

**UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA
FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CAMPUS DE JABOTICABAL**

**ESTIMATIVAS DE PARÂMETROS GENÉTICOS EM
LINHAGENS AVANÇADAS DE SOJA**

Bruno Henrique Pedroso Val

Engenheiro Agrônomo

2014

**UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA
FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CAMPUS DE JABOTICABAL**

**ESTIMATIVAS DE PARÂMETROS GENÉTICOS EM
LINHAGENS AVANÇADAS DE SOJA**

Bruno Henrique Pedroso Val

Orientador: Prof. Dr. Antonio Orlando Di Mauro

Coorientadora: Profa. Dra. Sandra Helena Unêda-Trevisoli

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

2014

V135e Val, Bruno Henrique Pedroso
Estimativas de parâmetros genéticos em linhagens avançadas de soja / Bruno Henrique Pedroso Val. -- Jaboticabal, 2014
xii, 68 p. : il. ; 28 cm

Dissertação (mestrado) - Universidade Estadual Paulista,
Faculdade de Ciências Agrárias e Veterinárias, 2014
Orientador: Antonio Orlando Di Mauro
Coorientadora: Sandra Helena Unêda-Trevisoli
Banca examinadora: Antonio Sergio Ferraudo, Everton Luis Finoto
Bibliografia

1. Análise de trilha. 2. *Glycine max*. 3. Índice de seleção. 4.
Melhoramento genético. 5. Mulamba e Mock. I. Título. II. Jaboticabal-
Faculdade de Ciências Agrárias e Veterinárias.

CDU 631.52:633.34

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação –
Serviço Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

CERTIFICADO DE APROVAÇÃO

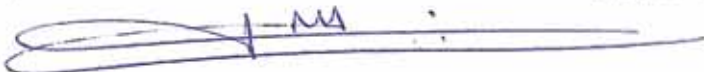
TÍTULO: ESTIMATIVAS DE PARÂMETROS GENÉTICOS EM LINHAGENS AVANÇADAS DE SOJA

AUTOR: BRUNO HENRIQUE PEDROSO VAL

ORIENTADOR: Prof. Dr. ANTONIO ORLANDO DI MAURO

CO-ORIENTADORA: Profa. Dra. SANDRA HELENA UNÊDA TREVISOLI

Aprovado como parte das exigências para obtenção do Título de MESTRE EM AGRONOMIA (GENÉTICA E MELHORAMENTO DE PLANTAS) , pela Comissão Examinadora:



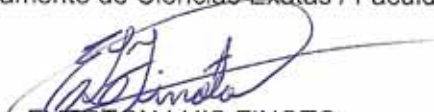
Prof. Dr. ANTONIO ORLANDO DI MAURO

Departamento de Produção Vegetal / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal



Prof. Dr. ANTONIO SERGIO FERRAUDO

Departamento de Ciências Exatas / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal



Prof. Dr. EVERTON LUIS FINOTO

Agência Paulista de Tecnologia dos Agronegócios / Pindorama/SP

Data da realização: 29 de julho de 2014.

DADOS CURRICULARES DO AUTOR

BRUNO HENRIQUE PEDROSO VAL – nascido em 9 de julho de 1990, em Lucélia, estado de São Paulo. Em agosto de 2008 ingressou no curso de Agronomia nas Faculdades Associadas de Uberaba (FAZU), Uberaba – MG. Em agosto de 2009 ingressou no curso de Tecnologia em Gestão Ambiental no Instituto Federal de Educação Ciência e Tecnologia do Triângulo Mineiro (IFTM), Uberaba – MG, obteve o título de engenheiro agrônomo em agosto de 2012 e o título de tecnólogo em gestão ambiental em 2013. Atuou como bolsista de iniciação científica modalidade PIBIC/IFTM no período de março de 2011 a fevereiro de 2012. Em agosto de 2012 ingressou no curso de Mestrado em Genética e Melhoramento de Plantas pela Universidade Estadual Paulista – Faculdade de Ciências Agrárias e Veterinárias (FCAV/UNESP), Jaboticabal – SP. Durante o mestrado atuou em projetos que abrangem as etapas do programa de melhoramento genético da soja da FCAV.

“Todo o homem que encontro me é superior em alguma coisa. E, nesse particular, aprendo com ele.”

Ralph Waldo Emerson

AGRADECIMENTOS

A Deus, pela força espiritual para a realização desse trabalho, pela coragem para superar os meus limites e chegar ao final desta jornada.

Aos meus pais Gláucia e Laudemir, pelo apoio, compreensão, ajuda. Vocês são responsáveis por cada sucesso obtido, cada degrau avançado durante todos esses anos, me ensinando a nunca desistir diante do primeiro obstáculo, meu exemplo de vitória, meus heróis e simplesmente aqueles que mais amo. Obrigado por estarem sempre comigo. Obrigado simplesmente por participarem comigo durante esta caminhada, me ajudando a construir os alicerces de um futuro que começa agora.

Ao meu irmão Victor pelo carinho, presença e por dividir comigo muitas das vitórias que hoje comemoro.

Aos meus avós Gildomar e Dáuria, Mário (In memoria) e Geny, tios e primos pela força e por sempre me desejarem o melhor.

A minha namorada Michelle, que esteve comigo durante esta jornada, me apoiando, pelo carinho, cumplicidade e amor a mim dedicados.

A Universidade Estadual Paulista – Faculdade de Ciências Agrárias e Veterinárias (FCAV/UNESP), Jaboticabal – SP e ao Departamento de Produção Vegetal.

Ao Prof. Dr. Antonio Orlando Di Mauro, pela amizade, conhecimentos transmitidos, orientação e realização deste trabalho.

A coordenadora Profa. Dra. Sandra Helena Unêda-Trevisoli, pela amizade, conhecimentos transmitidos e pelos esforços dedicados a mim neste trabalho.

A todos os professores do departamento de Produção Vegetal e também do Programa de Pós Graduação em Genética e Melhoramento de Plantas.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa de mestrado concedida e pelo recurso para execução deste trabalho.

Aos meus amigos José Arantes, Eduardo Henrique, Rodolfo Buzinaro, Fabiana Mota, Elise Pereira, Marcela Marconato, Daniel Leite, Alysson Jalles, Andréa Bastos, Paulo Selestrino.

A toda a equipe do programa de melhoramento genético da soja da FCAV/UNESP.

Ao Técnico Agrícola Geraldo Mangela de Assis pela amizade e dedicação às atividades do programa de melhoramento genético da soja FCAV/UNESP.

Aos membros das bancas de qualificação e defesa pelas contribuições e auxílios fornecidos.

Aos funcionários do Departamento de Fitotecnia, Rubens, Sebastião, Tito, Faco, Osmar, Mauro, Gabi e Mônica pela amizade, auxílio e os momentos de alegria.

Ao gerente da fazenda de pesquisa da FCAV/UNESP, Marcelo e toda a sua equipe que sempre esteve presente nos auxiliando com empenho e dedicação.

SUMÁRIO

RESUMO.....	xi
ABSTRACT	xii
CAPÍTULO 1 – CONSIDERAÇÕES GERAIS.....	1
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	2
2.1 A cultura da soja	2
2.2 Importância econômica.....	4
2.3 Melhoramento da cultura da soja.....	5
2.4 Caracteres agronômicos.....	6
2.4.1 Produtividade.....	6
2.4.2 Ciclo.....	7
2.4.3 Altura de inserção da primeira vagem	7
2.4.4 Altura da planta na maturação.....	8
2.4.5 Acamamento e valor agronômico	8
2.4.6 Número de ramos e vagens.....	9
2.4.7 Peso de cem sementes	9
2.4.8 Teor de óleo.....	9
2.5 Estimativas de parâmetros genéticos	10
2.5.1 Herdabilidade.....	11
2.5.2 Ganho com seleção	12
2.5.3 Correlações entre caracteres.....	12
2.6 Análise multivariada.....	13

2.6.1	Análise de componentes principais	14
2.6.2	Análise de agrupamento por método não hierárquico	14
2.6.3	Análise discriminante linear	15
3.	REFERÊNCIAS.....	16
CAPÍTULO 2 - ESTIMATIVAS DE PARÂMETROS GENÉTICOS E CORRELAÇÕES DE CARACTERÍSTICAS EM LINHAGENS AVANÇADAS DE SOJA		20
RESUMO.....		20
1.	INTRODUÇÃO	21
2.	MATERIAL E MÉTODOS	22
2.1	Manejo experimental e tratamentos utilizados.....	22
2.2	Caracteres agronômicos.....	24
2.3	Análises estatísticas	25
3.	RESULTADOS E DISCUSSÃO.....	26
4.	CONCLUSÕES	39
5.	REFERÊNCIAS.....	39
CAPÍTULO 3 – SELEÇÃO DE LINHAGENS AVANÇADAS DE SOJA POR		43
ABORDAGENS MULTIVARIADAS		43
RESUMO.....		43
1.	INTRODUÇÃO	44
2.	MATERIAL E MÉTODOS	46
2.1	Manejo experimental e tratamentos utilizados.....	46
2.2	Caracteres agronômicos.....	48
2.3	Análises estatísticas	49
2.3.1	Cálculo dos componentes principais.....	49
2.3.2	Cálculo da variância contida em cada componente principal	50
2.3.3	Correlações das características (variáveis) com os componentes principais.....	50

2.3.4 Análise de agrupamento por método não hierárquico	51
2.3.5 Análise discriminante linear	51
2.3.5.1 Análise discriminante por Kernel Gaussiano	52
3. RESULTADOS E DISCUSSÃO.....	52
4. CONCLUSÕES	67
5. REFERÊNCIAS.....	67

ESTIMATIVAS DE PARÂMETROS GENÉTICOS EM LINHAGENS AVANÇADAS DE SOJA

RESUMO – A soja é a oleaginosa mais importante cultivada no planeta. Dela são retirados inúmeros derivados, que compõem a alimentação humana e animal. Dentre os principais estão o farelo e o óleo que têm sua cotação influenciada pela razão demanda/produção no cenário mundial. Desde sua introdução no Brasil em 1882, a soja vem passando por um processo constante de melhoramento genético, o que permitiu que a cultura expandisse a área de cultivo que era basicamente a região Sul até a década de 70 para a região dos cerrados que corresponde ao Triângulo Mineiro, Mato Grosso do Sul, Mato Grosso, Goiás, Tocantins, sul do Maranhão, sul do Piauí e oeste da Bahia. Para um programa de melhoramento genético, a obtenção de estimativas de parâmetros genéticos é de grande importância, pois possibilita conhecer a magnitude dos efeitos que controlam o carácter além da possibilidade de prever qual será o ganho com a seleção. O ganho genético para o melhorista tem grande importância, pois com os valores deste parâmetro pode-se efetuar alterações no critério seletivo adotado, visando adequar a população selecionada aos objetivos do programa de melhoramento. Avaliar as características agrônomicas da soja e analisar todos os parâmetros ao mesmo tempo em função das informações coletadas durante o ensaio, se torna interessante quando o objetivo é selecionar genótipos superiores, uma vez que muitos fatores estão relacionados de maneira que seus diferentes efeitos são melhor interpretados se estudados em conjunto. A estatística multivariada permite identificar quais caracteres estão diretamente correlacionados com a produtividade de grãos e discriminar os genótipos levando em consideração a similaridade dos múltiplos caracteres avaliados, ajudando no processo seletivo. A análise, a descrição e a inferência são realizadas com base nas respostas simultâneas, valendo-se da estrutura de correlação entre as variáveis. Objetivou-se com este trabalho estimar parâmetros genéticos em linhagens avançadas de soja, identificar caracteres direta e indiretamente correlacionados com a produtividade de grãos e selecionar genótipos superiores de forma a maximizar o ganho para o conjunto de caracteres avaliados.

Palavras-chave: Análise de trilha, *Glycine max*, Índice de seleção, Melhoramento genético, Mulamba e Mock.

ESTIMATES OF GENETIC PARAMETERS IN SOYBEAN ADVANCED LINES

ABSTRACT – Soybean is the most important cultivated oilseed crop on the planet, it is extracted numerous derivatives that make up the human and animal food. Among the main ones are the bran and oil that have influence your quote according to the reason of demand and production on the world stage. Since its introduction in Brazil in 1882, the soybean has been undergoing a constant process of genetic breeding, which allowed the culture expand the area of cultivation which was basically South in the 70 years for the Cerrado region that corresponds to Triangulo Mineiro, Mato Grosso do Sul, Mato Grosso, Goias, Tocantins, Maranhao south, south west of Piaui and Bahia. For a breeding program, obtain estimates of genetic parameters is of great importance, as they allow to know the magnitude of the effects that control the character and the ability to forecast the gain with selection. The genetic gain for the breeder is very important, because with the values of this parameter can make changes to the selection criteria adopted in order to adjust the improved population for the objectives of the breeding program. Evaluate the agronomic characteristics of soybean and analyze all parameters at the same time in function on the information collected during the test becomes interesting when the goal is to select superior genotypes, since many factors are correlated so that their different effects are best interpreted if studied together. The Multivariate statistics allows to identify which characters are directly correlated with grain yield and discriminate the genotypes taking into account the similarity of multiple traits evaluated, helping in the selective process. The Analysis, description and inference are performed based on the simultaneous responses, taking advantage of the correlation structure between the variables. The objective of this study was to estimate genetic parameters in advanced soybean lines, identify characters directly and indirectly correlated with the grain yield and selecting superior genotypes maximizing the gain for the character set valuated.

Keywords: Genetic improvement, *Glycine max*, Index selection, Mulamba and Mock, Path analysis

CAPÍTULO 1 – CONSIDERAÇÕES GERAIS

1. INTRODUÇÃO

A soja tem se destacado no cenário mundial nas últimas décadas devido a sua capacidade de fornecer grandes quantidades proteína (40% em média) a um custo relativamente barato. Além da proteína a soja fornece o óleo (20% em média), que vem sendo utilizado largamente na alimentação humana, bem como para a produção de biocombustíveis (biodiesel).

Desde sua introdução no Brasil em 1882, a soja vem passando por um processo constante de melhoramento genético, o que permitiu que a cultura expandisse a área de cultivo que era basicamente a região Sul até a década de 70 para a região dos cerrados que corresponde ao Triângulo Mineiro, Mato Grosso do Sul, Mato Grosso, Goiás, Tocantins, sul do Maranhão, sul do Piauí e oeste da Bahia (SEDIYAMA et al., 2009).

Essa expansão da cultura gerou novas fronteiras agrícolas, levando consigo desenvolvimento econômico para diferentes regiões do país, possibilitando o desenvolvimento de cidades, o surgimento de indústrias nacionais e internacionais processadoras da leguminosa e de seus subprodutos e conseqüentemente um fortalecimento da economia local.

Existem inúmeros fatores que permitiram a expansão da cultura da soja para os cerrados, dentre elas podemos citar: Correção da acidez dos solos pelo processo da calagem e as condições do relevo que permitiram a mecanização de todo o processo produtivo da cultura. Todavia, o principal fator que possibilitou essa expansão foi o melhoramento genético que vem desenvolvendo cultivares cada vez mais adaptadas e produtivas às diferentes condições edafoclimáticas do território brasileiro.

Para um programa de melhoramento genético, a obtenção de estimativas de parâmetros genéticos é de grande importância, pois possibilitam conhecer a magnitude dos efeitos que controlam o carácter além da possibilidade de predizer qual será o ganho com a seleção.

Para estimar os parâmetros genéticos utilizam-se dos componentes de variância que permitem estimar a correlação entre os caracteres, a herdabilidade e

consequentemente o ganho com a seleção, os principais parâmetros que subsidiam o trabalho do melhorista durante o processo seletivo.

O ganho genético para o melhorista tem grande importância, pois com os valores deste parâmetro pode-se efetuar alterações no critério seletivo adotado, visando adequar à população selecionada aos objetivos do programa de melhoramento.

Avaliar as características agrônomicas da soja e analisar todos os parâmetros ao mesmo tempo em função das informações coletadas durante o ensaio, se torna interessante quando o objetivo é selecionar genótipos superiores, uma vez que muitos fatores estão relacionados de maneira que seus diferentes efeitos são melhor interpretados se estudados em conjunto (Vianna et al., 2013).

A estatística multivariada é uma área fascinante da estatística que estuda os fenômenos observando e analisando suas diversas respostas simultaneamente. Quando um pesquisador faz uma amostragem ou um experimento, dificilmente o interesse é focado em apenas uma variável. As variáveis mensuradas possuem relações entre si, que se forem exploradas, conduzirão a análises mais robustas e mais informativas (FERREIRA, 2011).

Objetivou-se neste trabalho estimar parâmetros genéticos em linhagens avançadas de soja, identificar caracteres direta e indiretamente correlacionados com a produtividade de grãos e selecionar genótipos superiores de forma a maximizar o ganho para o conjunto de caracteres avaliados.

2. REVISÃO DE LITERATURA

2.1 A cultura da soja

A soja é uma planta pertencente ao reino *Plantae*, divisão *Magnoliophyta*, classe *Magnoliopsida*, ordem *Fabales*, família *Fabaceae* (*Leguminosae*), subfamília *Faboideae* (*Papilionoideae*), gênero *Glycine*, espécie *Glycine max* e forma cultivada *Glycine max* (L.) Merrill. O gênero *Glycine* tem várias espécies originárias das regiões como África, Ásia oriental e Austrália. A palavra soja teve origem do japonês *Shoyu* (SEDIYAMA et al., 2009).

A soja (*Glycine max (L.) Merrill*) que hoje é cultivada mundo afora, é muito diferente dos ancestrais que lhe deram origem: espécies de plantas rasteiras que se desenvolviam na costa leste da Ásia, principalmente ao longo do Rio Amarelo, na China. Sua evolução começou com o aparecimento de plantas oriundas de cruzamentos naturais, entre duas espécies de soja selvagem, que foram domesticadas e melhoradas por cientistas da antiga China. Sua importância na dieta alimentar da antiga civilização chinesa era tal, que a soja, juntamente com o trigo, o arroz, o centeio e o milheto, era considerada um grão sagrado, com direito a cerimoniais ritualísticos na época da sementeira e da colheita (EMBRAPA, 2004).

A introdução da soja no Brasil deu-se por volta de 1882, e foi o professor Gustavo Dutra, da Escola de Agronomia da Bahia, o responsável pelos primeiros estudos com a cultura no país. Cerca de dez anos depois, o Instituto Agrônomo de Campinas (IAC), no Estado de São Paulo, também iniciou estudos para obtenção de cultivares aptos à região. Naquela época, porém, o interesse pela cultura não era pelo seu material nobre, o grão, era mais pela planta como uma espécie a ser utilizada como forrageira e na rotação de culturas (BONATO, 1987; SINDIMILHO E SOJA, 2011).

O cultivo da soja para produção comercial de grãos teve início no Rio Grande do Sul, por volta de 1935, cuja principal finalidade era para alimentação de suínos. A primeira exportação de soja brasileira ocorreu em 1938 para a Alemanha (SEDIYAMA et al., 2009). Nas estatísticas internacionais o Brasil começou aparecer como produtor de soja em 1949 (MIYASAKA, 1965).

A região sul foi responsável, até 1960 e 1970, por ser a produtora majoritária do país, sobretudo no Rio Grande do Sul e Paraná. A partir dos anos 80, a soja estendeu-se para o cerrado, uma vasta região que abrange o chamado polígono dos solos ácidos, ou seja: Triângulo Mineiro, Mato Grosso do Sul, Mato Grosso, Goiás, Tocantins, sul do Maranhão, sul do Piauí e oeste da Bahia. Com isso, a região do cerrado tornou-se a maior região produtora do país. A expansão para essa nova fronteira agrícola deveu-se, basicamente, aos estudos de fertilização dos solos do cerrado, à sua topografia plana e favorável à mecanização, e o desenvolvimento de plantas adaptadas à região (BONATO, 1987; SINDIMILHO E SOJA, 2011).

