	A	A4
2421	GOBR00-0325.24	EMGOPA-313 RCH x BR92-249	T	37,8	17,0	12,4	2,2	22,3	54,4	7,1	A	A2
2480	BRGO02-2172-3	{BR94-493*3 x [EMBR-61*3 x (E96- 246 x BRS 133)]} x BR96-9627-2	M	40,6	16,4	11,2	1,8	42,7	37,3	5,8	C	C1
2552	GRALHA	-	M	38,5	17,2	11,7	2,3	25,4	53,5	6,1	A	A1
2553	GOBR00- 464.007.01	-	T	39,0	16,8	12,4	2,0	21,2	56,3	6,8	A	A1
2591	GOBR03-2397-13	[MONSOY-6101 x (UFV TRIANGULO x BRSGO-PARAISO)]	T	39,2	17,4	9,9	1,7	44,4	36,8	5,1	C	C1
2592	GOBR04-3320-18	GOBR00-0410-053 x BR00-11971	T	40,2	16,3	11,4	1,9	27,9	51,7	6,2	A	A2
2593	GOBR03-2886-2	IAC-16 x GOBR95- 1337.00.02	T	41,1	16,9	11,9	2,1	26,0	53,2	6,0	A	A1
2594	GOBR03-3135-10	BR00-8118 x GOBR00-123-079	T	39,9	16,9	11,7	2,2	27,4	51,8	5,8	A	A1
2595	GOBR00-2727-14	SAMBAIBA x (TRACAJAS x PINTADO)	T	39,1	16,9	11,6	1,8	27,7	51,6	6,5	A	A2
2596	GONBR03-2692- 46	(PINTADO x TRACAJAS) x EMGOPA-302 BRSMG-68	M	39,2	16,6	11,7	2,2	28,4	49,8	6,3	A	A2
2599	GOBR03-2682-1	(VENCEDORA) x (PI-159322 x BRSGO- 204 (GOIANIA))	T	38,3	16,1	11,5	1,9	39,5	39,4	6,0	B	B3
2600	GOBR02-2672-35	BRSGO-MINEIROS x (BRS-134 x BRSGO- MINEIROS)	M	38,3	16,3	12,7	2,2	23,2	53,5	7,1	A	A2
2601	GOBR03-2786-28	BRS-213 x GOBR97- 056184	T	39,0	16,3	12,5	2,0	26,1	50,8	6,9	A	A2
2602	GOBR95-2802-21	GOBR95-13848 x OS 205	T	38,9	17,0	11,8	1,9	24,7	54,2	6,5	A	A1
2603	GOBR03-3061-22	GOBR00-345-028 x GOBR00-280-023	T	39,1	17,4	11,3	3,2	34,4	43,9	4,8	B	B5
2604	GONBR03-2732- 44	(PINTADO x TRACAJAS) x BRSGO-PARAISO	M	38,5	17,1	11,6	2,2	23,3	55,1	6,7	A	A1
2605	GONBR03-2772-5	BR86-1864 x BRSMS- BACURI	T	39,4	17,3	13,1	2,0	21,6	55,3	7,1	A	A1
2608	GOBR03-2934-2	GO99-03.013 x (GOBR97-061004 NC x PI-230970)	T	39,7	16,5	10,9	1,8	28,6	51,4	6,3	A	A2

(Continua)

Apêndice A. Conclusão.

Bag	Acesso	Genealogia	CM	PROT	OL	PAL	EST	OLC	LIN	LIL	G	SG
2609	GOBR03-3108-21	MONSOY-8411 x EMBRAPA-63 (MIRADOR) BRSGO- CHAPADOES x	T	38,7	17,1	12,3	2,3	25,8	51,4	6,6	A	A2
2613	GONBR03-3040- 11	(SAMBAIBA x TRACAJAS) x (PINTADO x TRACAJAS)	T	37,8	16,7	11,9	2,2	22,5	55,4	6,6	A	A1
2617	GOBR02-2247-14	SAMBAIBA x PI- 368038	M	37,5	16,3	12,8	1,9	20,6	56,3	7,3	A	A1
2619	GOBR03-2668-55	BRSGO-MINEIROS x (BRS-134 X OCEPAR-13)	M	37,9	17,5	11,5	1,8	31,2	47,7	5,9	B	B1

¹- Bag: número usado como código de cada acesso no banco de germoplasma; CM (ciclo de maturação) foi classificado em P, M e T, indicando, respectivamente, os ciclos precoce, médio e tardio; a análise de agrupamento resultou em cinco grupos (A, B, C, D e E), que se subdividiram, respectivamente, em (A1, A2 ... A8), (B1, B2 ... B5) e (C1 e C2); os grupos D e E não tiveram subdivisão.



Apêndice B. Análise de ácidos graxos em óleo de soja, em cromatógrafo gasoso (Focus GC Master), com identificação dos ácidos palmítico, esteárico, oleico, linoleico e linolênico e os respectivos valores de porcentagens e do tempo de retenção de cada ácido (éster metílico) pelo capilar do aparelho.

