



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS AGRÁRIAS
DEPARTAMENTO DE FITOTECNIA
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA/FITOTECNIA

TAMIRIS PEREIRA DA SILVA

**MODIFICAÇÃO DA METODOLOGIA DE ' $m + a$ ' E USO DO PEDIGREE PARA
SELEÇÃO PRECOCE DE PROGÊNIES DE SOJA**

FORTALEZA

2019

TAMIRIS PEREIRA DA SILVA

MODIFICAÇÃO DA METODOLOGIA DE ' $m + a$ ' E USO DO PEDIGREE PARA
SELEÇÃO PRECOCE DE PROGÊNIES DE SOJA

Tese apresentada ao Programa de Pós-Graduação em Agronomia/Fitotecnia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Doutora em Agronomia/Fitotecnia. Área de concentração: Genética e Melhoramento de Plantas.

Orientador: Prof. Dr. Júlio César do Vale Silva
Coorientador: Prof. Dr. Roberto Fritsche Neto

FORTALEZA

2019

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

S584m Silva, Tamiris Pereira da.
Modificação da metodologia de 'm + a' e uso do pedigree para seleção precoce de progênies de soja /
Tamiris Pereira da Silva. – 2019.
74 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências Agrárias, Programa de Pós-
Graduação em Agronomia (Fitotecnia), Fortaleza, 2019.

Orientação: Prof. Dr. Júlio César do Vale.

Coorientação: Prof. Dr. Roberto Fritsche Neto.

1. Glycine max [L.] Merrill. 2. Capacidade preditiva. 3. BLUEs. 4. BLUPs. 5. Matriz de parentesco. I.
Título.

CDD 630

TAMIRIS PEREIRA DA SILVA

MODIFICAÇÃO DA METODOLOGIA DE ' $m + a$ ' E USO DO PEDIGREE PARA
SELEÇÃO PRECOCE DE PROGÊNIES DE SOJA

Tese apresentada ao Programa de Pós-Graduação em Agronomia/Fitotecnia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Doutora em Agronomia/Fitotecnia. Área de concentração: Genética e Melhoramento de Plantas.

Aprovada em: 29 de agosto de 2019

BANCA EXAMINADORA

Prof. Dr. Júlio César do Vale Silva (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Roberto Fritsche Neto (Coorientador)
Universidade de São Paulo (USP-ESALQ)

Prof. Dra. Cândida Hermínia Campos de Magalhães Bertini
Universidade Federal do Ceará (UFC)

Pesq. Dr. Dheyne Silva Melo
Embrapa Agroindústria Tropical (CNPAT)

Pesq. Dra. Elaine Facco Celin
Universidade Federal do Ceará (UFC)

Aos meus avós Pedro, Solidade, Hilda e Manoel (*in memoriam*), meus agricultores mais queridos.

À Maria do Socorro, minha mãe, minha essência e motivo de tudo.

Ao meu pai Wilson Pereira (*in memoriam*), presente em minha vida por meio das boas lembranças deixadas entre os que com ele estiveram.

AGRADECIMENTOS

Pelo feliz motivo de nunca ter caminhado até aqui sozinha, não me permitirei nestas linhas ser breve...

À Deus, luz em meu caminho, que diante dos obstáculos me permitiu ser forte a ponto de não desistir em nenhum momento nesta árdua caminhada.

À minha mãe, Maria do Socorro, além de dedicar também agradeço. À pessoa mais especial de minha vida, por todo amor, dedicação, apoio e bons exemplos. Eu te amo!

Aos meus tios Santino, Zenaide, Zildene e Waverley, por tanto carinho ao fazer de mim a filha que nunca tiveram. Por sempre acreditarem na minha vitória.

À Ivone, Julio (Teddy), Beatriz e Vitória, por terem me recebido sempre de braços abertos em seu lar, que fiz de segunda casa em Fortaleza durante esses 6 anos. Não tenho palavras para descrever a importância de vocês para que eu concluísse esta etapa.

À toda minha família paterna, que a distância jamais impediu de que me dessem carinho e amor. Aos meus padrinhos Neide e Claudimir (Miro), meus exemplos de coragem.

À amiga Laianny (Sister), que esteve comigo durante a maior parte do período de pós-graduação, e nesta cidade dividiu comigo o lar, alegrias, tristezas, aventuras e desventuras.

À Joana Camila e Linda Brenna pelo abrigo nesta etapa final, pela companhia e pela amizade em períodos difíceis de escrita.

Aos meus amigos caririenses, que aqui não citarei todos os nomes, mas que para mim dão sentido a famosa frase “amigos são a família que Deus nos permitiu escolher”. Em especial a Débora, João Adolfo e Vanessa que se fizeram presentes sempre que precisei.

À toda equipe da INT Sementes, que contribuiu direta ou indiretamente para execução dos experimentos de campo. A tão importante companhia que puderam me oferecer também fora do ambiente de trabalho.

Ao pesquisador Lucas Silvério que abriu as portas da INT Sementes para que eu pudesse executar os ensaios experimentais.

Ao Jean Newton, pelo grande suporte durante os experimentos, contribuindo sempre de muita boa vontade não só com sua experiência com a cultura da soja, mas também com os esforços físicos para a realização deste trabalho. Pela sua amizade nesta jornada paranaense.