2.2 Importância econômica

A soja é a oleaginosa mais importante cultivada no planeta, dela são retirados inúmeros derivados, que compõem a alimentação humana e animal. Dentre os principais estão o farelo e o óleo que têm sua cotação influenciada pela razão demanda/produção no cenário mundial.

Segundo Sedyama et al. (2009), o grão possui em torno de 40% de proteínas de excelente qualidade e em torno de 20% de óleo, esses fatores criaram uma demanda pelo farelo da leguminosa que é utilizado principalmente na formulação de rações para aves e suínos.

Com o constante crescimento populacional e conseqüentemente a demanda por alimentos, a soja vem a ser o alimento do futuro. De acordo com o Sindimilho e soja (2011), a leguminosa é riquíssima em vitaminas do complexo B, A, C e E além dos minerais como o cálcio, ferro, fósforo e potássio. Em termos de proteínas um único kg do alimento equivalem a 2kg de carne de vaca, 5kg de arroz, 3kg de feijões/grãos ou ainda a 11 litros de leite.

Numa época em que se torna indispensável encontrar soluções que permitam alimentar uma população mundial cujo crescimento vertiginoso parece ser inversamente proporcional à disponibilidade de recursos, a soja e seus derivados apresentam-se como uma luz ao fundo do túnel, graças à sua capacidade de satisfação das necessidades calórico-proteicas do homem e ao seu baixo custo de produção (SINDIMILHO E SOJA, 2011).

O Brasil é o segundo maior produtor da cultura da soja, ficando atrás somente dos Estados Unidos. A produção brasileira estimada pela CONAB na safra 2013/2014 é de 85,442 milhões de toneladas representando um incremento de 4,8% em relação à safra passada, um aumento de 7,4% em relação a área plantada e uma redução prevista na produtividade em 2,4% (CONAB, 2014).

O crescimento do complexo soja no Brasil e no mundo é tanto que do ano de 2002 ao ano de 2012 o Brasil saltou de 30,413 milhões de toneladas exportadas para 48,956 milhões de toneladas, isso equivale a 26,114 bilhões de dólares no ano de 2012. A soja correspondeu na balança do agronegócio brasileiro no ano de 2011 a

26% das exportações do setor agrícola e 9,4% das exportações brasileiras (COSTA, 2012).

De acordo com os dados da Conab (2014), a área plantada com soja no estado de São Paulo na safra 2013/2014 foi de 734.500 hectares, representando apenas 2,46% da área plantada no País no mesmo período. Segundo a Faesp (2011) as principais áreas produtoras de soja no Estado estão nas regiões de Assis, Orlandia, Itapeva, Barretos, Ourinhos e Avaré. A Faesp (2011) afirma ainda que apesar do elevado nível de tecnologia empregado, as lavouras de soja no Estado não competem em termos de produtividade e rentabilidade com as lavouras da região Centro-Oeste. O clima, os ganhos de escala e o menor valor das terras naquela região conferem uma maior rentabilidade à produção da região Centro-Oeste. No entanto, a soja continua sendo uma alternativa viável no Estado.

2.3 Melhoramento da cultura da soja

Segundo Magalhães (1981), as pesquisas com soja no Brasil iniciaram-se na década de 1930, na antiga Estação Experimental Fitotécnica das colônias, no município de Veranópolis, Rio Grande do Sul.

Em 1946, as pesquisas foram estendidas a outras estações experimentais e também à Universidade Federal do Rio Grande do Sul, por meio da Faculdade de Agronomia, que intensificou seu programa de pesquisa, realizando experiências em colaboração com o Instituto Agrônomo de Campinas, em São Paulo, e a própria Secretaria de Agricultura do Estado do Rio Grande do Sul, com introduções de coleções procedentes da Universidade Federal de Viçosa (SEDIYAMA et al., 2005).

Os primeiros trabalhos constaram da manutenção de uma pequena coleção de cultivares, compreendendo o Laredo, Tokio, Biloxi, Prolific, Mammouth e outros, nos quais se observavam as características botânicas e agronômicas, realizando-se com eles, ensaios de competições para avaliação do rendimento. Em 1947, foram feitas as primeiras hibridações, na tentativa de obter cultivares por meio de cruzamentos artificiais. Na década de 50 ampliaram-se as pesquisas com soja, aumentando as coleções de cultivares, introduzindo vários tipos de cultivares e tipos de soja com o

objetivo de selecionar ou desenvolver cultivares com elevada produtividade, altura de planta e inserção das primeiras vagens adequadas à mecanização da lavoura, resistência ao acamamento e deiscência natural das vagens, com resistência às doenças, boa qualidade de sementes, alto rendimento de óleo e proteína (SEDIYAMA et al., 2005).

De acordo com Magalhães (1981), em 1960 surgiu a primeira variedade que foi lançada com o nome de Pioneira.

O resultado dessas pesquisas foi a substituição gradativa de cultivares ultrapassados por novos cultivares obtidos de seleções individuais, os quais contribuíram sobremaneira para aumentar o rendimento médio das lavouras no Rio Grande do Sul (FERES; GOMES, 1981).

A soja é considerada como planta de dias curtos (noites longas); por isso grande parte da área mundial cultivada com essa cultura está localizada em latitudes maiores que 30°, onde prevalecem condições de clima temperado. O Brasil representa uma exceção dentro desse contexto. Nas duas últimas décadas, com a expansão da cultura em grandes áreas dos Cerrados, o processo produtivo agrícola com a soja ocorre predominantemente em regiões de climas tropical e subtropical. A adaptação da soja às condições de latitudes das regiões Centro-Oeste, Norte e Nordeste foi um dos grandes desafios enfrentados pelo programa de melhoramento da Embrapa Soja. Essa expansão só foi possível pelo descobrimento do período juvenil longo, o que possibilitou desenvolvimento de cultivares melhoradas e adaptadas inclusive para zonas equatoriais. Atualmente, cerca de metade da produção brasileira é colhida nos estados compreendidos em latitudes menores que 20°. As regiões situadas em latitudes menores que 10° representam atualmente a área de expansão da soja, especialmente nos estados do Maranhão, Piauí, Tocantins e Pará (ALMEIDA et al., 1999).

2.4 Caracteres agronômicos

2.4.1 Produtividade

Nos programas de melhoramento de soja visando ao desenvolvimento de novos cultivares, além dos objetivos associados a caracteres agrônômicos, resistência a pragas e doenças, adaptação ao meio ambiente, tolerância aos herbicidas e obtenção de sementes com melhor qualidade fisiológica, genética, nutricional e bioquímica, evidentemente é necessário e de fundamental importância que o cultivar apresente produtividade economicamente viável (SEDIYAMA et al., 2005).

Em 1993 a produtividade média brasileira era em torno de 2000 kg/ha de acordo com Roessing e Guedes (1993), já no ano de 2012 a produtividade média brasileira fechou em 2938 kg/ha de acordo com a Conab (2014), isso equivale a um ganho em produtividade de aproximadamente 47%. Estima-se que metade desse ganho seja devido ao melhoramento genético.

2.4.2 Ciclo

O ciclo refere-se ao número de dias entre a emergência e a maturação de 95% das vagens (estádio R8). Até o ponto de colheita, dependendo das condições de umidade e de temperatura, pode ainda durar cerca de sete a dez dias. Na maioria dos cultivares, o ciclo varia entre 75 e 210 dias. Contudo, dentro de um mesmo cultivar, esse número pode também oscilar, dependendo da latitude, da altitude, da temperatura e da época de semeadura (SEDIYAMA et al., 2005).

O ciclo das cultivares no Brasil era denominado basicamente de: Precoce, médio e tardio. Com a necessidade de padronizar a nomenclatura adotou-se o sistema já utilizado nos Estados Unidos e na Argentina, denominado grupo de maturidade variando de VI à VIII no Brasil. Um estudo foi feito por Alliprandini et al. (2009), no qual foi estimado uma equação de regressão para a Região Sul e outra para a região Centro-Oeste, permitindo calcular qual o grupo de maturidade do genótipo em função do número de dias para sua maturação.

2.4.3 Altura de inserção da primeira vagem

Os cultivares mais produtivos tendem a formar vagens ao longo de toda a planta, inclusive até a parte mais baixa. Todavia, é desejável, para que não haja perda

na colheita pela barra de corte, que a inserção das primeiras vagens na haste principal da planta esteja pelo menos cerca de 10 a 12cm acima da superfície do solo de topografia plana e cerca de 15 cm para cultivo em solos mais inclinados (SEDIYAMA et al., 2005).

Almeida et al. (2011), relata que selecionando plantas de florescimento tardio e de altura de inserção da primeira vagem mais alta, seria possível o melhoramento indireto para o carácter produtividade de grãos.

2.4.4 Altura da planta na maturação

A altura média da planta pode variar de 20 à até mais de 150cm, conforme o cultivar, época de plantio, fertilidade, distribuição de chuvas, temperatura, luminosidade, fotoperíodo etc. A altura mínima desejável é em torno de 50 a 60 cm, no entanto, a planta da soja estando entre 70 a 80 cm de altura permite uma eficiente operação de colheita. Plantas muito acima de 100 cm tendem ao acamamento e, além de dificultarem a colheita também tendem a produzir menos (SEDIYAMA et al., 2005).

2.4.5 Acamamento e valor agronômico

O acamamento das plantas pode ocasionar perdas de grãos durante a colheita mecanizada. As plantas altas e, ou, de caule muito fino tendem ao acamamento com relativa facilidade do mesmo modo que cultivares de caule excessivamente grossos mesmo que produtivos atrapalham a colheita mecanizada (SEDIYAMA et al., 2005).

É avaliado no estágio R8 da planta por meio de uma escala de notas visuais, variando de 1 (todas as plantas eretas) a 5 (todas as plantas acamadas) (FEHR & CAVINESS, 1977).

O valor agronômico é avaliado no estágio R8 de desenvolvimento da planta, por meio de uma escala de notas visuais, a qual varia de 1 (plantas com características agronômicas ruins) a 5 (plantas com ótimas características agronômicas), sendo a nota atribuída representativa de um conjunto de caracteres visuais (arquitetura da planta, quantidade de vagens cheias, vigor e sanidade da planta, debulha prematura

das vagens, acamamento e retenção foliar na maturidade). De um modo geral a nota é particular dependendo do ponto de vista de cada melhorista.

2.4.6 Número de ramos e vagens

O número de ramos é obtido através da contagem feita a partir de uma amostra de plantas coletadas em cada parcela experimental. Navarro Júnior; Costa (2002) relatam que esta característica tem relação com o potencial produtivo da planta por apresentar maior superfície fotossintetizante e também maior área para o surgimento de flores.

O número de vagens também é obtido através da contagem feita a partir de uma amostra de plantas de cada parcela experimental. É uma característica que tem grande influência sobre a produtividade do genótipo.

Iqbal et al. (2003) observaram através de estudos envolvendo a análise de trilha, que quando a produtividade de grãos era considerada como característica principal, as principais variáveis que tinham efeito direto na produtividade era o número de vagens, o peso de cem sementes e a altura da planta na maturação.

2.4.7 Peso de cem sementes

O peso de cem sementes é uma das variáveis que mais influencia a produtividade de grãos na soja, porém não deve ser a única característica a ser dada atenção na hora de buscar por genótipos mais produtivos. O número de vagens por planta bem como a altura da planta na maturação também influencia o carácter produtividade de grãos, em vista deste ser de ordem quantitativa ou seja governado por vários genes e com forte influência ambiental.

2.4.8 Teor de óleo

O óleo extraído da semente de soja é um dos produtos mais importantes utilizados no preparo de alimentos, sendo o tipo mais consumido no mundo. Em média, os cultivares comerciais apresentam cerca de 20% de óleo; no entanto,

existem introduções com percentuais que variam de 13 a 28%. Embora o conteúdo de óleo possa ser efetivamente modificado pelo melhoramento, este caráter não tem sido enfatizado no desenvolvimento de cultivares porque o preço pago pela saca de soja ainda não é baseado na composição da semente (SEDIYAMA et al., 2005).

Atualmente o melhoramento genético da soja busca obter genótipos com maiores teores de óleo e proteína, isso é possível até um certo ponto visto que os dois caracteres possuem uma correlação negativa. Sabe-se ainda que a produtividade possui correlação negativa com proteína e positivamente correlacionada com o óleo (Wilcox; Goudong, 1997).

2.5 Estimativas de parâmetros genéticos

O sucesso do melhoramento genético de qualquer caráter requer, obrigatoriamente, que este seja herdável e que haja variação genética na população em que se pratica seleção. No estudo da herança e da variação de caracteres quantitativos adota-se o modelo básico $F = G + M$, que define o valor fenotípico (F), estimado a partir dos dados diretamente mensurados nos indivíduos, como o resultado da ação do genótipo (G), ou valor genotípico, sob influência do meio (M) (CRUZ, 2012).

Com base na fórmula descrita acima temos que a variância fenotípica (VF) é composta pela variância genotípica (VG) e pela variância atribuída aos desvios proporcionados pelo ambiente (VM). A variância genotípica por sua vez, é estabelecida por três outros componentes: A variância aditiva (VA), a variância atribuída aos desvios de dominância (VD) e a variância atribuída aos efeitos epistáticos resultantes de interações inter-alélicas (CRUZ, 2012).

Os componentes de variâncias podem ser estimados de diversas formas, porém a mais comum em programas de melhoramento genético é através da tabela anova, onde igualam-se os valores do quadrado médio às respectivas esperanças matemáticas.

De posse das informações sobre os componentes de variância, é possível obter estimativas dos parâmetros genéticos. A obtenção de tais estimativas são de grande importância para programas de melhoramento genético, pois possibilitam conhecer a

magnitude dos efeitos que controlam o carácter além da possibilidade de predizer qual será o ganho com a seleção.

2.5.1 Herdabilidade

A herdabilidade pode ser conceituada como a proporção da variância genética presente na variância fenotípica total. Portanto, ela estima a confiabilidade do valor fenotípico como indicador do valor reprodutivo. Por essa razão a herdabilidade participa quase sempre das expressões relacionadas com a predição de ganho dos diferentes métodos de melhoramento e, como consequência das decisões que os melhoristas tomam na condução dos seus programas de seleção (RAMALHO et al., 2012a).

Existem diversos métodos na literatura de como estimar a herdabilidade, os mais comuns são: Método da herdabilidade realizada, Método da regressão pai filho e o método dos componentes de variância.

A herdabilidade pode ser estimada no sentido amplo e no sentido restrito. No sentido amplo é a razão da variância genética pela variância fenotípica, no sentido restrito é a razão da variância aditiva pela variância fenotípica.

O coeficiente de herdabilidade pode variar de zero a um, tanto para o sentido amplo quanto para o restrito. No caso da herdabilidade ser igual a 1, as diferenças fenotípicas entre os indivíduos são causadas unicamente pela variância genotípica entre os mesmos. Quando a herdabilidade for igual a 0, significa que a variabilidade do carácter não tem origem genética. Neste caso não existe correlação alguma entre valor genético e valor fenotípico da unidade de seleção (ALLARD, 1971).

A herdabilidade no sentido restrito é mais importante, principalmente no melhoramento de plantas autógamas porque o que é hereditário (passado para os descendentes) é o alelo e não o genótipo em si, sendo assim os demais componentes que compõem a variância genotípica não são herdados. O único caso em que é interessante calcular a herdabilidade no sentido amplo é quando se trabalha com culturas de propagação vegetativa, onde através do clone, é transmitido o componente de variância genotípica inteiramente para o descendente.

2.5.2 Ganho com seleção

A possibilidade de predição dos ganhos obtidos por uma determinada estratégia de seleção constitui-se em uma das principais contribuições da Genética Quantitativa para o melhoramento. Com base nestas informações é possível orientar de maneira mais efetiva o programa de melhoramento, predizer o sucesso do esquema de seleção adotado e determinar, de forma científica, quais as técnicas que podem ser mais eficazes (CRUZ; REGAZZI, 1994). O ganho genético ou resposta à seleção é a diferença entre a média fenotípica dos descendentes dos progenitores selecionados e a média da geração parental (população) antes da seleção (FALCONER; MACKAY, 1996).

O ganho genético para o melhorista tem grande importância, pois com os valores deste parâmetro pode-se efetuar alterações no critério seletivo adotado, visando adequar à população selecionada aos objetivos do programa de melhoramento. Assim, a conexão entre o ganho e o diferencial de seleção expresso provém diretamente do significado de herdabilidade (RAMALHO et al., 2012b).

O ganho na seleção é uma função apenas da porção genética da variabilidade total. A obtenção de informações a respeito dos parâmetros do complexo genótipo ambiente é, portanto, de grande importância para o melhorista. Nas espécies cultivadas, quanto mais precisas forem essas estimativas, melhores serão as previsões do melhorista. Possibilitando assim, antever melhor o ganho esperado com a aplicação de diferentes tipos e intensidades de seleção (ALLARD, 1971).

2.5.3 Correlações entre caracteres

A correlação entre duas variáveis x e y pode ser medida através do coeficiente de regressão linear $r_{(xy)}$ que é a covariância de xy dividida pelo produto dos desvios padrões de x e de y .

Características correlacionadas são de interesse, por três razões principais. Primeiramente, em conexão com as causas genéticas de correlação, por meio da ação pleiotrópica dos genes. Em segundo lugar em conexão com as mudanças efetuadas pela seleção, é importante conhecer como o melhoramento de uma

característica pode causar mudanças simultâneas em outras. Em terceiro lugar, em conexão com a seleção natural, a relação entre uma característica métrica e o poder adaptativo é o agente primário que determina as propriedades genéticas desta característica, em uma população natural (FALCONER; MACKAY, 1996).

A importância do conhecimento da associação entre caracteres nos trabalhos de melhoramento se dá principalmente em caracteres de baixa herdabilidade. Assim, o caráter correlacionado com outro de alta herdabilidade e de fácil mensuração pode ser mais conveniente e conduzir a progressos mais rápidos para o melhoramento (CRUZ; REGAZZI, 1994).

2.6 Análise multivariada

A análise multivariada se refere a todas as técnicas estatísticas que simultaneamente analisam múltiplas medidas sobre indivíduos ou objetos sob investigação. Assim qualquer análise de mais do que duas variáveis pode ser considerada multivariada. Muitas técnicas multivariadas são extensões da análise univariada (análises de distribuições de uma única variável) e da análise bivariada (classificação cruzada, correlação, análise de variância, e regressão simples usadas para analisar duas variáveis) (HAIR et al., 2009).

A estatística multivariada tem como ingrediente básico a resposta de várias variáveis nas unidades amostrais ou experimentais. A análise, a descrição e a inferência são realizadas com base nas respostas simultâneas, valendo-se da estrutura de correlação entre as variáveis (FERREIRA, 2011).

Alguns pesquisadores usam multivariada simplesmente para se referirem ao exame de relações entre mais de duas variáveis. Para ser considerada verdadeiramente multivariada, todas as variáveis devem ser aleatórias e inter-relacionadas de tal maneira que seus diferentes efeitos não podem ser significativamente interpretados em separado (HAIR et al., 2009).

2.6.1 Análise de componentes principais

A análise de componentes principais é uma técnica multivariada de modelagem da estrutura de covariância. Foi introduzida por Pearson em 1901 e desenvolvida de forma independente por Hotelling em 1933 (FERREIRA, 2011).

É uma técnica para redução do hiperespaço do conjunto original de variáveis em dimensões menores e interpretáveis, projetando as informações que possuem maior relevância.