Apêndice C. Rotina computacional para implementação das análises estatísticas no aplicativo R, relativas ao estudo de caracterização de 527 acessos do banco de germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), acerca de sete atributos relacionados à qualidade do grão (teor de óleo – OL, teor de proteína – PROT, teores dos ácidos graxos palmítico – PAL, esteárico – EST, oleico – OLC, linoleico – LIN e linolênico – LIL, e teores totais dos ácidos saturados – SAT e poli-insaturados – POLI).

#Pacotes exigidos:

```
library(fpc)
library(A2R)
library(ape)
library(vegan)
```

1. Obtenção da matriz de distâncias

```
k1 <- read.table(file="medias.txt", header=TRUE, sep=";") # arquivo de médias
head(k1)
```

```
#CICLO BAG OL PROT PAL EST OLC LIN LIL SAT POLI FOL FPROT FPAL FEST FOLC FLIN FLIL FSAT FPOLI
# P 1 16.96 38.18 11.39 2.83 26.80 49.648 6.41 14.23 56.06 B C C B C B B B B
# P 2 16.67 38.59 11.69 1.90 24.08 53.91 7.04 13.60 60.95 B B C C C A A C A
```

#Leitura da matriz de distância de Mahalanobis obtida do programas Genes

```
d_mahal <- read.table(file="dist_mahalanobis.txt", header=F, sep=";")
d <- as.dist(d_mahal) #converte matriz simétrica em triangular inferior
```

Matriz de distâncias euclidianas a partir das médias

```
k2 <- k1[,3:9] # colunas das sete variáveis
k3 <- scale(k2) # padronização das variáveis
d_eucl <- dist(k3, method = "euclidean") # matriz de distâncias euclidianas
```

2. Agrupamento

```
fit <- hclust(d, method="ave") #Agrupamento hierárquico pelo critério de UPGMA
BAG <- k1[,2]
fit$labels <- BAG #identifica os acessos pelo número do BAG
```

```
coph <- cophenetic(fit) #matriz de distâncias cofenéticas
mantel(d, coph, method="spearman", permutations=10000) #correlação cofenética e
significância pelo teste de Mantel
```

3. Valores de Pseudo F

```
vetor_temp <- c(1:100)
for (i in 1:100)
{
  grupo_temp <- cutree(fit, k=i)
  vetor_temp[i] <- calinhara(k2, grupo_temp)
}
cont <- 1
pontos <- c(1:35)
for (i in 1:35)
{
  pontos [i] <- cont
  cont <- cont + 3
}
```

```
plot(vetor_temp [1:35],xlab='número de grupos', las=2, cex.axis=0.8, ylab='Pseudo F',xaxt=
"n", xaxs="p",xlim=c(35,1)) # gráfico grupos versus valor Pseudo F
abline(v=pontos, lty="dotted", col="lightgray")
axis(1, pontos, col.axis="1", las=1, cex.axis=0.8)
```

```
GRUPOS <-cutree(fit, k=5) # cinco grupos de similaridade
SUBGRUPOS <-cutree(fit, k=17) # dezessete subgrupos de similaridade
```

4. Dendrogramas

Dendrograma identificado os grupos

```
A2Rplot(fit, k=5,knot.pos="mean",boxes=T, main=" Grupos",show.labels=F,
lty.up="dotted",lty.down="solid", lwd.up=0.1, lwd.down=0.5, only.tree=F,
col.down= c("darkmagenta", 10, 12,"violetred","green4"))
```

Dendrograma identificado os subgrupos

```

A2Rplot(fit, k=17,knot.pos="mean",boxes=T, main="Subgrupos",show.labels=F,
lty.up="dotted",lty.down="solid", lwd.up=0.1, lwd.down=0.5, only.tree=F,
col.down= c("magenta","darkmagenta", 10, "darkred",10,"darkred",10, 12, "blue4", 12,
"blue4", 12, "blue4", 12, "blue4", "violetred3","darkgreen"))

# Dendrograma com identificação dos níveis (A, nível baixo; B, nível médio; e C, nível
alto) de cada variável
# Exemplo: para teor de óleo (coluna 12 ou fol), use "fact.sup=k1$fol"
A2Rplot(fit, k=17,fact.sup=k1$fol, outline=F,knot.pos="mean", show.labels=FALSE,
lty.up="dotted",lty.down="solid", lwd.up=1, lwd.down=1.3, only.tree=F,
col.down= c("magenta","darkmagenta", 10, "darkred",10,"darkred",10, 12, "blue4", 12,
"blue4", 12, "blue4", 12, "blue4", "violetred3","darkgreen"))

## 5. Boxplot

# Boxplot dos grupos
a<- cbind(BAG, GRUPOS)
grupos <- merge(k1[,2:11], a, by.x = "BAG", by.y = "BAG") # relação dos acessos com
repectivas médias originais e grupo

# Exemplo para a variável teor de óleo (OL)
bp<-boxplot(OL~GRUPOS, data=grupos, pch=8,cex=0.5, bg = "pink" , medcol = 1,main="ÓLEO",
cex.main=1, cex.lab=0.8, ylab="%", border=c(1:11), lwd=0.8)
meanOL <- tapply(grupos$OL,grupos$GRUPOS, mean)
bpw <- .8
for(i in 1:ncol(bp$stats))
  rect(i-bpw/2, bp$stats[2, i], i+bpw/2, bp$stats[4, i],
      density=30, col=i, angle=45, border=i, lwd=0.1)
points(seq(bp$n), meanOL, pch = 21, bg = c(1:5), lwd=1)

## 6. Análise em Componentes Principais

cp <- prcomp(k2, scale= T, center=T)
loads <- cp$r
escores<- cp$x
summary(cp)

## 7. Análise de variáveis canônicas

#leitura dos escores e dos coeficientes de ponderação obtidos no Genes
vc <- read.table(file="scores_vc.txt", header=TRUE, sep=";") # escores a análise em
variáveis canônicas, importados do programa Genes
ld_vc <- read.table(file="loads_vc.txt", header=TRUE, sep=";") # coeficientes de ponderação
rownames(ld_vc)<-c("OL", "PROT","PAL","EST", "OLC", "LIN", "LIL")

plot(ld_vc[, 1], ld_vc[, 2], pch = " ", cex=1, xlab="vc1 (58,5%)", ylab="vc2 (23,3%)")
abline(h = 0, v = 0, lty="dashed", col="gray")
text(ld_vc[, 1], ld_vc[,2], label = rownames(ld_vc), col = c("sienna3", "purple",
12,10,"violetred","green4", "darkmagenta"))

# dispersão dos escores na vc1 e vc2
plot(vc[, 1], vc[, 2], pch=8, cex=0.7, lwd=0.8, col=12, xlab="vc1 (58,5%)", ylab="vc2
(23,3%)", cex.main=1)
abline (v=0, h=0,lty="dotted",col="grey" )

# dispersão dos escores na vc1 e vc2 com identificação dos grupos
plot(vc[, 1], vc[, 2], pch="", cex=0.7,lwd=1, col="blue", xlab="vc1 (58,5%)", ylab="vc2
(23,3%)", cex.main=1, font.main=1, cex.lab=1, font.lab=1, cex.axis=0.8, font.axis=1, ann=T,
axes=T, family="mono" )
abline (v=0, h=0,lty="dotted",col="grey" )
text(vc[, 1], vc[, 2],label=grupos$GRUPOS, cex=0.6, col = c(grupos$GRUPOS))
leg<-c("Grupo 1", "Grupo 2", "Grupo 3", "Grupo 4", "Grupo 5")
legend(20, 34, leg, cex=0.6, col=c(1:5), pch="*", pt.cex=1, bty="o")

# dispersão dos escores na vc1 e vc2 com identificação dos ciclos de maturação
c4 <- as.character(k1$CICLO)
c5 <- c(1:527)
for (i in 1:527)
{
  if (c4[i] == "P") c5[i] <- "green4"
  if (c4[i] == "M") c5[i] <- "slateblue3"
  if (c4[i] == "T") c5[i] <- "red"
}
c3 <- cbind(c4, c5)

plot(vc[, 1], vc[, 2], pch="", cex=0.7,lwd=1, col="blue", xlab="vc1 (58,5%)", ylab="vc2
(23,3%)", cex.main=1, font.main=1, cex.lab=1, font.lab=1, cex.axis=0.8, font.axis=1, ann=T,
axes=T, family="mono" )
abline (v=0, h=0,lty="dotted",col="grey" )
text(vc[, 1], vc[, 2], label = c3[,1], col=c3[,2], cex=0.6, font=2)
legend(15, 35, c("P-> ciclo precoce", "M-> ciclo médio", "T-> ciclo tardio"), cex=0.8,
col=c("green4", "slateblue3", "red"), pch="*", pt.cex=1, bty="n");

# dispersão dos acessos na vc1 e vc2 com identificação dos níveis de concentração para cada
acesso