Os componentes principais possuem propriedades importantes, que são de grande interesse em certos estudos de melhoramento. Cada componente principal é uma combinação linear das variáveis originais, são ortogonais entre si e estimados com o propósito de reter em ordem de estimação o máximo da informação em termos de variação total, contida nos dados iniciais (CRUZ; REGAZZI, 1994).

A técnica dos componentes principais baseia-se apenas nas informações individuais de cada acesso, sem a necessidade de dados com repetições (CRUZ; CARNEIRO, 2003).

2.6.2 Análise de agrupamento por método não hierárquico

Proposto por J. MacQueen em 1967, o algoritmo de análise de agrupamento por método não hierárquico processado com o algoritmo k-médias inicia com a escolha da quantidade k de grupos a ser processada. O processo inicia-se por um dos métodos:

- Selecionando as p primeiras observações;
- Selecionando p observações aleatoriamente;
- Escolhendo p observações de modo que seus valores sejam bastante diferentes.

Escolhidas as sementes iniciais, é calculada a distância de cada elemento em relação às sementes, agrupando o elemento ao grupo que possuir a menor distância

(mais similar) e recalculando o centroide do mesmo. O processo é repetido até que todos os elementos façam parte de um dos grupos.

Após agrupar todos os elementos, procura-se encontrar uma partição melhor do que a gerada arbitrariamente. Para isto, calcula-se o grau de homogeneidade interna dos grupos utilizando-se da soma de Quadrados Residual (SQRes), que é a medida usada para avaliar o quão boa é uma partição.

Após o cálculo, move-se o primeiro objeto para os demais grupos e verifica-se se existe ganho na soma de quadrados residual, ou seja, se ocorre uma diminuição no valor da SQRes. Existindo, o objeto é movido para o grupo que produzir o maior ganho, a SQRes dos grupos é recalculada e passa-se ao objeto seguinte. Depois de um certo número de interações ou não havendo mais mudanças, o processo é interrompido.

2.6.3 Análise discriminante linear

A análise discriminante linear tem por finalidade classificar um objeto em k grupos, onde k tem que ser maior ou igual a dois grupos, com base em um vetor de observações multivariada x desse objeto. O objetivo é determinar a qual dos k grupos ou populações, cujas densidades de probabilidades são conhecidas, irá pertencer um novo objeto ou um novo conjunto de objetos (FERREIRA, 2011).

Conforme descrito, este trabalho demonstra a importância do complexo soja para o Brasil e o mundo, a dificuldade dos programas de melhoramento em selecionar genótipos superiores devido à grande complexidade das relações entre os caracteres selecionados.

O presente trabalho utilizou-se de métodos tradicionais e de técnicas de análise multivariadas descritas para a seleção de linhagens superiores.

3. REFERÊNCIAS

ALLARD, R. W. **Princípios de melhoramento genético das plantas**. São Paulo: Edgard Blücher, 1971. 381p.

ALLIPRANDINI, L. F.; ABATTI, C.; BERTAGNOLLI, P. F.; CAVASSIM, J. E.; GABE, H. L.; KUREK, A.; STECKLING, C. Understanding Soybean Maturity Groups in Brazil: Environment, Cultivar Classification, and Stability. **Crop Science**, Madison, v. 49, n. 3, p.801-808, maio/jun. 2009. Disponível em: <<https://www.crops.org/publications/cs/pdfs/49/3/801>>. Acesso em: 25 mar. 2014

ALMEIDA, R. D.; PELUZIO, J. M.; AFFÉRRI, F. S. Divergência genética entre cultivares de soja, sob condições de várzea irrigada, no sul do Estado Tocantins. **Revista Ciência Agronômica**, v. 42, n. 1, p. 108-115, 2011.

ALMEIDA, L. A.; SOUZA KIIHL, R. A.; MIRANDA, M. A. C.; AZEVEDO CAMPELO, G. J. **Melhoramento da soja para regiões de baixas latitudes**. 1999. Disponível em: <<http://www.cpatsa.embrapa.br/catalogo/livrorg/sojamelhoramento.pdf>>. Acesso em: 24 mar. 2014.

BONATO, E.R.; BONATO, A.L.V. **A soja no Brasil: História e Estatística**. Londrina, EMBRAPA-CNPSO, 1987. 61p. (EMBRAPA-CNPSO. Documentos, 21). Disponível em: < <http://www.infoteca.cnptia.embrapa.br/bitstream/doc/446431/1/Doc21.pdf> >. Acesso em: 03 jun. 2014.

CONAB (Companhia Nacional de Abastecimento). **Acompanhamento da Safra Brasileira de Grãos**. 2014. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/14_03_12_08_41_24_boletim_graos_marco_2014.pdf>. Acesso em: 24 mar. 2014.

COSTA, N.C. **Soja-Comercialização Safra 2011/12 e Cenário 2012/13**. 2012. Disponível em: <http://www.agricultura.gov.br/arq_editor/file/camaras_setoriais/Soja/16RO/App_Cenario_Soja.pdf>. Acesso em: 24 mar. 2014.

CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos Biométricos Aplicados ao Melhoramento Genético: Volume 2**. Viçosa: UFV, 2003. 585 p.

CRUZ, C.D.; REGAZZI, A.J. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, Imprensa Universitária da UFV, 1994. 390p.

CRUZ, C.D. Componentes da Variância Genotípica. In: CRUZ, C.D. **Princípios de Genética Quantitativa**. Viçosa: UFV, 2012. Cap. 3. p. 109-150.

EMBRAPA SOJA. **Tecnologias de Produção de Soja Região Central do Brasil**. 2004. Disponível em: <<http://www.cnpso.embrapa.br/producaosoja/SojanoBrasil.htm>>. Acesso em: 23 mar. 2014.

FALCONER, D. S.; MACKAY, T. F. C. **Introduction to quantitative genetics**. 4th ed. New York: Longman, 1996. 464p.

FEASP, Federação da Agricultura e Pecuária do Estado de São Paulo. **Soja**. 2011. Disponível em: <<http://www.faespsenar.com.br/faesp/pagina/exibe/faesp/produtos/soja/596>>. Acesso em: 05 ago. 2014.

FEHR, W. P.; CAVINESS, C. E. **Stages of soybean development**. Agriculture and Home Economics Experiment Station and Cooperative Extension Service. Ames: Iowa State University, 1977. 11p. (Special Report 80)

FERES, J.; GOMES, J.E.S. 1981. Melhoramento de Cultivares no Brasil – No Estado do Rio Grande do Sul. In: MIYASAKA, S.; MEDINA, J.C.(eds). **A soja no Brasil**. Campinas: Instituto de Tecnologia de Alimentos. p.279-282.

FERREIRA, D.F. **Estatística Multivariada**. 2. ed. Lavras: Ufla, 2011. 676 p.

HAIR, J.F.; BLACK, W.C.; BABIN, B.J.; ANDERSON, R.E.; TATHAM, R.L. **Análise Multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009. 688 p.

IQBAL, S.; MAHMOOD, T.; TAHIRA, M. A.; ANWAR, M.; SARWAR, M. Path coefficient analysis in different genotypes of soybean (*Glycine max* (L.) Merrill). **Pakistan Journal of Biological Science**, v. 12, p. 1085-1087, 2003.

MACQUEEN, J.B. Some methods for classification and analysis of multivariate observations. In Proc 5th Berkeley symposium on mathematical statistics and probability, Univ of California Press, Berkeley, 1967.

MAGALHÃES, C.M. 1981. Introdução e evolução da soja no Brasil – No estado do Rio Grande do Sul. In: MIYASAKA, S.; MEDINA, J.C. (eds.). **A soja no Brasil**. Campinas: Instituto de Tecnologia de Alimentos. P.18-20.

MIYASAKA, S. **Instruções para a cultura da soja**. Campinas, Instituto Agrônomo, 1965. 27p. (Boletim, 12).

NAVARRO JÚNIOR, H.; M.; COSTA, J.; A. Expressão do potencial de rendimento de cultivares de soja. **Pesquisa Agropecuária Brasileira**, v.37, p.275-279, 2002.

RAMALHO, M.A.P.; FERREIRA, D.F.; OLIVEIRA, A.C. Covariância e seu Emprego na Genética e na Experimentação Agrícola. In: RAMALHO, M.A.P.; FERREIRA, D.F.; OLIVEIRA, A.C. **Experimentação em genética e melhoramento de plantas**. 3. ed. Lavras: Ufla, 2012a. Cap. 13. p. 239-263.

RAMALHO, M.A.P.; BARBOSA-ABREU, A.F.; SANTOS, J.B.; NUNES, J.A.R. Conceito de herdabilidade e sua estimativa. In: RAMALHO, M.; BARBOSA-ABREU, A.F.; SANTOS, J.B.; NUNES, J.A.R. **Aplicações da genética quantitativa no melhoramento de plantas autógamas**. Lavras: UFLA, 2012b. Cap. 7. 123-148.

RAMALHO, M.A.P.; SANTOS, J.B.; PINTO, C.A.B.P.; SOUZA, E.A.; AVELAR-GONÇALVES, F.M.; SOUZA, J.C. Genética Quantitativa. In: RAMALHO, M.A.P.; SANTOS, J.B.; PINTO, C.A.B.P.; SOUZA, E.A.; AVELAR-GONÇALVES, F.M.; SOUZA, J.C. **Genética na agropecuária**. 5 ed. Lavras: UFLA, 2012. Cap. 12. 271-312.

ROESSING, A.C.; GUEDES, L.C.A. 1993. Aspectos Econômicos do complexo soja: sua participação na economia brasileira e evolução na região do Brasil Central. In: ARANTES, N.E.; SOUZA, P.I.M. (eds). **Cultura da soja nos cerrados**. Piracicaba: Associação Brasileira de Pesquisa da Potassa e do Fosfato. P.01-69.

SOJA, Sindmilho. **Soja e suas riquezas – História**. 2011. Disponível em: <<http://www.fiesp.com.br/sindmilho/sobre-o-sindmilho/curiosidades/soja-e-suas-riquezas-historia/>>. Acesso em: 23 mar. 2014.

SEDIYAMA, T.; TEIXEIRA, R.C.; REIS, M.S.. Melhoramento da Soja. In: BORÉM, A. (Ed.). **Melhoramento de Espécies Cultivadas**. 2. ed. Viçosa: UFV, 2005. Cap. 14. p. 562-572.

SEDIYAMA, T.; TEIXEIRA, R.C.; BARROS, H.B. Origem, Evolução e Importância Econômica. In: SEDIYAMA, T. (Ed.). **Tecnologias de produção e usos da soja**. Londrina: Mecenias, 2009. Cap. 1. p. 1-5.

VIANNA, V.F.; UNÊDA-TREVISOLI, S.H.; DESIDÉRIO, J.A.; SANTIAGO, S.; CHARNAI, K.; FERREIRA JUNIOR, J.A.; FERRAUDO, A.S.; MAURO, A.O. The multivariate approach and influence of characters in selecting superior soybean genotypes. **Afr. J. Agric. Res**, v. 8, p. 4162-4169, 2013.

WILCOX, J.R.; GOUDONG, Z. 1997. Relationships between seed yield and seed protein in determinate and indeterminate soybean populations. *Crop Sci.* 37:361–364. doi:10.2135/cropsci1997.0011183X003700020009x

CAPÍTULO 2 - ESTIMATIVAS DE PARÂMETROS GENÉTICOS E CORRELAÇÕES DE CARACTERÍSTICAS EM LINHAGENS AVANÇADAS DE SOJA

RESUMO – O presente trabalho teve como objetivo a estimação de parâmetros genéticos em linhagens avançadas de soja, identificação de caracteres correlacionados de forma direta e indireta com a produtividade de grãos, além da seleção de genótipos superiores de forma a maximizar o ganho para os caracteres avaliados. Foram conduzidos dois experimentos no ano agrícola 2012/2013 no delineamento de blocos ao acaso com três repetições. Foram semeadas 23 linhagens no primeiro experimento e 46 linhagens no segundo experimento, ambas pertencentes ao programa de melhoramento genético de soja da UNESP, Câmpus de Jaboticabal. Para todas as características analisadas nos dois experimentos, houve significância do teste F a 5% de probabilidade. A razão entre o coeficiente de variação genético e o coeficiente de variação ambiental ficou acima de 1 para altura de inserção da primeira vagem, altura da planta na maturidade, número de dias para maturidade, valor agrônômico, número de ramos, número de vagens, peso de cem sementes, teor de óleo e produção de grãos, para o e primeiro experimento. Para o segundo experimento apenas os caracteres altura da planta na maturidade e peso de cem sementes apresentaram a razão acima de 1, indicando a possibilidade de ganhos com a seleção nos respectivos caracteres. O índice de Mulamba e Mock detectou linhagens superiores, destacando as linhagens 10,15,17 e 20 para o experimento I e as linhagens 32, 41,48, e 57 para o experimento II.

Palavras-chave: Mulamba e Mock. Análise de trilha. *Glycine max*. Melhoramento genético. Índice de seleção.

1. INTRODUÇÃO

O melhoramento genético vem contribuindo significativamente para o sucesso da cultura da soja no país, o que tem possibilitado o desenvolvimento de cultivares cada vez mais adaptadas e produtivas às mais variadas condições edafoclimáticas (SEDIYAMA et al., 2005).

A obtenção de estimativas de parâmetros genéticos, tais como herdabilidade, ganho esperado com a seleção e correlações genéticas e fenotípicas, tem grande importância em programas de melhoramento genético, pois possibilita a tomada de decisões relacionadas com a escolha do método de melhoramento genético mais apropriado e permite o monitoramento dos demais caracteres durante o processo seletivo (ROSSMANN, 2001).

A produtividade é um caráter complexo e resultante da expressão e associação de diferentes fatores. O conhecimento do grau dessa associação, por meio de estudos de correlações, possibilita identificar caracteres que podem ser usados em processos de seleção indireta visando ganhos de produtividade (CARVALHO et al., 2002).

Embora seja de grande importância o conhecimento das correlações, alguns cuidados devem ser considerados, pois elas podem não ser medidas de causa e efeito e a interpretação direta das suas magnitudes pode resultar em equívocos na estratégia de seleção, visto que a correlação alta entre dois caracteres pode ser resultado do efeito indireto de um terceiro caráter ou de um grupo de caracteres (NOGUEIRA et al., 2012).

Para solucionar este problema, Wright (1921) desenvolveu a análise de trilha ou "path analysis", que consiste no estudo dos efeitos diretos e indiretos de caracteres sobre uma variável básica, cujas estimativas são obtidas por meio de equações de regressão, em que as variáveis são previamente padronizadas (CRUZ; REGAZZI, 2001).

Objetivou-se neste trabalho estimar parâmetros genéticos em linhagens avançadas de soja, identificar caracteres direta e indiretamente correlacionados com a produtividade de grãos e selecionar genótipos superiores de forma a maximizar o ganho para o conjunto de caracteres avaliados.

2. MATERIAL E MÉTODOS

2.1 Manejo experimental e tratamentos utilizados

Foram conduzidos dois experimentos no ano agrícola 2012/2013 na Fazenda de Ensino, Pesquisa e Extensão (FEPE), da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Jaboticabal, localizada ao norte do Estado de São Paulo, a 21°15' de latitude sul e 48°18' de longitude oeste, com altitude aproximada de 595 m. A divisão em dois experimentos se fez necessária pois 23 linhagens estavam na geração F₇-F₈ (I) e 46 estavam na geração F₆-F₇ (II).

O solo da área experimental é classificado como Latossolo Vermelho Eutroférico Argiloso, com relevo suavemente ondulado e o clima é do tipo Cwa, segundo a classificação de Köppen.

A semeadura foi realizada mecanicamente no dia 27 de novembro de 2012, para o primeiro experimento e no dia 29 de novembro de 2012 para o segundo experimento, ambos foram desbastados e o estande final ficou com densidade de 15 plantas m⁻¹. No plantio foram utilizados 350 kg ha⁻¹ da fórmula 00-30-15, sendo que as parcelas experimentais foram mantidas durante todo o ciclo da cultura, com rigoroso controle de pragas, doenças e plantas infestantes, conforme recomendações para a cultura da soja (EMBRAPA, 2011).

Tabela 1. Linhagens pertencentes ao Programa de melhoramento genético de soja da Unesp-FCAV. Ano agrícola 2012/2013, Jaboticabal-SP

Experimento I		Experimento II	
Linhagens		Linhagens	
1		24	47
2		25	48
3		26	49
4		27	50
5		28	51
6		29	52
7		30	53
8		31	54
9		32	55
10		33	56
11		34	57
12		35	58
13		36	59
14		37	60
15		38	61
16		39	62
17		40	63
18		41	64
19		42	65
20		43	66
21		44	67
22		45	68
23		46	69

Linhagens proveniente do cruzamento PI 200456 x MGBR46 (Conquista)

Os experimentos foram instalados no delineamento de blocos ao acaso com três repetições. Cada parcela experimental consistiu de quatro linhas de 5 m de comprimento, espaçadas de 0,5 m entre si, sendo considerada como área útil as duas linhas centrais, desprezando-se 0,5 m de cada extremidade, totalizando 4 m². Foram semeadas 23 linhagens no primeiro experimento e 46 linhagens no segundo experimento (Tabela 1), ambas pertencentes ao programa de melhoramento genético de soja da UNESP, Câmpus de Jaboticabal.

2.2 Caracteres agronômicos

A caracterização fenotípica foi realizada por meio da avaliação dos seguintes caracteres agromorfológicos: Altura de inserção da primeira vagem (AIV): obtido pela medida, do colo da planta até a altura da inserção da primeira vagem (cm), considerando a média de seis plantas dentro da área útil. Altura de planta na maturidade (APM): obtido pela medida, da altura do colo da planta até o ápice da haste principal (cm), considerando a média de seis plantas tomadas ao acaso dentro da área útil. Número de dias para maturidade (NDM): definido como o período em dias entre a data de emergência das plantas até a data em que aproximadamente 95% das vagens apresentaram-se maduras (estádio R8) (FEHR & CAVINESS, 1977). Acamamento (AC): caráter avaliado no estágio R8 (FEHR & CAVINESS, 1977) da planta por meio de uma escala de notas visuais, variando de 1 (todas as plantas eretas) a 5 (todas as plantas acamadas); valor agronômico (VA): caráter avaliado no estágio R8 de desenvolvimento da planta, por meio de uma escala de notas visuais, a qual varia de 1 (plantas com características agronômicas ruins) a 5 (plantas com ótimas características agronômicas), sendo a nota atribuída representativa de um conjunto de caracteres visuais (arquitetura da planta, quantidade de vagens cheias, vigor e sanidade da planta, debulha prematura das vagens, acamamento e retenção foliar na maturidade); número de ramos (NR): obtido por meio da contagem de ramos de cada planta, sendo tomado o valor médio de seis plantas dentro da área útil; número de vagens (NV): obtido por meio da contagem da quantidade de vagens de cada planta sendo tomado o valor médio de seis plantas dentro da área útil; peso de cem sementes (PCS): obtido por meio de três amostras de 100 sementes, utilizando uma balança de precisão de um grama; teor de óleo (TO): porcentagem de óleo contida nos grãos de soja, sendo obtida com o equipamento NIR Bruker modelo Tango, que realiza a medição dos teores de óleo sem que ocorra a destruição das sementes; produtividade de grãos (PG): obtida a partir da colheita das plantas da área útil da parcela, sendo as mesmas trilhadas e o peso dos grãos corrigido para 13% de umidade, posteriormente o peso corrigido é convertido em kg. ha⁻¹.

2.3 Análises estatísticas

Os dados foram analisados no software R (R Core Team 2013) para exclusão dos “outliers” ou valores discrepantes. Para a transformação, foi utilizada a metodologia de Box e Cox (1964) quando necessário.