```

```

# no script deve substituir fSAT pelas outras variáveis fPROT, fPAL, fEST...
z10 <- as.character(k1$fSAT)
z11 <- c(1:527)
for (i in 1:527)
{
  if (z10[i] == "A") z11[i] <- 12
  if (z10[i] == "B") z11[i] <- 10
  if (z10[i] == "C") z11[i] <- "green4"
}
z12 <- cbind(z11, z10)
zleg_g<-c("A-> alto","B-> médio","C-> baixo")
zleg_grup<-c( 12, 10, "green4")

plot(vc[, 1], vc[, 2], main="Ác. Saturados", pch=" ", cex=0.8,lwd=0.2, col="blue",
xlab='CP1', ylab='CP2', cex.main=1, font.main=1, cex.lab=1, font.lab=1, cex.axis=0.8,
font.axis=1, ann=T, axes=T, family="mono", )
text(vc[, 1], vc[, 2], label = z12[,2], col =z12[,1], cex=0.6, font=1)
legend(15, 30, zleg_g, cex=0.7, col=zleg_grup, pch="*", pt.cex=1, bty="n")

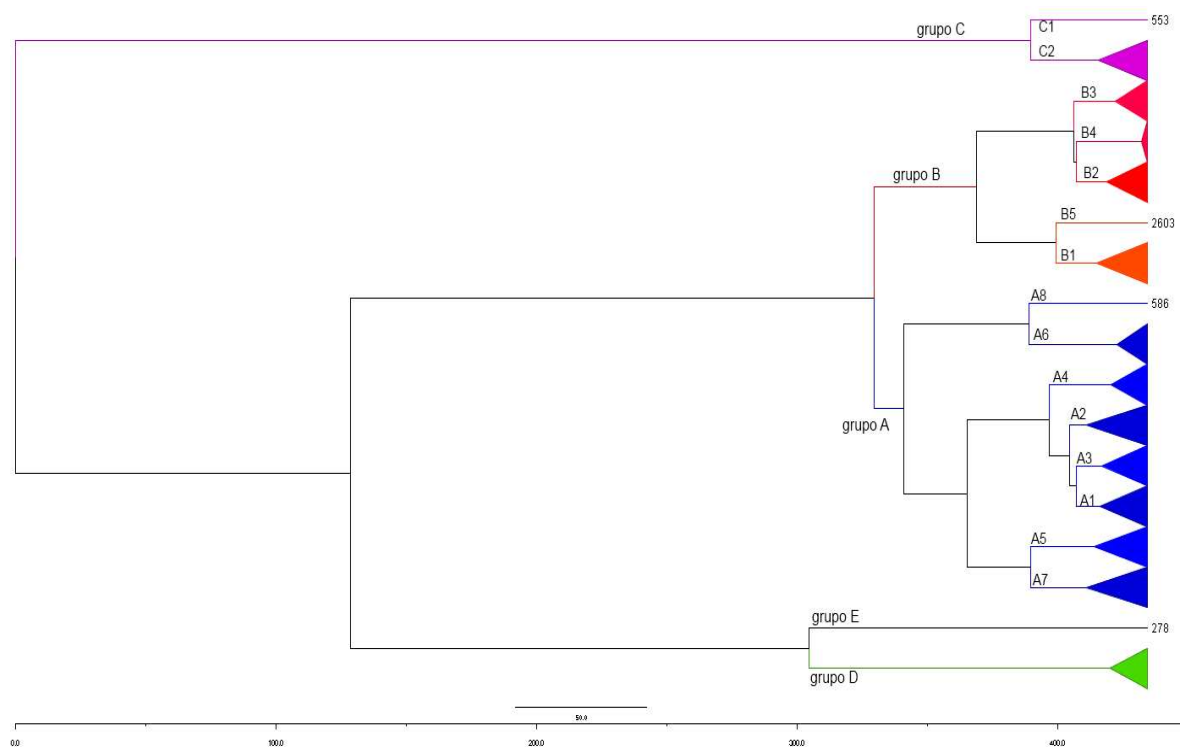
# 8. Correlação fenotípica

correl<-cor(k2)