A seguir os dados foram submetidos à análise de variância com o uso do programa SAS (Statistical Analysis System, 2010), segundo o modelo de delineamento em blocos ao acaso (DBC) conforme descrito por Cruz e Regazzi (2001).

$$Y_{ij} = \mu + \beta_i + \tau_i + \varepsilon_{ij}$$

Onde:

Y_{ij} = observação do tratamento i , na repetição j ;

μ = média geral do caráter;

β_i = efeito do j – ésimo bloco;

τ_i = efeito do i – ésimo tratamento;

ε_{ij} = efeito do erro aleatório.

A obtenção dos componentes de variância também foi feita com o auxílio do programa SAS (Statistical Analysis System, 2010) pelo método da máxima verossimilhança restrita, adotando o efeito de blocos como fixo e de genótipos como aleatório.

Foram obtidas as matrizes de correlação genética, covariâncias fenotípicas e genotípicas com o uso do software Selegen (Resende, 2008). A matriz de correlação genética e as matrizes de covariância fenotípica e genotípica foram empregadas no programa GENES (Cruz, 2007) para a realização da análise de trilha, diagnóstico de multicolinearidade e aplicação do índice de seleção baseado na soma de “ranks” de Mulamba e Mock (1978).

O desdobramento das correlações em efeitos diretos e indiretos foram realizados por meio da análise de trilha proposta por Wright (1921), tomando os devidos cuidados para que o grau de multicolinearidade não afetasse a análise.

Assim, foi calculado o grau de multicolinearidade da matriz $X'X$ estabelecido de acordo com Montgomery e Peck (1981), baseando nos valores do determinante e do número de condição ($NC = \text{razão entre o maior e o menor autovalor das matrizes}$). Foi efetuada a análise dos elementos dos autovetores associados aos autovalores com o intuito de identificar os caracteres que mais contribuíram para a multicolinearidade, conforme descrito por Belsley et al. (1980).

Foi aplicado o índice de seleção baseado em soma de “ranks” proposto por Mulamba e Mock (1978).

$$I = p_1r_1 + p_2r_2 + \dots + p_n r_n$$

Onde:

I = valor do índice para determinado indivíduo ou família;

r_n = classificação (ou “rank”) de um indivíduo em relação ao n -ésimo caráter;

n = número de caracteres considerado no índice;

p = peso econômico atribuído pelo usuário para cada caráter (n).

A pressão de seleção foi de 30% e realizada com base nos caracteres produção de grãos (peso econômico 2), valor agrônômico e teor de óleo (peso econômico 1).

3. RESULTADOS E DISCUSSÃO

Para todas as características analisadas nos dois experimentos, houve significância do teste F a 5% de probabilidade. Os coeficientes de variação experimental oscilaram de 2,18% (TO) (Tabela 2, experimento I) a 32,09% (PG) (Tabela 3, experimento II).

A razão entre o coeficiente de variação genético (C_{vg}) e o coeficiente de variação ambiental (C_{ve}) ficou acima de 1 para AIV, APM, NDM, VA, NR, NV, PCS, TO e PG (Tabela 2, experimento I) e APM, PCS (Tabela 3, experimento II), indicando situação favorável na obtenção de ganhos com a seleção baseada nos respectivos caracteres. O coeficiente de herdabilidade foi de 22,11% para AC (Tabela 2, experimento I), valor baixo indicando poucos ganhos para o caráter que não apresenta

situação favorável para a seleção visto que a razão (Cvg/Cve) ficou abaixo de 1. Tal fato não acarretará em grandes problemas visto que a média do caráter na população é 1,25, bem próximo do ideal que é 1. O mesmo é válido para o caráter AC no experimento II (Tabela 3, experimento II).

O coeficiente de herdabilidade para os caracteres NR, NV, PCS e PG no experimento I (Tabela 2, experimento I) apresentaram valores médios à moderados, e estão em situação favorável para seleção, sendo possível a obtenção de ganhos significativos nestes caracteres.

Para o experimento II o coeficiente de herdabilidade apresentou valores baixos para oito caracteres, com exceção dos caracteres APM que apresentou um valor moderado e PCS que apresentou um valor médio, os dois também foram os únicos que apresentaram situação favorável para seleção. Na situação do experimento II, os ganhos com a seleção tendem a ser baixos, visto que oito dos dez caracteres avaliados apresentaram baixos valores para o coeficiente de herdabilidade e uma situação desfavorável para a seleção.

Silveira et al. (2006) obtiveram estimativas de herdabilidade de 59% para AIV e 78% para APM possibilitando a obtenção de estimativas de ganho com a seleção significativa para os dois caracteres. Muniz et al. (2007) encontraram valores de herdabilidade variando 0 a 53% para APM, e 0 a 19% para PG. Valores bem próximos foram encontrados neste trabalho, o autor obteve estimativas de ganho com a seleção baixas para os caracteres, variando de 0 a 6,87% para APM e 0,57 a 5,76% para PG, indicando que a situação não está favorável à seleção para todos os cruzamentos observando os dois caracteres. Oliveira et al. (2000) encontraram valores de herdabilidade para NV de 11% e 57%. Tais valores diferem dos observados nesse trabalho. As estimativas de herdabilidade são intrínsecas a população em estudo, podendo variar muito ou quase nada de uma população para a outra, além disso o parâmetro pode variar muito dentro da mesma população de acordo com o método utilizado para sua estimação, neste trabalho foi utilizada a metodologia da máxima verossimilhança restrita, que é mais indicada quando se tem dados desbalanceados.

Tabela 2. Quadrados médios e parâmetros genéticos obtidos na análise de variância para 10 caracteres em 23 linhagens F7-F8 de soja. Experimento I, safra 2012/2013, UNESP – Jaboticabal

FV	GL	AIV (cm) ¹	APM (cm)	NDM	AC ²	VA ³
BLOCOS	2	0,44**	132,46**	46,36	0,07	0,00
GENÓTIPOS	22	0,60**	321,71**	80,68**	0,05**	0,0041**
RESÍDUOS	41	0,11	32,50	16,65	0,03	0,00
CVg/CVe		1,22	1,74	1,13	0,53	1,17
S ² genética		0,17	98,29	21,34	0,01	0,00
S ² ambiental		0,11	32,46	16,65	0,03	0,00
Herdabilidade (%)		59,80	75,17	56,17	22,11	57,65
MÉDIA***		9,18	77,99	111,29	1,25	2,83
CV(%)		11,08	7,31	3,67	21,2	4,38
FV	GL	NR ⁴	NV	PCS (g)	TO (%)	PG (kg,ha ⁻¹)
BLOCOS	2	0,01	1009,10	1,00	0,18	41574,22
GENÓTIPOS	22	0,22**	2603,65**	4,06**	1,84**	2110400,40**
RESÍDUOS	41	0,06	467,96	0,78	0,20	352087,77
CVg/CVe		1,04	1,31	1,40	1,87	1,43
S ² genética		0,06	794,27	1,51	0,70	693248,60
S ² ambiental		0,06	466,13	0,77	0,20	338253,80
Herdabilidade (%)		51,87	63,02	66,20	77,74	67,21
MÉDIA***		3,34	107,43	17,34	20,62	3707,64
CV(%)		12,22	20,14	5,11	2,18	16,00

Significativos a 5% de probabilidade pelo teste F. As transformações foram feitas de acordo com a necessidade pela metodologia de Box e Cox (1964). ¹ Transformação $Y^{0,5}$; ² Transformação Y^{-1} ; ³ Transformação $Y^{-0,4}$; ⁴ Transformação $\sqrt{X + 0,5}$; * Valores apresentados sem a transformação dos dados. AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

Tabela 3. Quadrados médios e parâmetros genéticos obtidos na análise de variância para 10 caracteres em 46 linhagens F₆-F₇ de soja. Experimento II, safra 2012/2013, UNESP – Jaboticabal

FV	GL	AIV (cm) ¹	APM (cm)	NDM	AC ¹	VA ¹
BLOCOS	2	0,01	12,29	132,64**	0,01	0,06**
GENÓTIPOS	43	0,42**	405,91**	78,98**	0,03**	0,04**
RESÍDUOS	73	0,19	58,41	38,52	0,01	0,02
CVg/CVe		0,66	1,48	0,59	0,73	0,72
S ² genética		0,08	128,05	13,49	0,01	0,01
S ² ambiental		0,19	58,36	38,52	0,01	0,02
Herdabilidade (%)		30,42	68,69	25,93	34,80	33,87
MÉDIA***		7,34	70,89	111,36	1,28	2,59
CV(%)		16,27	10,78	5,57	10,22	8,44
FV	GL	NR ¹	NV	PCS (g)	TO (%)	PG (kg,ha ⁻¹)
BLOCOS	2	0,19	668,28	0,97	2,12	1276843,00
GENÓTIPOS	43	0,20**	1284,48**	4,27**	2,25**	1910844,91**
RESÍDUOS	73	0,08	726,99	1,20	0,91	1029143,10
CVg/CVe		0,75	0,53	1,04	0,79	0,50
S ² genética		0,04	201,22	1,29	0,56	287937,20
S ² ambiental		0,08	727,37	1,20	0,91	1160137,00
Herdabilidade (%)		36,00	21,67	51,83	38,24	19,88
MÉDIA***		3,17	97,64	17,08	20,16	3161,19
CV(%)		15,66	27,61	6,42	4,73	32,09

Significativos a 5% de probabilidade pelo teste F. As transformações foram feitas de acordo com a necessidade pela metodologia de Box e Cox (1964). ¹Transformação Y^{0,5}. * Valores apresentados sem a transformação dos dados. AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônomo; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

As estimativas de correlações genéticas, para os caracteres avaliados nos experimentos I e II estão apresentadas na Tabela 4.

Verifica-se que as estimativas das correlações genéticas entre PG x AIV, PG x APM, PG x NDM, PG x TO para o Experimento I (Tabela 4) apresentam significativo grau de associação e encontram-se no mesmo sentido. Isto indica que selecionando maiores valores para esses caracteres poderá ocorrer um ganho indireto em produtividade. O carácter VA apresentou baixa magnitude de correlação genotípica com PG em mesmo sentido, apresentou também correlações variando de média a alta magnitude para outros caracteres de interesse agrônomo a saber: VA x AIV (correlação média e mesmo sentido), VA x APM (alta correlação e mesmo sentido),

NR x VA (média correlação e sentido oposto), NV x VA (média correlação e sentido oposto) (Experimento I).

No Experimento II as correlações PG x AIV, PG x APM, PG x NDM foram de baixas magnitudes e mesmo sentido, a correlação PG x AC foi de baixa magnitude e em sentido oposto, significando que ganhos no caráter produtividade de grãos poderá implicar numa redução do caráter acamamento. Os maiores coeficientes de correlação para o segundo experimento foram entre as variáveis APM x VA, NR x NV, AIV x VA ambos no mesmo sentido.

Muniz et al. (2002) obtiveram correlação de baixa magnitude e em sentido oposto para PG x APM, onde tal valor difere do obtido nesse trabalho que foram correlações baixas e médias, ambas no mesmo sentido para os dois experimentos (Tabela 4). A Correlação entre PG x APM em sentido oposto pode ser interessante quando na população a média do caráter APM é alta (próxima a 100 cm), porque selecionando plantas mais produtivas, também selecionaria plantas mais baixas, facilitando a colheita mecanizada e buscando um estande mais ereto visto que em muitos casos existe correlação no mesmo sentido entre APM x AC.

Neto e Vello (2001) encontraram correlações genótípicas de PG x APM de alta magnitude e de mesmo sentido. Neste estudo, as correlações encontradas para PG x APM foram médias e baixas, ambas de mesmo sentido (Tabela 4). O autor também encontrou correlações genéticas de 60% para PG x VA no mesmo sentido.

Tabela 4. Estimativas dos coeficientes de correlação genotípica Experimento I (acima da diagonal) e Experimento II (abaixo da diagonal) entre 10 caracteres avaliados em linhagens avançadas de soja, safra 2012/2013, UNESP – Jaboticabal

	AIV (cm)	APM (cm)	NDM	AC	VA	NR	NV	PCS (g)	TO (%)	PG (kg.ha ⁻¹)
AIV	1	0,78	0,37	-0,03	0,47	-0,30	-0,29	0,03	-0,07	0,40
APM	0,33	1	0,40	-0,01	0,60	-0,39	-0,34	0,04	0,17	0,58
NDM	0,00	0,35	1	0,12	-0,02	-0,07	0,09	0,57	0,27	0,57
AC	-0,17	0,22	0,29	1	-0,59	0,44	0,28	0,40	0,09	-0,02
VA	0,49	0,66	-0,04	-0,34	1	-0,51	-0,45	-0,29	-0,04	0,24
NR	-0,05	-0,11	-0,19	0,38	-0,27	1	0,90	0,41	0,10	-0,05
NV	-0,41	-0,08	0,04	0,48	-0,35	0,55	1	0,51	0,24	0,12
PCS	-0,20	-0,16	0,07	0,25	-0,33	0,21	0,04	1	0,19	0,21
TO	-0,30	-0,40	-0,07	-0,03	-0,36	0,32	0,11	0,11	1	0,44
PG	0,18	0,37	0,14	-0,24	0,37	-0,35	-0,39	-0,08	-0,07	1

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

Com o intuito de identificar e quantificar os efeitos diretos e indiretos que outros caracteres apresentam sobre a variável produção de grãos, foi realizado um desdobramento dos coeficientes de correlação genotípico pela análise de trilha com uma cadeia proposta por Wright (1921). São apresentados, na Tabela 5, os efeitos diretos e indiretos das variáveis explicativas sobre a produção de grãos para os 23 genótipos de soja (Experimento I) e para os 46 genótipos de soja (Experimento II).

O coeficiente de determinação revela que 63,03% da produção de grãos é explicada pelo efeito das variáveis analisadas com base na matriz de correlação genotípica, sucessivamente para o segundo experimento temos 33,10% (Tabela 5).

Para o experimento I a variável AIV apresentou efeito direto sobre PG em sentido oposto (-0,07), os efeitos indiretos mais pronunciados foram APM e NDM e NV. A correlação total entre AIV e PG foi positiva e bem superior ao efeito direto, indicando neste caso que a causa da correlação se deve aos efeitos indiretos exercidos pelas demais variáveis. A variável APM para o experimento I apresentou correlação total com PG de 0,58, o efeito direto exercido de APM sobre PG foi igual à correlação total, indicando neste caso que a correlação total explica a verdadeira relação entre as variáveis e a seleção direta sobre APM será efetiva (Tabela 5).

A variável NDM para o experimento I, apresentou correlação total de 0,57 e efeito direto sobre PG de 0,41, como o efeito direto de NDM sobre PG é alto e próximo

da correlação total, podemos dizer que a correlação explica a verdadeira relação entre as variáveis e neste caso a seleção direta sobre NDM também será efetiva (Tabela 5).

Para o caráter VA a correlação total foi de 0,24 e o efeito direto sobre PG foi de -0,05 indicando que neste caso o efeito indireto é a causa da correlação, isso já poderia ser esperado visto que o caráter VA é uma nota visual atribuída pelo pesquisador levando em consideração vários outros caracteres (Tabela 5).

Rigon et al. (2012) encontraram correlação total e efeitos diretos negativos sobre a produção de grãos para NDM, AC, APM, onde a correlação explicou a verdadeira relação entre as variáveis e a seleção direta sobre essa característica poderá ser efetiva para ganhos indiretos em produtividade, os autores encontraram efeito direto positivo somente para o caráter PCS.

Para o experimento II a variável APM apresentou correlação total de 0,37 e efeito direto sobre PG de 0,48, indicando que a correlação explica a verdadeira relação e a seleção direta sobre a característica é efetiva. Nenhuma variável apresentou efeito indireto pronunciado (Tabela 5). Perini et al. (2012) encontraram para cultivares de hábito indeterminado correlação total do caráter massa de mil grãos com a produtividade de grãos de -0,785 e efeito direto sobre a produtividade de grãos de -0,020 indicando que neste caso a correlação não explica a verdadeira relação e a seleção direta sobre a característica não é efetiva, o que pode ser confirmado pelo efeito indireto de número de grãos por planta de -0,764. Para o caráter número de grãos por planta os autores encontraram correlações totais bem próximas do efeito direto sobre a produtividade de grãos indicando a possibilidade de se utilizar esse caráter para ganhos indiretos na produtividade de grãos.

Observando o caráter VA para o segundo experimento (Tabela 5), pode se chegar à conclusão de que a correlação existente entre VA x PG é devida ao efeito indireto da variável APM. O caráter AC apresentou correlação total com PG bem próximo do efeito direto sobre PG indicando que a correlação existente entre os dois caracteres explica a verdadeira relação e a seleção direta sobre a característica é efetiva.

O caráter APM pode ser utilizado em índices de seleção no sentido superior visando a obtenção de ganhos indiretos para a produtividade de grãos e o caráter AC também porém no sentido inferior.

Analisando os resultados dos dois experimentos (Tabela 5), pode-se verificar que as correlações, bem como seus efeitos diretos e indiretos são intrínsecas à população em estudo e que não podem ser extrapoladas para outras populações, sendo assim faz-se necessário o estudo em cada população para que se possa adotar uma estratégia de seleção onde se consiga maximizar os ganhos de acordo com o objetivo do programa.

Tabela 5. Análise de Trilha: desdobramento das correlações genótípicas em componentes de efeito direto e indireto envolvendo a variável dependente principal (PG) e as variáveis independentes explicativas (AIV, APM, NDM, AC, VA, NR, NV, PCS, TO), avaliados no ano agrícola 2012/13 para 23 genótipos de soja (Experimento I) e para 46 genótipos de soja (Experimento II), UNESP – Jaboticabal

Componentes	Estimativa Experimento I	Estimativa Experimento II
AIV		
EFEITO DIRETO SOBRE PG	-0,07	-0,02
EFEITO INDIRETO VIA APM	0,45	0,16
EFEITO INDIRETO VIA NDM	0,15	0,00
EFEITO INDIRETO VIA AC	0,00	0,04
EFEITO INDIRETO VIA VA	-0,02	-0,02
EFEITO INDIRETO VIA NR	0,05	0,01
EFEITO INDIRETO VIA NV	-0,15	0,07
EFEITO INDIRETO VIA PCS	-0,01	-0,01
EFEITO INDIRETO VIA TO	-0,01	-0,05
Correlação (Total)	0,40	0,18
APM		
EFEITO DIRETO SOBRE PG	0,58	0,48
EFEITO INDIRETO VIA AIV	-0,06	-0,01
EFEITO INDIRETO VIA NDM	0,17	0,01
EFEITO INDIRETO VIA AC	0,00	-0,05
EFEITO INDIRETO VIA VA	-0,03	-0,02
EFEITO INDIRETO VIA NR	0,07	0,02
EFEITO INDIRETO VIA NV	-0,17	0,02
EFEITO INDIRETO VIA PCS	-0,01	-0,01
EFEITO INDIRETO VIA TO	0,03	-0,07
Correlação (Total)	0,58	0,37
NDM		
EFEITO DIRETO SOBRE PG	0,41	0,01

	EFEITO INDIRETO VIA AIV	-0,03	0,00
	EFEITO INDIRETO VIA APM	0,23	0,17
	EFEITO INDIRETO VIA AC	-0,01	-0,06
	EFEITO INDIRETO VIA VA	0,00	0,00
	EFEITO INDIRETO VIA NR	0,01	0,04
	EFEITO INDIRETO VIA NV	0,04	-0,01
	EFEITO INDIRETO VIA PCS	-0,14	0,00
	EFEITO INDIRETO VIA TO	0,05	-0,01
	Correlação (Total)	0,57	0,14
AC	EFEITO DIRETO SOBRE PG	-0,08	-0,21
	EFEITO INDIRETO VIA AIV	0,00	0,00
	EFEITO INDIRETO VIA APM	0,00	0,10
	EFEITO INDIRETO VIA NDM	0,05	0,00
	EFEITO INDIRETO VIA VA	0,03	0,01
	EFEITO INDIRETO VIA NR	-0,08	-0,07
	EFEITO INDIRETO VIA NV	0,14	-0,09
	EFEITO INDIRETO VIA PCS	-0,10	0,02
	EFEITO INDIRETO VIA TO	0,02	0,00
	Correlação (Total)	-0,02	-0,24
VA	EFEITO DIRETO SOBRE PG	-0,05	-0,04
	EFEITO INDIRETO VIA AIV	-0,03	-0,01
	EFEITO INDIRETO VIA APM	0,35	0,31
	EFEITO INDIRETO VIA NDM	-0,01	0,00
	EFEITO INDIRETO VIA AC	0,04	0,07
	EFEITO INDIRETO VIA NR	0,09	0,05
	EFEITO INDIRETO VIA NV	-0,22	0,06
	EFEITO INDIRETO VIA PCS	0,07	-0,02
	EFEITO INDIRETO VIA TO	-0,01	-0,06
	Correlação (Total)	0,24	0,37
NR	EFEITO DIRETO SOBRE PG	-0,17	-0,19
	EFEITO INDIRETO VIA AIV	0,02	0,00
	EFEITO INDIRETO VIA APM	-0,23	-0,05
	EFEITO INDIRETO VIA NDM	-0,03	0,00
	EFEITO INDIRETO VIA AC	-0,03	-0,08
	EFEITO INDIRETO VIA VA	0,02	0,01
	EFEITO INDIRETO VIA NV	0,45	-0,10
	EFEITO INDIRETO VIA PCS	-0,10	0,01
	EFEITO INDIRETO VIA TO	0,02	0,05
	Correlação (Total)	-0,05	-0,35

	EFEITO DIRETO SOBRE PG	0,50	-0,18
	EFEITO INDIRETO VIA AIV	0,02	0,01
	EFEITO INDIRETO VIA APM	-0,20	-0,04
	EFEITO INDIRETO VIA NDM	0,04	0,00
NV	EFEITO INDIRETO VIA AC	-0,02	-0,10
	EFEITO INDIRETO VIA VA	0,02	0,01
	EFEITO INDIRETO VIA NR	-0,15	-0,10
	EFEITO INDIRETO VIA PCS	-0,12	0,00
	EFEITO INDIRETO VIA TO	0,04	0,02
Correlação (Total)		0,12	-0,39
	EFEITO DIRETO SOBRE PG	-0,24	0,06
	EFEITO INDIRETO VIA AIV	0,00	0,00
	EFEITO INDIRETO VIA APM	0,02	-0,08
	EFEITO INDIRETO VIA NDM	0,23	0,00
PCS	EFEITO INDIRETO VIA AC	-0,03	-0,05
	EFEITO INDIRETO VIA VA	0,01	0,01
	EFEITO INDIRETO VIA NR	-0,07	-0,04
	EFEITO INDIRETO VIA NV	0,25	-0,01
	EFEITO INDIRETO VIA TO	0,03	0,02
Correlação (Total)		0,21	-0,08
	EFEITO DIRETO SOBRE PG	0,17	0,17
	EFEITO INDIRETO VIA AIV	0,00	0,01
	EFEITO INDIRETO VIA APM	0,10	-0,19
	EFEITO INDIRETO VIA NDM	0,11	0,00
TO	EFEITO INDIRETO VIA AC	-0,01	0,01
	EFEITO INDIRETO VIA VA	0,00	0,01
	EFEITO INDIRETO VIA NR	-0,02	-0,06
	EFEITO INDIRETO VIA NV	0,12	-0,02
	EFEITO INDIRETO VIA PCS	-0,05	0,01
Correlação (Total)		0,44	-0,07
Coeficiente de Determinação (R ²)		0,6303	0,331
Efeito Residual (p _E)		0,6081	0,818
Determinante da Matriz		0,0008	0,0219
Colinearidade da Matriz		Fraca	Fraca
Número de Condição		64,7368	25,237

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

Com relação à análise de ganhos genéticos, o índice baseado na soma de “ranks” de Mulamba e Mock (1978), possibilitou um ganho total de 111,80% distribuído para os dez caracteres no experimento I. O maior ganho foi para PG (32,15%), seguido de TO e AIV com 19,40% e 15,06% respectivamente. O caráter NDM foi o que apresentou o menor ganho, de 0,65% (Tabela 6).

Analisando a Tabela 7, observamos que o índice teve boa aplicabilidade na seleção de linhagens superiores, visto que a seleção direta sobre PG também selecionou as mesmas linhagens. Conforme Cruz (1990), a utilização dos índices, mesmo considerando apenas um caráter como principal, mostra-se vantajosa em relação à seleção direta, uma vez que possibilita ganhos mais distribuídos em todos os caracteres avaliados, com ganhos totais maiores, sem proporcionar perda significativa no caráter principal.

As linhagens que se destacaram no caráter PG para o experimento I foram 10,15,17 e 20 (Tabela 7).

Para o Experimento II, foram observados ganhos totais inferiores (58,57%) quando comparado ao Experimento I. O ganho para o caráter PG foi muito inferior ao primeiro experimento. Os caracteres APM, PG e PCS apresentaram os maiores ganhos que foram de 15,46%, 9,86% e 9,34% respectivamente. O menor ganho foi constatado para o caráter NDM de 0,19%, sendo constatado uma perda de -1% e -0,91% para os caracteres NR e NV respectivamente, essa perda pode ter sido recompensada com o ganho obtido no caráter PCS que foi de 9,34% (Tabela 8).

Como no experimento I, o índice também teve boa aplicabilidade na seleção de linhagens superiores para o experimento II, diferindo da seleção direta sobre PG apenas nos 3 últimos indivíduos menos produtivos do rank. Destacam-se as linhagens 41,57,32 e 48 que apresentaram maiores valores para o caráter PG (experimento II).

Costa et al. (2004), observaram ganhos de 46,67% para produção de grãos usando o índice de Mulamba e Mock (1978), peso econômico 1 e considerando os caracteres PG, VA e APM como principais. O autor concluiu que o índice de Mulamba e Mock revelou-se mais adequado nas condições de seu trabalho, com progressos superiores em várias situações.

Bizari (2014), obteve resultado bastante favorável no processo seletivo em populações segregantes de soja, com ganho total de 39,17%, sendo que para o

caráter PG o ganho foi de 34,45%, utilizando o índice baseado na soma de “ranks”, tendo como caracteres principais VA e PG, adotando como peso econômico o coeficiente de variação genética (Cvg).

Tabela 6. Estimativas das médias da população original (Xo) e da população selecionada (Xs), herdabilidade (h^2 %) e ganho com a seleção (GS) em 23 linhagens F7-F8 de soja (Experimento I), safra 2012/2013, UNESP – Jaboticabal

Variáveis	Xo	Xs	h^2 %	GS	GS %
AIV	9.49	11.88	59.80	1.43	15.06
APM	74.74	85.25	75.17	7.90	10.57
NDM	111.29	112.57	56.17	0.72	0.65
AC	1.31	1.41	22.11	0.02	1.60
VA	2.82	3.10	57.65	0.16	5.65
NR	3.10	3.46	51.87	0.19	6.09
NV	97.55	110.85	63.02	8.38	8.60
PCS	12.59	14.88	66.20	1.52	12.04
TO	15.25	19.06	77.74	2.96	19.40
PG	2719.88	4020.90	67.21	874.42	32.15
GANHO TOTAL				897.69	111.80

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

Tabela 7. Médias das 10 características dos indivíduos selecionados pelo índice baseado na soma de “ranks” de Mulamba e Mock (Experimento I), safra 2012/2013, UNESP – Jaboticabal

Linhagens	AIV (cm)	APM (cm)	NDM	AC	VA	NR	NV	PCS (g)	TO (%)	PG (kg.ha ⁻¹)
10	15.80	103.60	118.33	1.17	3.67	3.27	94.33	17.80	21.52	5534.53
17	8.67	77.07	105.00	1.17	3.17	3.27	116.73	15.55	21.90	4492.83
15	9.60	80.40	107.67	1.17	3.33	3.08	96.18	15.97	20.83	3941.40
18	16.93	91.73	108.67	1.67	3.17	3.93	100.73	16.12	19.67	2993.99
20	12.93	78.13	114.67	1.67	2.33	3.53	105.93	12.55	21.60	4647.71
22	7.80	73.27	118.00	1.33	3.17	4.47	161.53	12.99	14.04	3244.35
2	11.40	92.53	115.67	1.67	2.83	2.67	100.53	13.15	13.85	3291.51

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

Tabela 8. Estimativas das médias da população original (Xo) e da população selecionada (Xs), herdabilidade (h^2 %) e ganho com a seleção (GS) em 46 linhagens F₆-F₇ de soja (Experimento II), safra 2012/2013, UNESP – Jaboticabal

Variáveis	Xo	Xs	h^2 %	GS	GS %
AIV	6.82	8.53	30.42	0.52	7.63
APM	63.81	78.17	68.69	9.86	15.46
NDM	111.22	112.02	25.93	0.21	0.19
AC	1.15	1.21	34.80	0.02	1.95
VA	2.29	2.87	33.87	0.20	8.61
NR	3.03	2.95	36.00	-0.03	-1.00
NV	91.58	87.73	21.67	-0.83	-0.91
PCS	13.43	15.85	51.83	1.25	9.34
TO	16.05	19.17	38.24	1.19	7.44
PG	2316.36	3465.75	19.88	228.50	9.86
GANHO TOTAL				240.90	58.57

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

Tabela 9. Médias das 10 características dos indivíduos selecionados pelo índice baseado na soma de “ranks” de Mulamba e Mock (Experimento II), safra 2012/2013, UNESP – Jaboticabal

Linhagens	AIV (cm)	APM (cm)	NDM	AC	VA	NR	NV	PCS (g)	TO (%)	PG (kg.ha ⁻¹)
41	6.72	72.67	123.33	1.00	3.00	1.94	76.72	16.17	21.03	4114.00
48	9.44	65.94	111.67	1.00	2.67	3.11	77.78	16.58	21.72	3806.59
33	6.06	91.22	111.67	1.00	3.17	1.83	78.94	16.34	19.74	3740.64
64	10.72	72.78	108.67	1.17	3.00	4.78	85.94	20.57	20.85	3309.71
34	9.06	85.11	107.33	1.00	3.33	3.50	91.17	15.91	18.99	3374.53
40	5.72	84.83	108.33	1.67	2.83	3.06	114.00	16.62	20.01	3366.34
29	9.89	81.56	114.00	1.17	2.83	2.94	96.56	16.73	20.07	3355.58
39	8.33	94.61	122.67	1.50	2.83	2.83	100.33	17.29	19.30	3737.25
57	7.94	74.67	111.33	1.17	2.67	2.56	95.44	12.17	13.99	4070.08
32	6.56	69.39	115.00	1.33	2.67	2.22	83.00	11.58	13.62	3819.28
25	11.67	72.72	108.00	1.17	2.83	3.00	89.11	16.56	18.28	3227.18
67	8.56	80.61	110.67	1.17	2.67	4.28	85.78	16.16	20.78	2946.26
56	9.89	83.94	108.33	1.67	3.00	4.06	102.56	17.77	20.49	2704.25
26	8.83	64.39	107.33	1.00	2.67	1.17	50.94	11.42	19.54	2948.76

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

4. CONCLUSÕES

As estimativas de parâmetros genéticos permitiram identificar que os caracteres APM, PCS, TO e PG para o experimento I e APM e PCS para o experimento II estão mais propensos para à seleção por apresentarem valores de herdabilidade acima de 0,5 e razão entre o coeficiente de variação genético (Cvg) e o coeficiente de variação ambiental (Cve) acima de 1.

Para o experimento I foi possível identificar que os caracteres APM e NDM estão diretamente correlacionados com PG e que a seleção direta sobre eles é efetiva. Para o experimento II foi possível identificar que o caractere APM está diretamente correlacionado com PG e que o caractere AC está indiretamente correlacionado com PG.

O índice de Mulamba e Mock detectou linhagens superiores, destacando as linhagens 10,15,17 e 20 para o experimento I e as linhagens 32, 41,48, e 57 para o experimento II.

5. REFERÊNCIAS

BELSLEY, D.A.; KUH, E.; WELCH, R.E. **Regression diagnostics**: identifying data and sources of collinearity. New York: John Wiley e Sons, 1980.

BIZARI, E. H. **Índices de seleção para caracteres agronômicos em populações segregantes de soja**. 2014. 44 f. Dissertação (Mestrado em Agronomia: Área de concentração em Genética e Melhoramento de Plantas) – Universidade Estadual Paulista – UNESP, Jaboticabal, 2014.

BOX, G. E. P.; COX, D. R. An analysis of transformations (with discussion). **Journal of the Royal Statistical Society** v.26, 211–252, 1964.

CARVALHO, C.G.P.; ARIAS, C. A. A.; TOLEDO, J. F. F.; OLIVEIRA, M. F.; VELLO, N. A. Correlações e análise de trilha em linhagens de soja semeadas em diferentes épocas. **Pesquisa Agropecuária Brasileira**, Brasília, v. 37, n. 3, p. 311-320, 2002.

COSTA, M.M.; MAURO, A. O.; UNÊDA-TREVISOLI, S. H.; ARRIEL, N. H. C.; BÁRBARO, I. M.; MUNIZ, F. R. S. Ganho genético por diferentes critérios de seleção em populações segregantes de soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 39, n.11, p.1095-1102, 2004.

CRUZ, C.D. **Aplicação de algumas técnicas multivariadas no melhoramento de plantas**. 1990. 188f. Tese (Doutorado) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

CRUZ, C. D.; REGAZZI, A. J. **Modelos biométricos aplicado no melhoramento genético**. 2ª ed.rev. Viçosa:UFV, 2001. 390p.

CRUZ, C. D. **Programa GENES** - Aplicativo computacional em genética e estatística. Disponível em: <www.ufv.br/dbg/genes/genes.htm>. Versão Windows 2007, 2007.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Tecnologias de produção de Soja Região Central do Brasil 2012 e 2013**. Londrina, 2011. 262 p. Disponível em: <<http://www.cnpso.embrapa.br/download/SP15-VE.pdf>>. Acesso em: 08 maio 2014.

FEHR, W. R.; CAVINESS, J. A. **Stages of soybean development**. Ames: Iowa State University, 1977, 11p. (Special Report, 80).

MONTGOMERY, D.C.; PECK, E.A. **Introduction to linear regression analysis**. New York: John Wiley e Sons, 1981.

MULAMBA, N.N.; MOCK, J.J. Improvement of yield potential of the Ecto Blanco maize (*Zea mays* L.) population by breeding for plant traits. **Egypt J. Gen. Cytol.**, Alexandria, v. 7, p.40-51, 1978.

MUNIZ, F.R.S.; MAURO, A. O.; UNÊDA-TREVISOLI, S. H.; OLIVEIRA, J. A.; BÁRBARO, I. M.; ARRIEL, N. H. C.; COSTA, M. M. Parâmetros genéticos e fenotípicos em populações segregantes de soja. **Revista brasileira de oleaginosas e fibrosas**, Campina Grande, v.6, n.3, 2002.

MUNIZ, F. R. S.; MAURO, A. O.; UNÊDA-TREVISOLI, S. H.; BÁRBARO, I. M.; COSTA, M. M. Seleção indireta e estudo genético entre gerações de cruzamentos biparentais em soja. **Revista Ceres**, Viçosa, v.54, n.315, p.434-440, 2007.

NOGUEIRA, A. P. O.; SEDIYAMA, T.; SOUSA, L. B.; HAMAWAKI, O. T.; CRUZ, C. D.; PEREIRA, D. G.; MATSUO, É. Análise de trilha e correlações entre caracteres em soja cultivada em duas épocas de semeadura. **Bioscience Journal**, Uberlândia, v. 28, n. 6, 2012.

NETO, J.T.F.; VELLO, N.A. Avaliação de progênies F4:3 e F5:3 e estimativas de parâmetros genéticos com ênfase para porcentagem de óleo, produtividade de grãos e óleo em soja. **Ciência e Agrotecnologia**, Lavras, v. 25, n. 4, p. 812-820, 2001.

OLIVEIRA, A. C. B.; VIANA, J. M. S.; CRUZ, C. D.; SEDIYAMA, C. S. Herdabilidade e correlações em plantas F2 de soja cultivadas em diferentes condições ambientais. **Acta Scientiarum**, Maringá, v. 22, n. 4, p. 889-893, 2000.

PERINI, L. J.; JÚNIOR, N. D. S. F.; DESTRO, D.; PRETE, C. E. C. Componentes da produção em cultivares de soja com crescimento determinado e indeterminado. **Semina: Ciências Agrárias**, Londrina, v. 33, n. 65(Suplemento 1), p. 2531-2544, 2012.

R Core Team (2013). R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. URL <http://www.R-project.org/>.

RESENDE, M. D. V. **Selegen-remi/blup**: Sistema estatístico e seleção genética computadorizada. Disponível em: <http://www.det.ufv.br/ppestbio/docs/Selegen_RemIBlup.zip>. Versão Janeiro 2008, 2008

RIGON, J. P. G.; CAPUANI, S.; BRITO NETO, J. F.; ROSA, G. M.; WASTOWSKI, A. D.; RIGON, C. A. G. Dissimilaridade genética e análise de trilha de cultivares de soja avaliada por meio de descritores quantitativos. **Revista Ceres**, Viçosa, v.59, n.2, p.233-240, 2012.

ROSSMANN, H. **Estimativas de parâmetros genéticos e fenotípicos de uma população de soja avaliada em quatro anos**. 2001 91f. Dissertação (Mestrado em Agronomia: Área de concentração em Genética e Melhoramento de Plantas). Universidade de São Paulo – Esalq, Piracicaba, 2001.

SAS Institute INC.SAS/STAT™ **SAS user's guide for windowns environment**. 12.1 ed. Cary : SAS Institute, 2010.

SEDIYAMA, T; TEIXEIRA, R.C.; REIS, M.S. Melhoramento da Soja. In: BORÉM, A. (Ed.). **Melhoramento de Espécies Cultivadas**. 2. ed. Viçosa: UFV, 2005. Cap. 14. p. 562-572.

SILVEIRA, G.D.; DI MAURO, A.O.; CENTURION, M.A.P.C. Seleção de genótipos de soja para a região de Jaboticabal (SP) (Ano agrícola 2003-2004). **Científica**, Jaboticabal, v.34, n.1, p.92-98, 2006.

WRIGHT, S. Correlation and causation. **Journal of Agricultural Research**, v. 20, p. 557-585, 1921.

CAPÍTULO 3 – SELEÇÃO DE LINHAGENS AVANÇADAS DE SOJA POR ABORDAGENS MULTIVARIADAS

RESUMO – Objetivou-se no presente estudo a seleção de genótipos superiores de soja, com ênfase na produtividade de grãos, por meio de análises multivariadas. Foram conduzidos dois experimentos no ano agrícola 2012/2013 na Fazenda de Ensino, Pesquisa e Extensão (FEPE), da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Jaboticabal. A divisão em dois experimentos se fez necessária pois 23 linhagens estavam na geração F₇-F₈ (I) e 46 estavam na geração F₆-F₇ (II). Os experimentos foram instalados em um delineamento de blocos ao acaso com três repetições e a caracterização fenotípica foi realizada por meio da avaliação dos seguintes caracteres agromorfológicos: altura de inserção da primeira vagem (AIV), altura de planta na maturidade (APM), número de dias para maturidade (NDM), acamamento (AC), valor agrônômico (VA), número de ramos (NR), número de vagens (NV), peso de cem sementes (PCS), teor de óleo (TO) e produtividade de grãos (PG). Os dados foram submetidos as análises de componentes principais, de agrupamento pelo método não hierárquico, discriminante linear e discriminante por Kernel Gaussiano. O primeiro experimento ficou melhor descrito quando subdividido em 5 grupos e o segundo em 3 grupos. As variáveis AIV, APM, NR, NV e PG foram importantes nos dois experimentos por componentes principais enquanto que TO e PCS teve importância somente no primeiro e AC, VA e NDM somente no segundo experimento. As análises discriminantes linear e utilizando Kernel confirmaram as subdivisões em 5 e 3 populações nos experimentos 1 e 2. As análises multivariadas confirmaram ser importantes ferramentas na separação em subpopulações do conjunto original multivariado bem como compreender a relação entre genótipos e variáveis. Foram destacados 4 genótipos no experimento I e 14 genótipos no experimento II.