print(correl)

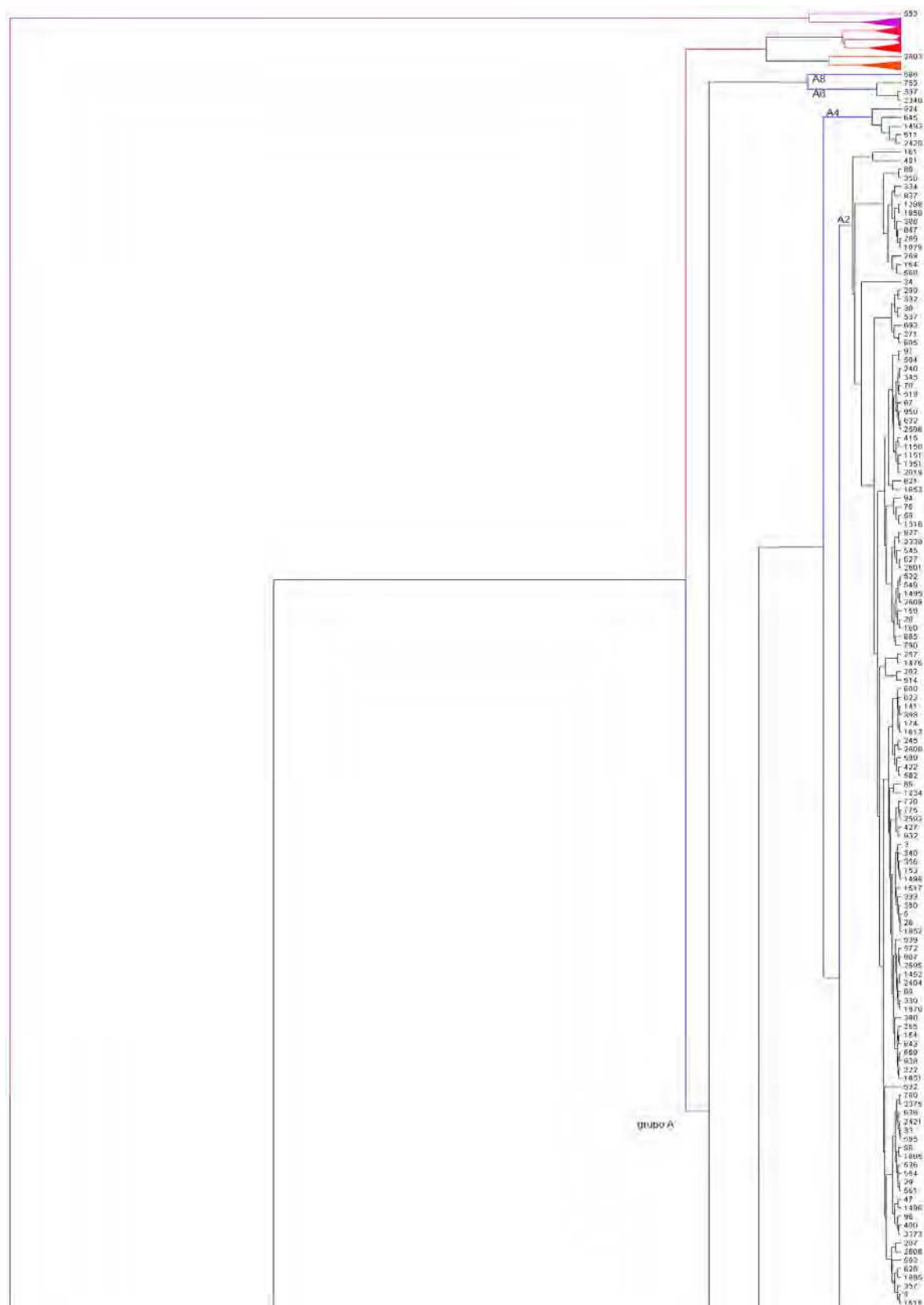
#dispersão da correlação
panel.cor <- function(x, y, digits=2, prefix="", cex.cor, ...)
{
  usr <- par("usr"); on.exit(par(usr))
  par(usr = c(0, 1, 0, 1))
  r <- (cor(x, y))
  txt <- format(c(r, 0.123456789), digits=digits)[1]
  txt <- paste(prefix, txt, sep="")
  if(missing(cex.cor)) cex.cor <- 0.8/strwidth(txt)
  text(0.5, 0.5, txt, cex = 1.2)
}

# na diagonal superior estão as correlações, e abaixo a dispersão identificando dos ciclos
de maturação
# para identificar outros fatores substituir no comando "unclass" a variável de
classificação
pairs(k1[,3:9], upper.panel=panel.cor, main="Correlação Fenotípica",cex.main=1,
cex.labels=1,
pch = ".",cex.axis=0.6, lwd=0.5, cex=3, col = c("red", "green3",
"blue")[unclass(k1$CICLO)])

```

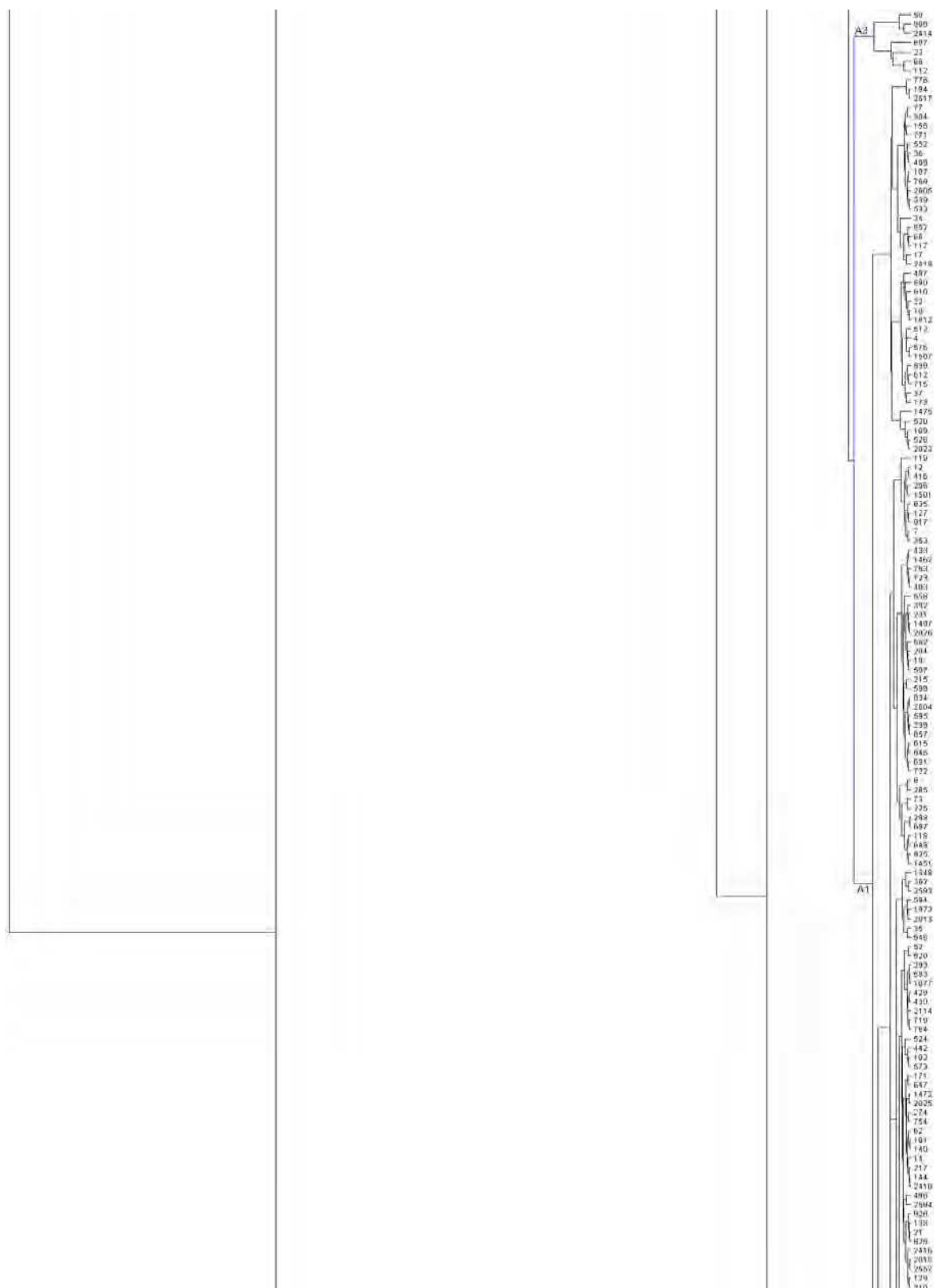


Apêndice D. Identificação dos grupos (A, B, C, D e E) e subgrupos de similaridade de 527 acessos do germoplasma de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), estabelecidos pelo agrupamento via UPGMA em relação à sete variáveis relacionadas à qualidade do grão.



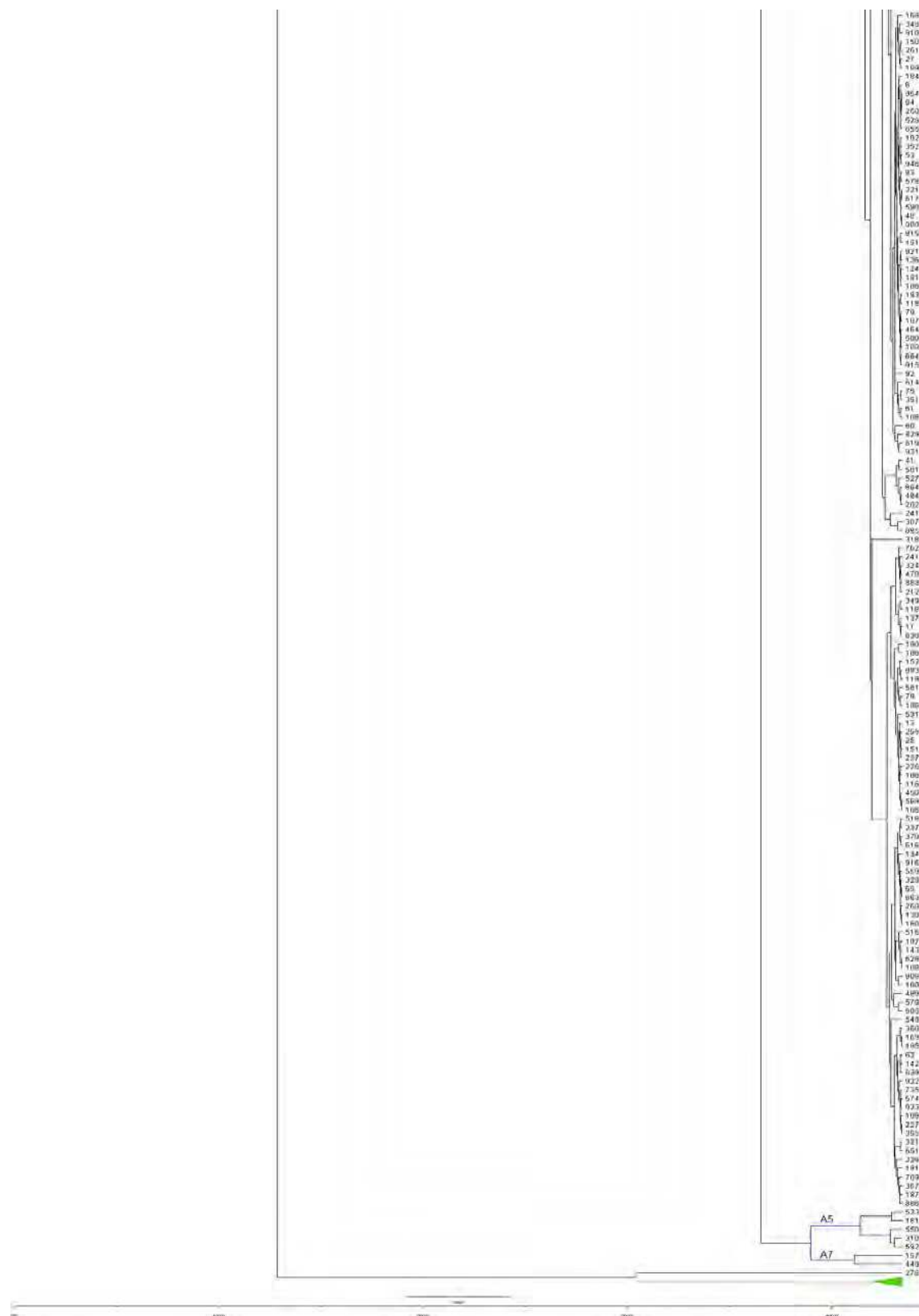
Apêndice E. Detalhamento da similaridade ou divergência entre os acessos de soja do grupo A (426 acessos), com subdivisão nos subgrupos A1, A2 ... A8, estabelecidos pelo agrupamento via UPGMA da subcoleção de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO) em relação à sete variáveis relacionadas à qualidade do grão.