Palavras-chave: Análise discriminante, análise de agrupamento, componentes principais, *Glycine max*, melhoramento genético.

1. INTRODUÇÃO

A soja *Glycine max* (L.) Merrill é originária de clima temperado, com ampla adaptabilidade aos climas tropicais e subtropicais. O interesse pela cultura se deve principalmente por ser um alimento que fornece grandes quantidades de óleo e proteína (aproximadamente 20% e 40% respectivamente) a um custo relativamente barato, além de seus subprodutos compor grande parte da alimentação humana e animal.

De acordo com Sedyama et al. (2005) desde o início das pesquisas com a oleaginosa no Rio Grande do Sul em 1930, o melhoramento genético vem contribuindo significativamente para o sucesso da cultura no país, o que tem possibilitado o desenvolvimento de cultivares cada vez mais adaptadas e produtivas às mais variadas condições edafoclimáticas.

Dentre os procedimentos de um programa de melhoramento genético, a seleção de progênies e linhagens superiores é uma das etapas mais complexas, isto porque grande parte dos caracteres de interesse agrônomo como a produtividade de grãos são caracteres quantitativos, ou seja governado por vários genes e com forte influência ambiental, existindo ainda quase sempre uma inter-relação entre os caracteres dificultando o processo seletivo.

Normalmente as características analisadas em um programa de melhoramento genético de soja são: ciclo, hábito de crescimento, altura da planta na maturidade, altura de inserção da primeira vagem, resistência ao acamamento, resistência à deiscência das vagens, qualidade da semente, resistência a pragas e doenças, resistência a herbicidas, proteína e óleo (SEDIYAMA et al., 2009). As características avaliadas podem variar de um programa de melhoramento para outro e também em função do objetivo do programa.

Avaliar características agrônomicas da soja e analisar todos os parâmetros ao mesmo tempo em função das informações coletadas durante o ensaio, se torna interessante quando o objetivo é selecionar genótipos superiores, uma vez que muitos fatores estão relacionados de maneira que seus diferentes efeitos são melhores interpretados se estudados em conjunto (Vianna et al., 2013).

Uma das técnicas que permite analisar todos os parâmetros conjuntamente é a análise de componentes principais, que tem sido amplamente utilizada no melhoramento genético por permitir a simplificação do conjunto de dados resumindo as informações originalmente contidas em um grupo de v variáveis, em poucos componentes, os quais apresentam as propriedades de reter o máximo da variação originalmente disponível e ser independentes entre si (CRUZ, 2006).

Análise de Componentes Principais (ACP) é considerada uma técnica estatística exploratória utilizada na tentativa de compreender o inter-relacionamento entre as variáveis originais. Os componentes principais são não correlacionados o que é interessante pois um pesquisador estando com um problema envolvendo variáveis originais inter-relacionadas pode analisar um conjunto menor de variáveis não correlacionadas que são os componentes principais. Num conjunto grande de variáveis nem todas têm quantidade de informação relevante podendo através da ACP selecionar aquelas que mais possuem quantidade de informação relevante (FERRAUDO, 2010).

Embora, do ponto de vista matemático, a Análise de Agrupamento apresente uma metodologia bastante simples e a Análise de Componentes Principais uma metodologia mais complexa, espera-se uma boa concordância entre os resultados de ambas, devendo uma ser utilizada como complemento da outra (FERRAUDO, 2010).

O uso de técnicas multivariadas como a análise de componentes principais e a análise de agrupamento ganharam destaque em programas de melhoramento genético e vem sendo muito utilizada, devido à grande eficiência que essas análises tem na exploração da variabilidade existente dentro da população. O trabalho de Vianna et al. (2013) e o trabalho de Dallastra et al. (2014) comprovam a eficiência das análises na seleção de progênies superiores de soja.

Diante do exposto, o objetivo do presente trabalho consistiu na seleção de genótipos superiores de soja, com ênfase na produtividade de grãos, por meio de análises multivariadas.

2. MATERIAL E MÉTODOS

2.1 Manejo experimental e tratamentos utilizados

Foram conduzidos dois experimentos no ano agrícola 2012/2013 na Fazenda de Ensino, Pesquisa e Extensão (FEPE), da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Jaboticabal, localizada ao norte do Estado de São Paulo, a 21°15' de latitude sul e 48°18' de longitude oeste, com altitude aproximada de 595 m. A divisão em dois experimentos se fez necessária pois 23 linhagens estavam na geração F₇-F₈ (I) e 46 estavam na geração F₆-F₇ (II).

O solo da área experimental é classificado como Latossolo Vermelho Eutroférrico Argiloso, com relevo suavemente ondulado e o clima é do tipo Cwa, segundo a classificação de Köppen.

A semeadura foi realizada mecanicamente no dia 27 de novembro de 2012, para o primeiro experimento e no dia 29 de novembro de 2012 para o segundo experimento, ambos foram desbastados e o estande final ficou com densidade de 15 plantas m⁻¹. No plantio foram utilizados 350 kg ha⁻¹ da fórmula 00-30-15, sendo que as parcelas experimentais foram mantidas durante todo o ciclo da cultura, com rigoroso controle de pragas, doenças e plantas infestantes, conforme recomendações para a cultura da soja (EMBRAPA, 2011).

Tabela 1. Linhagens pertencentes ao Programa de melhoramento genético de soja da Unesp-FCAV. Ano agrícola 2012/2013, Jaboticabal-SP

Experimento I		Experimento II	
Linhagens		Linhagens	
1		24	47
2		25	48
3		26	49
4		27	50
5		28	51
6		29	52
7		30	53
8		31	54
9		32	55
10		33	56
11		34	57
12		35	58
13		36	59
14		37	60
15		38	61
16		39	62
17		40	63
18		41	64
19		42	65
20		43	66
21		44	67
22		45	68
23		46	69

Linhagens proveniente do cruzamento PI 200456 x MGBR46 (Conquista)

Os experimentos foram instalados considerando delineamento de blocos ao acaso com três repetições. Cada parcela experimental consistiu de quatro linhas de 5 m de comprimento, espaçadas de 0,5 m entre si, sendo considerada como área útil as duas linhas centrais, desprezando-se 0,5 m de cada extremidade, totalizando 4 m². Foram semeadas 23 linhagens no primeiro experimento e 46 linhagens no segundo experimento (Tabela 1), ambas pertencentes ao programa de melhoramento genético de soja da UNESP, Campus de Jaboticabal.

2.2 Caracteres agronômicos

A caracterização fenotípica foi realizada por meio da avaliação dos seguintes caracteres agromorfológicos: Altura de inserção da primeira vagem (AIV): obtido pela medida, do colo da planta até a altura da inserção da primeira vagem (cm), considerando a média de seis plantas dentro da área útil. Altura de planta na maturidade (APM): obtido pela medida, da altura do colo da planta até o ápice da haste principal (cm), considerando a média de seis plantas tomadas ao acaso dentro da área útil. Número de dias para maturidade (NDM): definido como o período em dias entre a data de emergência das plantas até a data em que aproximadamente 95% das vagens apresentaram-se maduras (estádio R8) (FEHR & CAVINESS, 1977). Acamamento (AC): caráter avaliado no estágio R8 (FEHR & CAVINESS, 1977) da planta por meio de uma escala de notas visuais, variando de 1 (todas as plantas eretas) a 5 (todas as plantas acamadas); valor agronômico (VA): caráter avaliado no estágio R8 de desenvolvimento da planta, por meio de uma escala de notas visuais, a qual varia de 1 (plantas com características agronômicas ruins) a 5 (plantas com ótimas características agronômicas), sendo a nota atribuída representativa de um conjunto de caracteres visuais (arquitetura da planta, quantidade de vagens cheias, vigor e sanidade da planta, debulha prematura das vagens, acamamento e retenção foliar na maturidade); número de ramos (NR): obtido por meio da contagem de ramos de cada planta, sendo tomado o valor médio de seis plantas dentro da área útil; número de vagens (NV): obtido por meio da contagem da quantidade de vagens de cada planta sendo tomado o valor médio de seis plantas dentro da área útil; peso de cem sementes (PCS): obtido por meio de três amostras de 100 sementes, utilizando uma balança de precisão de um grama; teor de óleo (TO): porcentagem de óleo contida nos grãos de soja, sendo obtida com o equipamento NIR Bruker modelo Tango, que realiza a medição dos teores de óleo sem que ocorra a destruição das sementes; produtividade de grãos (PG): obtida a partir da colheita das plantas da área útil da parcela, sendo as mesmas trilhadas e o peso dos grãos corrigido para 13% de umidade, posteriormente o peso corrigido é convertido em kg. ha⁻¹.

2.3 Análises estatísticas

Inicialmente os dados foram padronizados resultando em média nula e variância unitária para todas as variáveis, de acordo com a equação:

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

Onde:

$i = 1, 2, \dots, n$ objetos;

$j = 1, 2, \dots, p$, variáveis;

\bar{x}_j e s_j = média e o desvio padrão da coluna j .

2.3.1 Cálculo dos componentes principais

A variabilidade original foi decomposta em dois autovetores (componentes principais) construídos com os autovalores da matriz de covariância. Os autovetores são combinações lineares das variáveis originais buscando maximizar a informação relevante (Hair, et. al, 2005).

2.3.2 Cálculo da variância contida em cada componente principal

A % da variância total contida em cada componente principal é obtida pela fórmula:

$$CP_h = \frac{\lambda_h}{\text{Traço}(C)} 100$$

Onde:

C = matriz de covariâncias dos dados originais padronizados;

λ_h = h -ésima raiz característica (autovalor) da matriz C ,

onde $\text{Traço}(C) = \lambda_1 + \lambda_2 + \dots + \lambda_h$

De acordo com o critério proposto por Kaiser (1958), somente são considerados autovalores acima de 1 pois geram componentes com quantidade relevante de informação das variáveis originais. Abaixo de 1 a quantidade de informação retida no componente não é relevante.

2.3.3 Correlações das características (variáveis) com os componentes principais

A correlação das características (variáveis) com os componentes principais é obtida através da fórmula:

$$r_{xj}(cp_h) = \frac{a_{jh}\sqrt{\lambda_h}}{S_j}$$

Onde:

S_j = desvio padrão da variável j ;

a_{jh} = coeficiente da variável j no h -ésimo componente principal;

λ_h = h -ésima raiz característica (autovalor) da matriz de covariância.

Segundo Hair, et. al, 2005 na fase exploratória pode ser aceito valores de cargas nos componentes a partir de 0,35. Isto foi considerado no experimento I e 0,55 no experimento II.

2.3.4 Análise de agrupamento por método não hierárquico

A análise de agrupamento por método não hierárquico é uma técnica que permite classificar (indivíduos ou objetos) em subgrupos excludentes. É um problema de otimização em que se pretende de uma forma geral maximizar a homogeneidade de objetos ou indivíduos dentro de grupos e maximizar a heterogeneidade entre os grupos (FERRAUDO, 2010). A análise foi processada com o algoritmo K-médias.

2.3.5 Análise discriminante linear

Uma variável estatística discriminante é a combinação linear das duas (ou mais) variáveis independentes que melhor discriminarão entre os objetos (genótipos) nos grupos definidos a priori. A discriminação é conseguida estabelecendo-se os pesos da variável estatística para cada variável independente para maximizar as diferenças entre os grupos. A variável estatística para uma análise discriminante, também conhecida como função discriminante é determinada a partir da seguinte equação (Hair et al., 2009):

$$z_{jk} = a + w_1x_{1k} + w_2x_{2k} + \dots + w_nx_{nk}$$

Onde:

z_{jk} = Escore Z discriminante da função discriminante j para o objeto k;

a = Intercepto;

w_i = Peso discriminante para a variável independente i;

x_{ik} = Variável independente i para o objeto k.

Segundo Ferraudo (2010) a análise discriminante é aplicada quando a variável dependente é categórica (nominal ou não métrica) e as variáveis independentes são métricas. Uma vantagem da análise discriminante é reduzir o espaço dimensional das variáveis independentes para G-1 dimensões, onde G é o número de grupos estabelecidos a priori.

2.3.5.1 Análise discriminante por Kernel Gaussiano

Esta análise deve ser aplicada para avaliar a densidade das populações de k-grupos definidos a priori. Para a realização desta análise seguiu-se as recomendações do artigo de Duong (2007), com a utilização do software estatístico R Core Team (2013) e do pacote estatístico “ks” (DUONG, 2014). A análise discriminante por Kernel Gaussiano foi processado com os escores dos dois primeiros componentes principais.

3. RESULTADOS E DISCUSSÃO

No experimento I, os dois primeiros componentes principais explicaram 58,87% da variância contida nas 10 variáveis originais (Tabela 2).

Tabela 2. Matriz de autovalores e estatística, para 10 características avaliadas em 23 genótipos de soja (Experimento I), safra 2012/2013, UNESP - Jaboticabal

Número do autovalor	Autovalor	Variância Total (%)	Autovalor acumulado	Variância acumulada (%)
1	3,04	30,44	3,04	30,44
2	2,84	28,42	5,89	58,87
3	1,38	13,84	7,27	72,70
4	1,22	12,24	8,49	84,95
5	0,63	6,34	9,13	91,29
6	0,40	3,99	9,53	95,28
7	0,30	2,97	9,83	98,25
8	0,09	0,92	9,92	99,17
9	0,05	0,46	9,96	99,64
10	0,04	0,36	10,00	100,00

O primeiro componente principal (CP1) reteve 30,44% da variância original e conforme pode ser observado na Tabela 3, as principais variáveis que explicaram essa retenção foram: AIV, APM, NR, e NV apresentando valores de correlações com o CP1 de -0,463, -0,454, 0,408 e 0,401 respectivamente.

O segundo componente principal (CP2), reteve 28,42% da variância original e conforme tabela 3, as principais variáveis que ficam retidas nesse componente foram: TO, PCS e PG apresentando valores de correlações com o CP2 de -0,519, -0,511 e -0,399 respectivamente (Tabela 3).

Tabela 3. Correlação entre as 10 variáveis e os componentes principais dos genótipos de soja (Experimento I), safra 2012/2013, UNESP – Jaboticabal

	CP1	CP2
AIV (cm)	-0,463	-0,168
APM (cm)	-0,454	-0,133
NDM	-0,253	0,073
AC	0,144	-0,120
VA	-0,285	-0,201
NR	0,408	-0,326
NV	0,401	-0,314
PCS (g)	0,058	-0,511
TO	-0,029	-0,519
PG (kg,ha ⁻¹)	-0,285	-0,399

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agronômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

O plano bidimensional formado pelos componentes CP1 (30,44%) e CP2 (28,42%) reteve no total 58,87% da variância original e ficou caracterizado pelas variáveis AIV, APM, NR, NV, TO, PCS e PG, discriminando os genótipos 3,10,15,16,17,18,20,22,23 (Figura 1). O plano bidimensional formado pelos componentes CP1 e CP3 reteve 44,28% da variância remanescente e discriminou os mesmos genótipos do plano bidimensional CP1 e CP2, não contribuindo com informações relevantes para o objetivo do trabalho.

As variáveis que permitiram identificar os melhores genótipos foram VA, PG, TO, e NV.

Com base na análise de correlações genóticas feita no capítulo anterior sabemos que existe correlação genética entre VA e PG de 0,24, para o caráter TO a correlação com PG é 0,44, e o caráter NV apresentou correção com PG de 0,12.

Isto comprova que selecionando genótipos com base nos caracteres VA, TO, e NV, podem ser obtidos ganhos indiretos para produtividade de grãos, além é claro da seleção direta sobre o caráter PG.

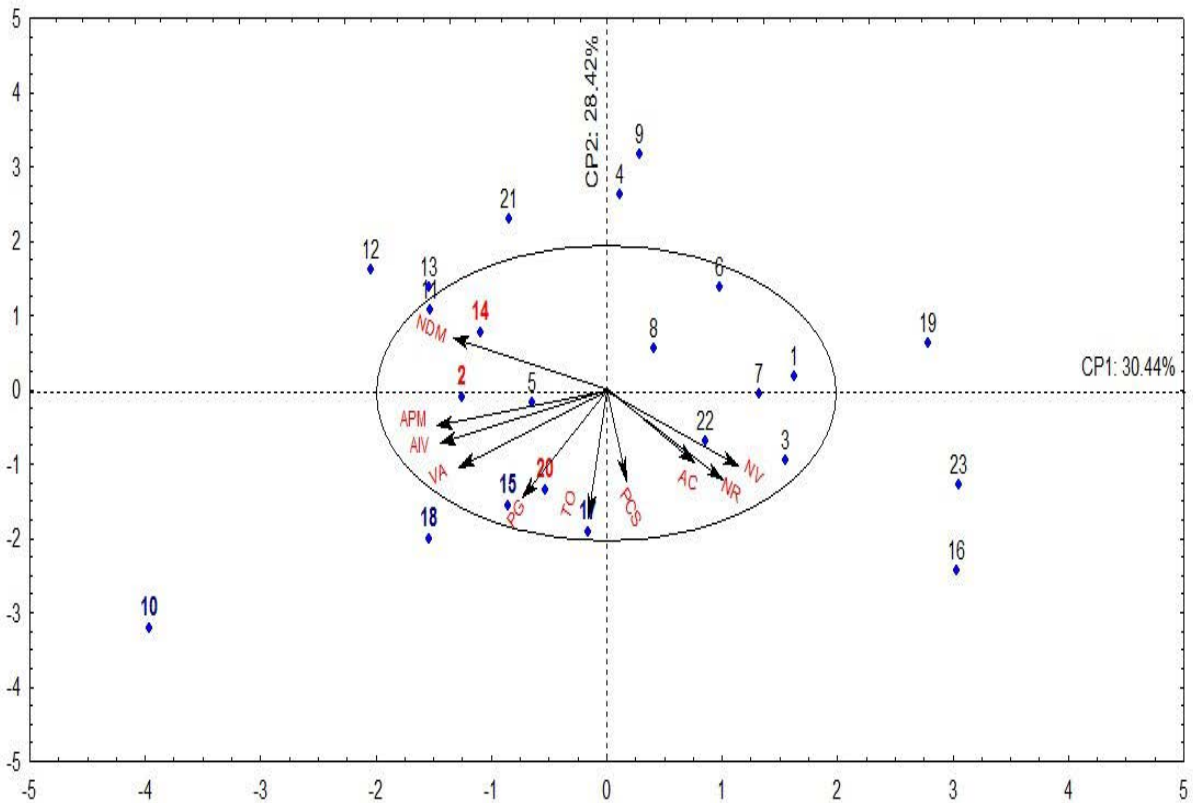


Figura 1. Gráfico biplot com dispersão de 23 genótipos de soja (Experimento I) em função dos componentes principais CP1 e CP2, contendo a projeção dos vetores das variáveis: AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônomico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos. Safra 2012/2013, UNESP - Jaboticabal

A análise de agrupamento por método não hierárquico K-médias apresentada na Figura 2, mostra que dois grupos se destacaram para o caráter produtividade de grãos, o principal caráter buscado em programas de melhoramento genético.

O grupo 1 ilustrado pela coloração azul, foi composto pelos genótipos 10, 15, 17 e 18 sendo esse grupo o que apresentou maior produtividade de grãos, valor agrônomico, altura de inserção da primeira vagem e altura da planta na maturidade, apresentando ainda valores intermediários de acamamento e número de dias para maturidade (Tabela 6).

É interessante a seleção de genótipos que apresentem menores valores para acamamento e número de dias para maturidade porque, assim o pesquisador vai estar contribuindo para a diminuição das perdas com a colheita mecanizada visto que selecionará genótipos com o perfil mais ereto, e também mais precoces diminuindo os custos com aplicações de inseticidas e fungicidas. Viabiliza ainda o plantio da

cultura da soja em áreas de renovação de canavial e permite o plantio do milho na safrinha.

O grupo 2 ilustrado pela coloração vermelha, foi composto pelos genótipos 2,14 e 20 que apesar de apresentar o menor valor agrônômico e maiores valores para acamamento e número de dias para maturidade, apresentou valores da altura de inserção da primeira vagem e altura da planta na maturidade menores que o grupo 1. Ainda o grupo 2 conteve o genótipo 20 que apresentou o segundo maior valor de produtividade de grãos no experimento I (Tabela 6).

A média para o caráter produtividade de grãos no grupo 1 foi de 4240,69 kg.ha⁻¹, o valor do caráter oscilou de 2993,99 a 5534,53 kg.ha⁻¹ (Tabela 6). No grupo 2, a média para o caráter produtividade de grãos foi de 3553,66 kg.ha⁻¹, o valor do caráter oscilou de 2721,76 a 4647,71 kg.ha⁻¹.

Os demais grupos apresentaram valores muito baixo para produtividade de grãos e poucos valores ideais para os demais caracteres, podendo ser descartados.

É importante observarmos a concordância da análise de K-médias (Figura 2), com a análise de componentes principais (Figura 1), onde genótipos do grupo 1 são destacados quanto a produtividade de grãos, juntamente com o genótipo 20 do grupo 2.

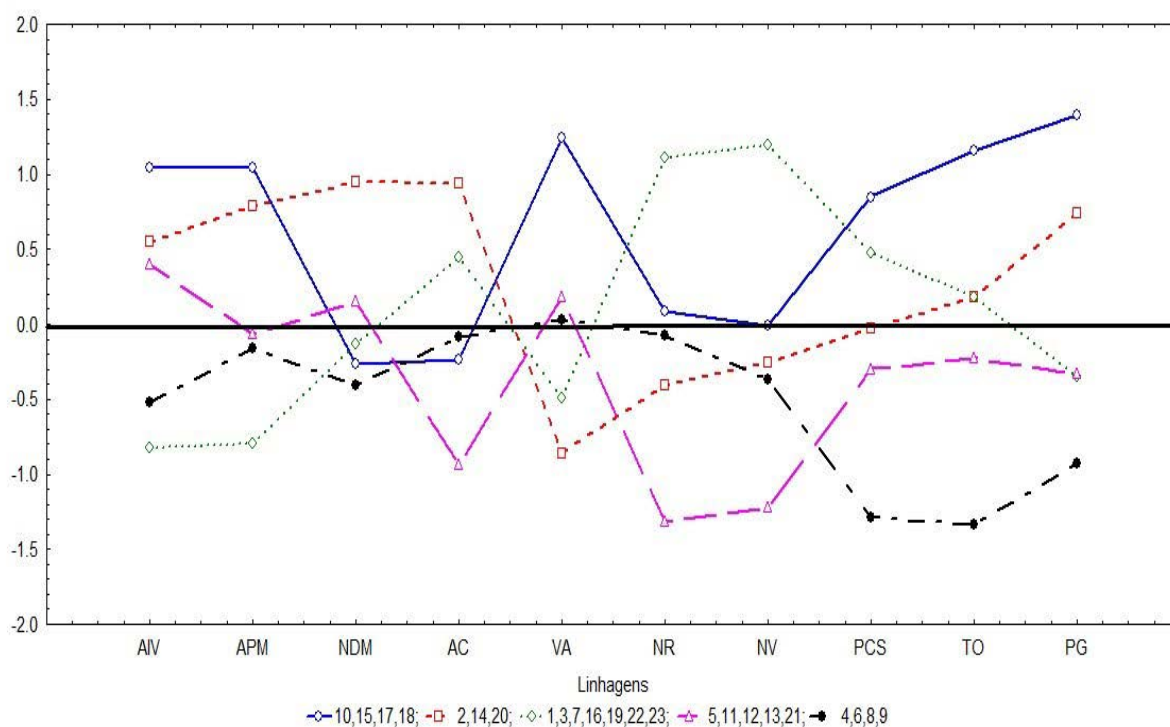


Figura 2. Distribuição dos centroides dos grupos na análise de agrupamento por k-médias formados a partir das variáveis: AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos. Experimento I. Safra 2012/2013, UNESP – Jaboticabal

As tabelas de classificação resultantes das análises discriminantes linear e utilizando Kernel (Tabelas 4 e 5) confirmam a existência de 5 populações com padrões distintos no experimento I tendo melhor eficiência na classificação a análise discriminante utilizando Kernel, pois separa populações com fronteiras não lineares.

Tabela 4. Matriz de Classificação resultante da análise discriminante linear processada com os escores dos componentes principais 1 e 2, (Experimento I), safra 2012/2013, UNESP - Jaboticabal

Grupos	Porcentagem Correta	G1	G2	G3	G4	G5
G1	100,00	4	0	0	0	0
G2	66,67	0	2	0	1	0
G3	100,00	0	0	7	0	0
G4	80,00	0	1	0	4	0
G5	100,00	0	0	0	0	4
Total	91,30	4	3	7	5	4

Tabela 5. Matriz de Classificação resultante da análise discriminante com Kernel Gaussiano processada com os escores dos componentes principais 1 e 2 (Experimento I), safra 2012/2013, UNESP - Jaboticabal

Grupos	Porcentagem Correta	G1	G2	G3	G4	G5
G1	100,00	4	0	0	0	0
G2	100,00	0	3	0	0	0
G3	100,00	0	0	7	0	0
G4	100,00	0	0	0	5	0
G5	100,00	0	0	0	0	4
Total	100,00	4	3	7	5	4

A Figura 3 mostra que é possível separar a população de genótipos em 5 subgrupos cujas densidades possuem padrões específicos.

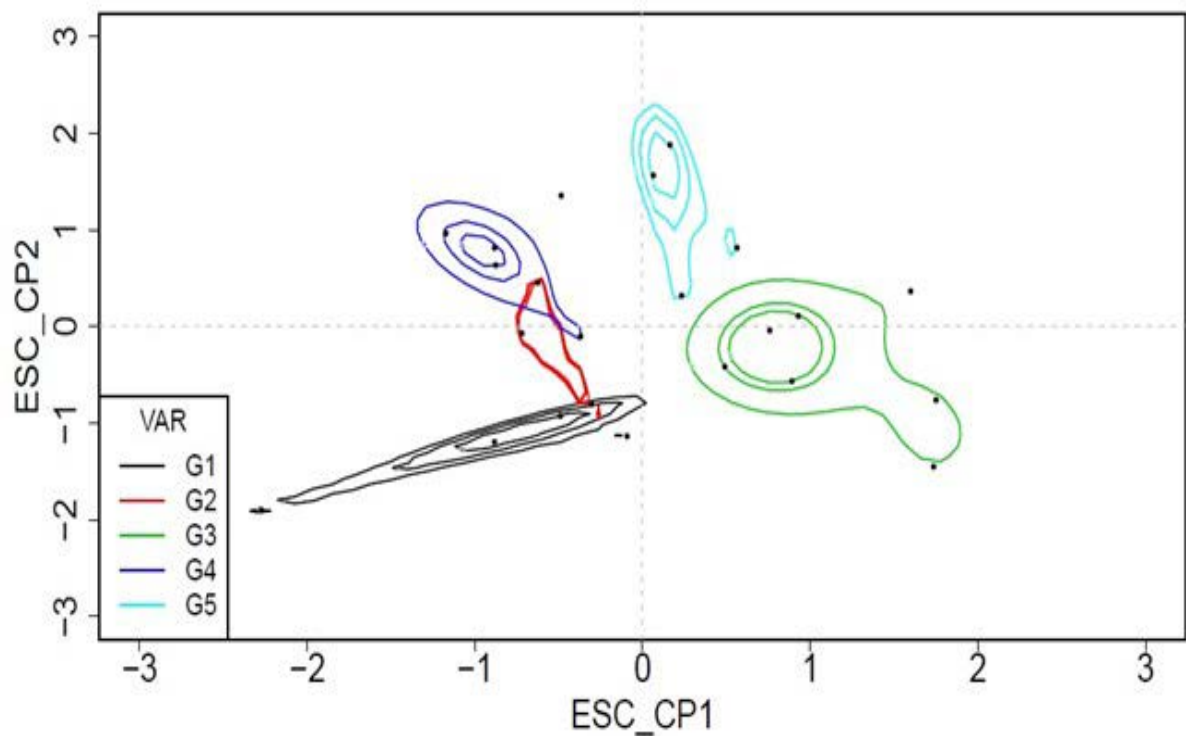


Figura 3. Densidade das subpopulações formada pelos 23 genótipos de soja, segundo análise discriminante com um Kernel Gaussiano processada com os escores dos componentes principais 1 e 2. Experimento I, safra 2012/2013, UNESP – Jaboticabal

Tabela 6. Média dos 10 caracteres avaliados em 23 genótipos de soja (Experimento I), safra 2012/2013, UNESP - Jaboticabal

	AIV	APM	NDM	AC	VA	NR	NV	PCS	TO	PG
1	7,60	77,27	110,00	1,50	2,67	4,33	146,27	11,50	13,21	2345,28
2	11,40	92,53	115,67	1,67	2,83	2,67	100,53	13,15	13,85	3291,51
3	7,13	67,80	114,67	1,50	2,67	4,00	115,60	18,59	20,91	1950,50
4	8,53	85,27	106,00	1,33	3,17	3,20	92,93	5,18	6,61	1144,08
5	9,47	78,47	106,33	1,33	3,17	2,00	68,40	15,99	19,70	2144,65
6	6,07	69,60	108,00	1,17	2,83	3,37	83,48	10,91	13,48	2042,11
7	8,80	72,80	116,33	1,67	2,50	4,27	140,13	12,48	14,19	2787,36
8	8,93	74,40	107,33	1,50	3,00	3,40	95,53	11,63	12,76	2631,12
9	9,47	69,87	115,33	1,33	2,67	2,80	86,60	5,80	6,55	1406,25
10	15,80	103,60	118,33	1,17	3,67	3,27	94,33	17,80	21,52	5534,53
11	13,00	82,67	109,67	1,17	3,00	2,27	72,47	10,13	13,11	2590,00
12	11,53	78,33	118,67	1,33	3,00	1,47	43,27	11,18	13,14	2544,93
13	10,67	80,20	108,00	1,00	3,33	1,53	63,53	10,73	14,00	2078,58
14	9,67	85,40	118,33	1,50	2,67	2,13	74,27	13,52	14,32	2721,76
15	9,60	80,40	107,67	1,17	3,33	3,08	96,18	15,97	20,83	3941,40
16	8,80	62,80	105,33	1,50	3,00	6,40	161,60	20,88	18,68	1991,75
17	8,67	77,07	105,00	1,17	3,17	3,27	116,73	15,55	21,90	4492,83
18	16,93	91,73	108,67	1,67	3,17	3,93	100,73	16,12	19,67	2993,99
19	5,40	64,33	104,67	1,83	2,50	4,40	111,07	11,47	14,88	1931,38
20	12,93	78,13	114,67	1,67	2,33	3,53	105,93	12,55	21,60	4647,71
21	9,80	59,27	117,67	0,67	2,33	0,90	54,87	12,36	13,98	2811,78
22	7,80	73,27	118,00	1,33	3,17	4,47	161,53	12,99	14,04	3244,35
23	6,07	55,67	105,33	1,00	2,67	4,80	165,33	16,92	20,32	2617,30

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônomo; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

Para o experimento II, os dois primeiros componentes explicam 48,97% da variância contida nas 10 variáveis (Tabela 7).

Tabela 7. Matriz de autovalores e estatística, para 10 características avaliadas em 46 genótipos de soja (Experimento II), safra 2012/2013, UNESP - Jaboticabal

Número do autovalor	Autovalor	Variância Total (%)	Autovalor acumulado	Variância acumulada (%)
1	3,14	31,37	3,14	31,37
2	1,76	17,60	4,90	48,97
3	1,24	12,40	6,14	61,37
4	0,97	9,70	7,11	71,07
5	0,88	8,82	7,99	79,89
6	0,72	7,24	8,71	87,13
7	0,57	5,72	9,29	92,85
8	0,39	3,86	9,67	96,71
9	0,19	1,93	9,86	98,64
10	0,14	1,36	10,00	100,00

O primeiro componente principal (CP1) reteve 31,37% da variância original e conforme pode ser observado na Tabela 7, as principais variáveis que explicaram essa retenção foram: VA, NV, AIV, NR e PG apresentando valores de correlações com o CP1 de 0,795, -0,679, 0,639, -0,596 e 0,553 respectivamente (Tabela 8).

O segundo componente principal (CP2), reteve 17,60% da variância original e conforme tabela 7, as principais variáveis que ficam retidas nesse componente foram: AC, APM e NDM apresentando valores de correlações com o CP2 de -0,725, -0,711 e -0,562 respectivamente (Tabela 8).

Tabela 8. Correlação entre as 10 variáveis e os componentes principais dos genótipos de soja (Experimento II), safra 2012/2013, UNESP – Jaboticabal

CARACTERÍSTICAS	CP1	CP2
AIV (cm)	0,639	-0,123
APM (cm)	0,573	-0,711
NDM	0,074	-0,562
AC	-0,470	-0,725
VA	0,795	-0,168
NR	-0,596	-0,283
NV	-0,679	-0,400
PCS (g)	-0,395	-0,125
TO	-0,515	0,338
PG (kg,ha ⁻¹)	0,553	0,010

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

O plano bidimensional formado pelos componentes CP1 (31,37%) e CP2 (17,60%) reteve no total 48,97% da variância original e ficou caracterizado pelas variáveis AIV, APM, AC, VA, NDM, NR, NV, e PG, discriminando os genótipos 24,26,39,40,46,51,55,56,63,65,68,25,27,34,35,41,43,45,57,58,64,67 (Figura 4). O plano bidimensional formado pelos componentes CP1 e CP3 reteve 43,77% da variância remanescente e permitiu a discriminação dos mesmos genótipos do plano bidimensional CP1 e CP2, não contribuindo com informações relevantes para o objetivo do trabalho. A dispersão dos 46 genótipos de soja no gráfico biplot da Figura 4 permitiu identificar genótipos com propriedades específicas (Figura 5).

As variáveis que permitiram identificar os melhores genótipos foram NDM, APM, VA, AIV, AC e PG.

Com base na análise de correlações genóticas feita no capítulo anterior sabemos que existe correlação genética entre NDM x PG de 0,14, APM x PG de 0,37, VA x PG de 0,37, AIV x PG de 0,18 e AC x PG de -0,24.

As correlações genéticas comprovam que selecionando genótipos com base nos caracteres NDM, APM, VA, AIV, e AC podem ser obtidos ganhos indiretos para produtividade de grãos, além é claro da seleção direta sobre o caráter PG.

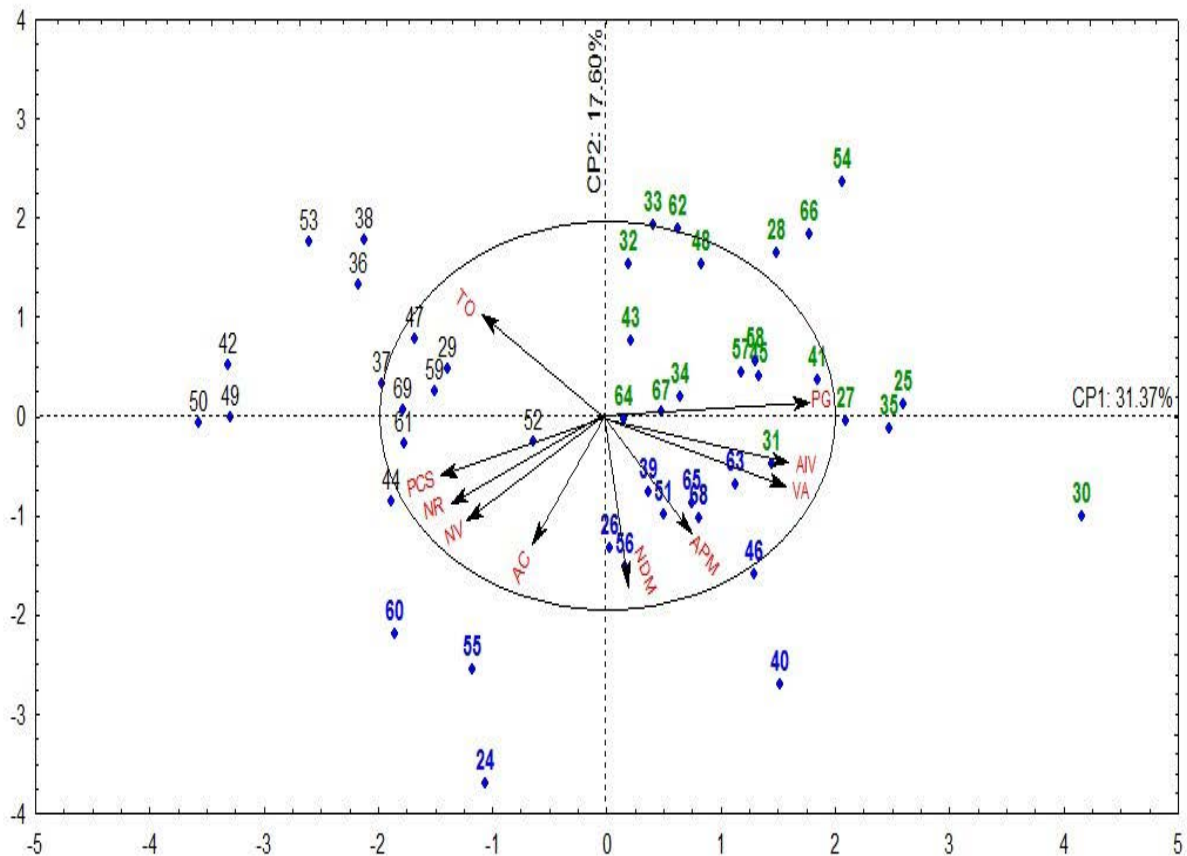


Figura 4. Gráfico biplot com dispersão de 46 genótipos de soja (Experimento II) em função dos componentes principais CP1 e CP2, contendo a projeção dos vetores das variáveis: AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônomo; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos. Safra 2012/2013, UNESP - Jaboticabal

A análise de K-médias apresentada na Figura 5, mostra a formação de 3 grupos, sendo que dois deles se destacaram para o caráter produtividade de grãos.

O grupo 3 ilustrado pela coloração verde, foi composto pelos genótipos 25, 27, 28, 30, 31, 32, 33, 34, 35, 41, 43, 45, 48, 54, 57, 58, 62, 64, 66 e 67, grupo esse caracterizado pela maior produtividade de grãos, valor agrônomo e altura de inserção da primeira vagem, apresentando valores intermediários para teor de óleo, número de dias para maturidade e altura de planta na maturidade.

A média para o caráter produtividade de grãos no grupo 3 foi de 3409,68 kg.ha⁻¹, o valor do caráter oscilou de 1153,63 a 6105,13 kg.ha⁻¹, indicando a presença de grande variabilidade para o caráter dentro do grupo (Tabela 11).