(Continua)



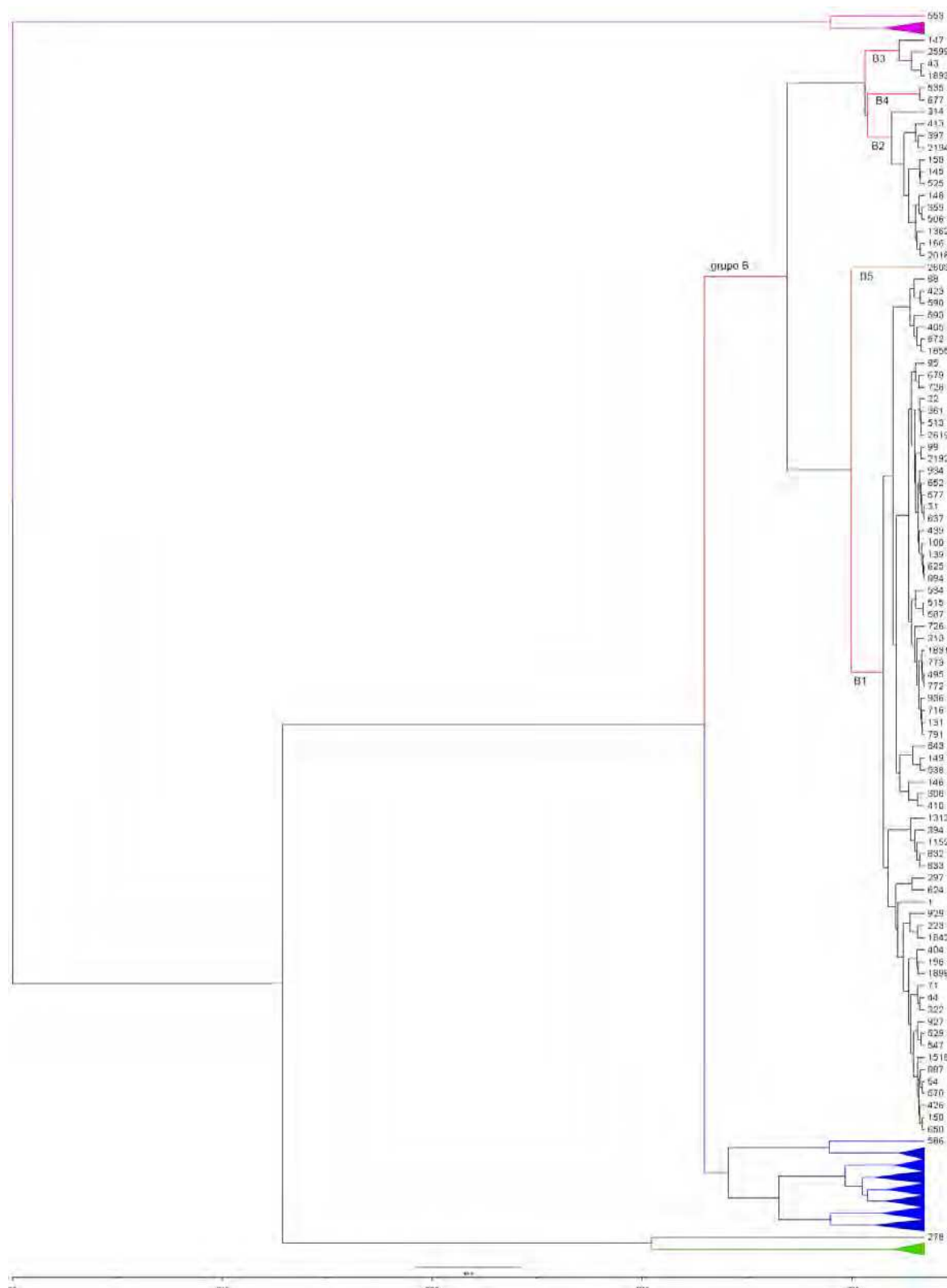
Apêndice E. Detalhamento da similaridade ou divergência entre os acessos de soja do grupo A (426 acessos), com subdivisão nos subgrupos A1, A2 ... A8, estabelecidos pelo agrupamento via UPGMA da subcoleção de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO) em relação à sete variáveis relacionadas à qualidade do grão.

(Continuação)

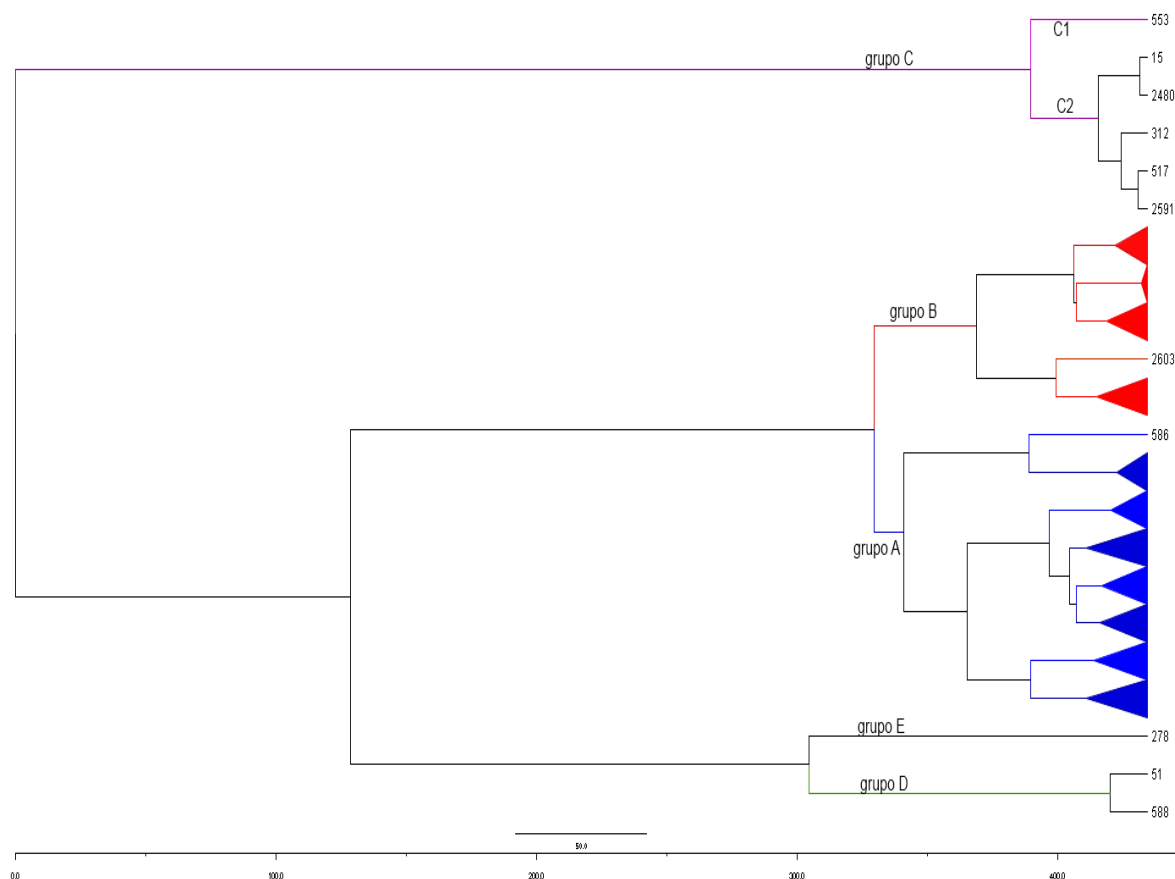


Apêndice E. Detalhamento da similaridade ou divergência entre os acessos de soja do grupo A (426 acessos), com subdivisão nos subgrupos A1, A2 ... A8, estabelecidos pelo agrupamento via UPGMA da subcoleção de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO) em relação à sete variáveis relacionadas à qualidade do grão.

(Conclusão)



Apêndice F. Detalhamento da similaridade ou divergência entre os acessos de soja do grupo B (92 acessos), com subdivisão nos subgrupos B1, B2 ... B5, estabelecidos pelo agrupamento via UPGMA da subcoleção de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO) em relação à sete variáveis relacionadas à qualidade do grão.



Apêndice G. Detalhamento da similaridade ou divergência entre os acessos de soja do grupo C (6 acessos), com subdivisão nos subgrupos C1 e C2; e dos grupos D e E, que não tiveram subdivisão, estabelecidos pelo agrupamento via UPGMA da subcoleção de soja da Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO) em relação à sete variáveis relacionadas à qualidade do grão.