O grupo 1 ilustrado pela coloração azul, foi composto pelos genótipos 24, 26, 39, 40, 46, 51, 55, 56, 60, 63, 65 e 68 que apresentaram o maior valor para altura de planta na maturidade, número de dias para maturidade e acamamento, menor valor para teor de óleo e valores intermediários para produtividade de grãos, altura de inserção da primeira vagem, valor agrônômico, número de ramos, número de vagens e peso de cem sementes.

A média para o caráter produtividade de grãos no grupo 1 foi de 3226,16 kg.ha⁻¹, o valor do caráter oscilou de 972,38 a 4883,08 kg.ha⁻¹, indicando também a presença de grande variabilidade para o caráter dentro do grupo (Tabela 11).

A média para o caráter produtividade de grãos no grupo 2 foi a menor entre os grupos, de 2301,75 kg.ha⁻¹, o valor do caráter oscilou de 854,48 a 3793,93 kg.ha⁻¹, este grupo pode ser descartado por apresentar média muito baixa para o caráter produtividade de grãos e também por apresentar valores muito inferiores para os demais caracteres de interesse agrônômico.

Como existe muita variabilidade para o caráter produtividade de grãos dentro dos grupos uma alternativa seria fazer uma seleção dentro de cada grupo, excluindo os indivíduos com valores abaixo da média para o caráter, assim os genótipos selecionados ficariam 26, 32, 46, 51, 63, 68 para o grupo 1 e 25, 30, 33, 34, 41, 48, 57, 62 para grupo 3, resultando numa pressão de seleção de 30.43%, em relação ao total de genótipos (Tabela 11).

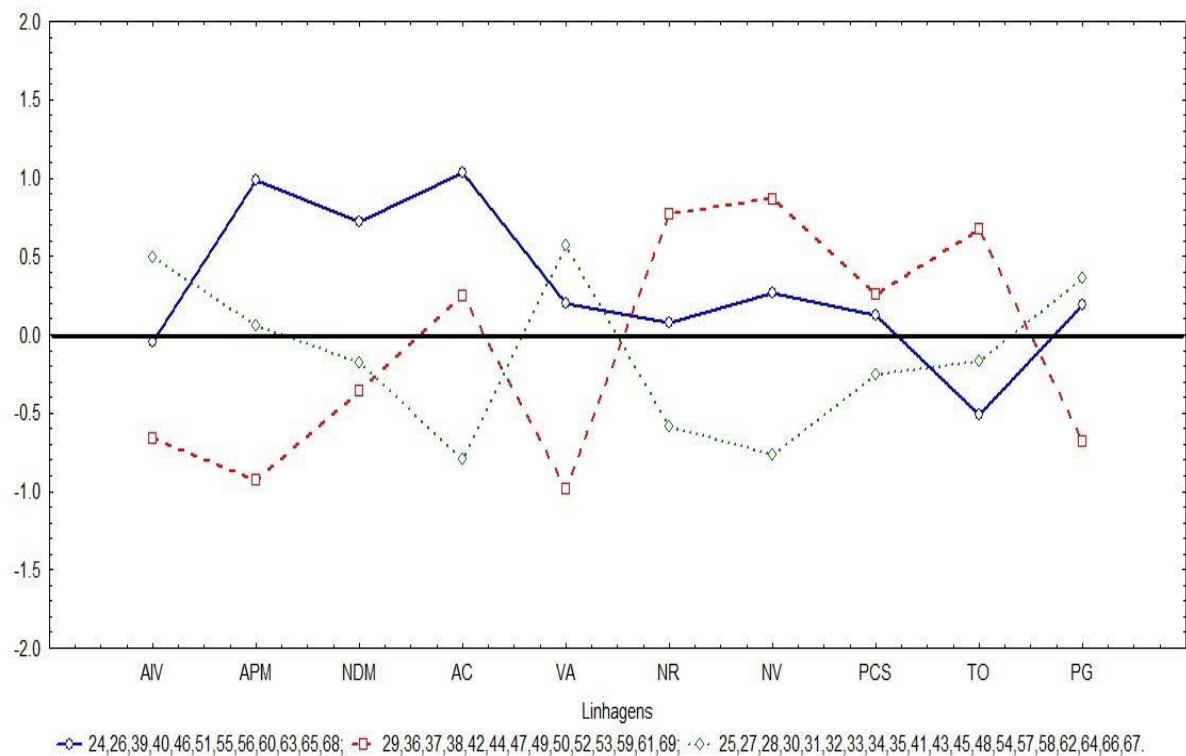


Figura 5. Distribuição dos centroides dos grupos na análise de agrupamento por k-médias formados a partir das variáveis: AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônômico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos. Experimento II. Safra 2012/2013, UNESP – Jaboticabal

As tabelas de classificação resultantes das análises discriminantes linear e utilizando Kernel Gaussiano (Tabelas 9 e 10) caracterizam 3 populações com padrões distintos tendo melhor eficiência na classificação a análise discriminante utilizando Kernel Gaussiano, pois separa populações com fronteiras não lineares.

O classificador que engloba tanto análise discriminante e componentes principais mostrou ser uma importante ferramenta na separação em subpopulações do conjunto original multivariado.

Tabela 9. Matriz de Classificação resultante da análise discriminante linear processada com os escores dos componentes principais 1 e 2, (Experimento II), safra 2012/2013, UNESP - Jaboticabal

Grupos	Porcentagem Correta	G1	G2	G3
G1	91,67	11	0	1
G2	100,00	0	14	0
G3	100,00	0	0	20
Total	97,83	11	14	21

Tabela 10. Matriz de Classificação resultante da análise discriminante com Kernel Gaussiano processada com os escores dos componentes principais 1 e 2 (Experimento II), safra 2012/2013, UNESP – Jaboticabal

Grupos	Porcentagem Correta	G1	G2	G3
G1	100,00	20	0	0
G2	100,00	0	15	0
G3	100,00	0	0	11
Total	100,00	20	15	11

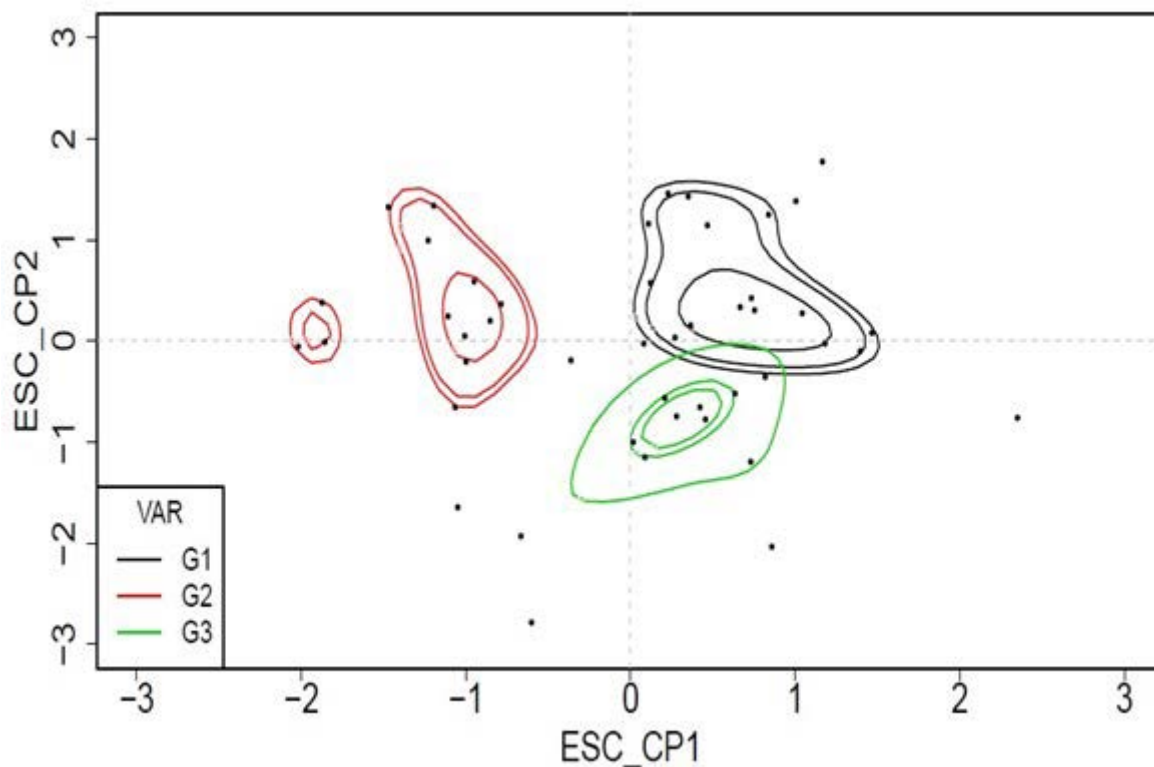


Figura 6. Densidade das subpopulações formada pelos 46 genótipos de soja, segundo análise discriminante utilizando um Kernel Gaussiano processada com os escores dos componentes principais 1 e 2. Experimento II, safra 2012/2013, UNESP – Jaboticabal

Tabela 11. Média dos 10 caracteres avaliados em 46 genótipos de soja (Experimento II), safra 2012/2013, UNESP - Jaboticabal

	AIV	APM	NDM	AC	VA	NR	NV	PCS	TO	PG
24	8,61	91,11	112,67	2,00	2,67	4,75	145,18	17,07	19,51	2729,89
25	6,06	91,22	111,67	1,00	3,17	1,83	78,94	16,34	19,74	3740,64
26	5,72	84,83	108,33	1,67	2,83	3,06	114,00	16,62	20,01	3366,34
27	11,67	72,72	108,00	1,17	2,83	3,00	89,11	16,56	18,28	3227,18
28	8,83	64,39	107,33	1,00	2,67	1,75	76,42	17,13	19,54	2948,76
29	5,61	60,78	111,67	1,33	2,17	3,72	96,50	19,10	20,56	3793,93
30	17,94	80,39	115,00	1,25	3,25	3,22	65,94	14,99	18,86	3414,50
31	9,89	81,56	114,00	1,17	2,83	2,94	96,56	16,73	20,07	3355,58
32	6,44	60,00	106,00	1,17	2,50	2,33	65,28	20,08	19,40	3021,93
33	6,50	62,61	107,00	1,00	2,50	3,67	79,06	15,66	20,88	4293,35
34	6,56	69,39	115,00	1,33	2,67	2,22	83,00	17,38	20,43	3819,28
35	9,06	85,11	107,33	1,00	3,33	3,50	91,17	15,91	18,99	3374,53
36	4,42	56,75	107,67	1,25	2,50	3,08	117,67	17,41	21,49	1187,85
37	6,02	64,58	112,33	1,16	2,61	4,23	138,63	16,11	21,79	1615,30
38	6,11	47,72	107,67	1,25	2,00	4,48	95,31	17,11	21,27	3159,23
39	8,22	77,39	115,00	1,50	2,50	3,17	79,17	16,67	20,00	2519,45
40	8,33	94,61	122,67	1,50	2,83	2,83	100,33	17,29	19,30	3737,25
41	6,72	72,67	123,33	1,00	3,00	1,94	76,72	16,17	21,03	4114,00
42	7,89	55,28	107,33	1,50	2,00	5,83	111,39	17,22	21,89	1571,13
43	5,50	67,00	115,00	1,25	2,50	1,75	77,83	17,11	20,39	2436,20
44	4,58	71,33	109,00	1,50	2,25	3,94	150,39	14,99	19,63	2097,41
45	8,67	75,94	107,33	1,25	2,75	2,33	91,28	15,66	19,72	2942,58
46	8,78	75,33	114,67	1,50	2,67	2,67	96,94	18,79	17,74	4153,55
47	6,93	63,83	106,00	1,16	2,36	3,65	124,38	17,11	20,39	854,48
48	9,44	65,94	111,67	1,00	2,67	3,11	77,78	16,58	21,72	3806,59
49	4,83	51,78	113,67	1,23	2,00	4,22	165,89	17,90	20,33	2450,46
50	4,25	51,08	108,33	1,75	1,50	4,22	109,39	19,05	19,34	2082,03
51	5,89	83,06	117,33	1,33	2,50	3,17	89,72	17,52	19,85	3924,64
52	7,17	75,83	104,67	1,25	2,75	5,33	124,08	16,47	20,39	3652,00
53	5,42	49,33	109,00	1,25	1,75	2,94	106,28	18,70	21,20	2291,54
54	7,58	60,25	106,33	1,00	2,75	1,58	69,00	13,87	19,70	3388,81
55	5,11	75,83	123,00	1,67	2,67	4,06	131,28	16,47	19,85	2539,95
56	9,89	83,94	108,33	1,67	3,00	4,06	102,56	17,77	20,49	2704,25
57	7,94	74,67	111,33	1,17	2,67	2,56	95,44	18,25	20,98	6105,13
58	7,06	74,11	109,00	1,00	3,17	2,67	76,83	17,73	18,88	1153,63
59	6,83	57,83	113,00	1,50	2,00	3,44	94,44	17,20	20,23	2355,45
60	6,56	74,94	123,67	1,83	2,17	3,00	102,17	18,57	20,50	972,38
61	6,93	61,83	117,00	1,41	2,11	3,40	115,05	17,11	20,59	1646,93
62	7,67	58,17	104,33	1,17	2,67	2,89	80,61	15,65	20,33	3451,80
63	7,92	83,17	114,67	1,50	2,25	2,08	89,08	15,73	20,09	4552,44
64	10,72	72,78	108,67	1,17	3,00	4,78	85,94	20,57	20,85	3309,71

65	6,42	77,17	110,00	1,25	3,25	3,78	118,22	16,08	18,89	2630,71
66	8,83	56,67	107,00	1,00	2,83	2,00	93,22	14,17	19,28	3343,13
67	8,56	80,61	110,67	1,17	2,67	4,28	85,78	16,16	20,78	2946,26
68	6,42	87,67	108,33	1,50	2,75	3,78	92,94	17,11	19,89	4883,08
69	5,42	63,25	104,00	1,75	2,75	4,17	114,42	16,98	21,77	3466,78

AIV: altura de inserção da primeira vagem; APM: altura da planta na maturidade; NDM: número de dias para a maturidade; AC: acamamento; VA: valor agrônomico; NR: número de ramos; NV: número de vagens; PCS: peso de cem sementes; TO: teor de óleo; PG: produtividade de grãos.

A eficiência das análises multivariadas tem sido demonstrada em trabalhos que abrangem áreas bem distintas dentro da genética e melhoramento de plantas. A exemplo disso temos o estudo da diversidade genética conforme Almeida, Peluzio e Afférri (2011), a discriminação de genótipos parcialmente resistentes à ferrugem asiática da soja conforme Santos et al., (2007), o estudo dos componentes da resistência à ferrugem-asiática em genótipos de soja conforme Koga et al., (2008) e a identificação e avaliação de novas características para fins de diferenciação de cultivares de soja conforme Nogueira et al., (2008).

4. CONCLUSÕES

As análises multivariadas contribuíram na seleção de genótipos superiores para ambos os experimentos, destacando 4 genótipos no experimento I e 14 genótipos no experimento II.

Os caracteres que permitiram selecionar genótipos superiores no experimento I foram: “valor agrônomico”, “produtividade de grãos”, “teor de óleo” e número de vagens” enquanto que no experimento II foram: “número de dias para maturidade”, “altura da planta na maturidade”, “valor agrônomico”, “altura de inserção da primeira vagem”, “acamamento” e “produtividade de grãos”.

5. REFERÊNCIAS

ALMEIDA, R.D.; PELUZIO, J.M.; AFFÉRI, F.S. divergência genética entre cultivares de soja, sob condições de várzea irrigada, no sul do estado Tocantins. **Revista Ciência Agronômica**, v.42, n.1, p.108-115, 2011.

CRUZ, C.D. Programa Genes: análise multivariada e simulação. Viçosa: UFV. p.175, 2006.

DALLASTRA, A.; UNÊDA-TREVISOLI, S. H.; FERRAUDO, A. S.; MAURO, A. O. Multivariate approach in the selection of superior soybean progeny which carry the RR gene. **Revista Ciência Agronômica**, v.45, n.3, p.588-597, 2014.

DUONG, T. ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R. **Journal of Statistical Software**, v. 21, n. 7, p. 1-16, 2007

DUONG, T. (2014). ks: Kernel smoothing. R package version 1.9.2. <http://CRAN.R-project.org/package=ks>

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Tecnologias de produção de Soja Região Central do Brasil 2012 e 2013**. Londrina, 2011. 262 p. Disponível em: <<http://www.cnpso.embrapa.br/download/SP15-VE.pdf>>. Acesso em: 08 maio 2014.

FEHR, W. R.; CAVINESS, J. A. **Stages of soybean development**. Ames: Iowa State University, 1977, 11p. (Special Report, 80).

FERRAUDO, A.S. **Técnicas de análise multivariada**: uma introdução. São Caetano: StatSoft South América, 2010.

Green, P.E.; and J. Douglas Carrol. **Mathematical Tools for Applied Multivariate Analysis**. New York: Academic Press, 1978

HAIR, J.F.; BLACK, W.C.; BABIN, B.J.; ANDERSON, R.E.; TATHAM, R.L. **Análise Multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009. 688 p.

KAISER, H.F. The varimax criterion for analytic rotation in factor analysis. **Psychometrika**, v.23, p. 187-200, 1958.

KOGA, L.J.; CANTERI, M. G.; CALVO, E. S.; UNFRIED, J. R.; GARCIA, A.; HARADA, A.; KIIHL, R. D. S. Análise multivariada dos componentes da resistência à ferrugem asiática em genótipos de soja. **Pesq. agropec. bras.**, Brasília, v.43, n.10, p.1277-1286, 2008.

MANLY, B.J.F. Análise de Agrupamentos. In: MANLY, B.J.F. **Métodos Multivariados: Uma Introdução**. 3. ed. Porto Alegre: Bookman, 2008. Cap. 9. p. 139-155.

NOGUEIRA, A. P. O.; SEDIYAMA, T.; CRUZ, C. D.; REIS, M. S.; PEREIRA, D. G.; JANGARELLI, M. Novas características para diferenciação de cultivares de soja pela análise discriminante. **Ciência Rural**, Santa Maria, v.38, n.9, p.2427-2433, 2008.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

SANTOS, J. A.; JULIATTI, F. C.; SANTOS, V. A.; POLIZEL, A. C.; JULIATTI, F. C.; HAMAWAKI, O. T. Caracteres epidemiológicos e uso da análise de agrupamento para resistência parcial à ferrugem da soja. **Pesq. agropec. bras.**, Brasília, v.42, n.3, p.443-447, 2007.

SEDIYAMA, T; TEIXEIRA, R.C.; REIS, M.S.. Melhoramento da Soja. In: BORÉM, A. (Ed.). **Melhoramento de Espécies Cultivadas**. 2. ed. Viçosa: UFV, 2005. Cap. 14. p. 562-572.

SEDIYAMA, T; TEIXEIRA, R.C.; BARROS, H.B. Cultivares. In: SEDIYAMA, T. (Ed.). **Tecnologias de produção e usos da soja**. 1. ed. Londrina: Mecenias, 2009. Cap. 8. p. 77-91

StatSoft, Inc. (2004). STATISTICA (data analysis software system), version 7. www.statsoft.com.

VIANNA, V.F.; UNÊDA-TREVISOLI, S.H.; DESIDÉRIO, J.A.; SANTIAGO, S.; CHARNAI, K., FERREIRA JUNIOR, J.A.; FERRAUDO, A.S.; MAURO, A.O. The multivariate approach and influence of characters in selecting superior soybean genotypes. **Afr. J. Agric. Res**, v. 8, p. 4162-4169, 2013.