Aos amigos Álvaro, Carol, Jaqueline, Luan, Sâmia e Vânia, que foram minha família durante os “dois anos” a tantos quilômetros de casa, estando ao meu lado nos bons e maus momentos vividos em Campo Mourão/PR.

Ao professor Júlio César DoVale, pela paciência em orientar-me durante toda a pós-graduação, por acreditar em mim ao me impor desafios que me fizeram crescer não só como profissional, mas também como pessoa. Pelo exemplo de dedicação com a docência mesmo em meio as dificuldades que lhe são impostas.

Ao professor Roberto Fritsche-Neto, pela coorientação nas análises deste trabalho, estando sempre disponível para contribuir na melhoria dele.

Aos membros da banca avaliadora, professora Cândida Bertini e pesquisadores Elaine Celin e Dheyne Melo, pelas valiosas contribuições para melhoria do trabalho.

Aos professores e amigos Fábio Rossi e Silvério Paiva, que me iniciaram no instigante mundo da pesquisa científica.

Ao grupo de pesquisa em melhoramento vegetal, pelos árduos e felizes momentos em que trabalhamos, estudamos e juntos aprendemos desde o mestrado: Antônio Neto, Fernanda Carla, Ingrid, Jéssica, Kadson, Linda Brenna, Leane, Marcelo e Renata.

À Universidade Federal do Ceará, em especial ao programa de Pós-Graduação em Agronomia/Fitotecnia, pela oportunidade de realização do curso, assim como a todo corpo docente que contribuiu para a construção do meu conhecimento. À FUNCAP, pela concessão da bolsa de estudos.

Aos colegas e amigos que fiz na pós-graduação em Fitotecnia, para onde quer que eu vá vou sempre levar comigo todo companheirismo e carinho que de alguns recebi aqui.

Muito obrigada!

RESUMO

Elevados esforços são hoje aplicados em etapas do melhoramento da cultura da soja, que se encontra em constante expansão econômica no país e no mundo. Para tanto, ao se gerar progênies segregantes, é necessária custosa e morosa fase de aumento de homozigose para se identificar aquelas com potencial para extração de novas linhagens. A metodologia de $m + a$, que estima a concentração de alelos favoráveis em duas gerações consecutivas, consiste numa alternativa para isto, contudo, demanda de bastante recursos para execução. Uma modificação desta metodologia, que vise contornar esses “entraves”, pode ser bastante útil aos programas de melhoramento. Nesse contexto, objetivou-se propor uma modificação para esta metodologia a fim de torná-la mais facilmente exequível. Para obtenção dos dados fenotípicos foram realizados experimentos simultâneos, requeridos pela metodologia tradicional, e sequenciais, pela metodologia modificada, com três níveis consecutivos de homoziguidade: F_2 , F_3 e F_4 . Esses experimentos foram conduzidos nas localidades de Luiziana/PR e Campo Mourão/PR, em blocos casualizados com três repetições e como tratamentos as 20 progênies nos níveis supracitados de homoziguidade mais quatro testemunhas comerciais. Para a obtenção dos dados usados na metodologia tradicional, avaliaram-se as progênies nos três níveis de homoziguidade simultaneamente nas duas localidades mencionadas. Para gerar dados para modificação proposta da metodologia, instalaram-se experimentos sequenciais com esses níveis em épocas distintas, juntamente com as testemunhas em cada um deles. Dados a respeito do parentesco das progênies foram utilizados para montagem de matriz e criação de cenários preditivos. Os caracteres avaliados para ambas metodologias foram: dias para a floração (FLR); altura de planta (ALT); dias para a maturação (MAT); peso de 1000 sementes (PMS) e produtividade (PROD). Os dados foram analisados pelo *software R*. Com o uso de estimativas do melhor estimador não-viesado (BLUEs) das progênies nos níveis de homoziguidade foram obtidas correlações superiores entre estas ao longo das gerações F_2 e F_3 para maioria dos caracteres, com exceção do caráter PMS em ambas as metodologias. As correlações obtidas com as estimativas de $m + a$ foram mais baixas que aquelas observadas com os BLUEs em ambas as metodologias. Contrastes com correlações positivas ou negativas também coincidiram entre as metodologias. Estimativas de $m + a$ entre as gerações F_2 e F_3 para prever médias das progênies em F_4 não apresentaram correlações fortes o suficiente para se fazer inferências precisas a respeito desta última geração com base naquelas. Os cenários preditivos elaborados com a incorporação da matriz de parentesco para geração de BLUPs também não foram informativos para prever o comportamento das gerações. Houve semelhança entre os

resultados obtidos pela metodologia de $m + a$ tradicional e a proposta sugerida neste estudo com base nas magnitudes e direções (sinais) das estimativas geradas. No entanto, a metodologia tradicional gerou maiores capacidades preditivas para PROD e ALT, enquanto a modificada para MAT e PMS, que também foram as maiores correlações obtidas no geral. É muito provável que interações epistáticas negligenciadas na concepção do método tenham sido cruciais para a expressão de caracteres de alta complexidade e, em razão disto, as predições apresentaram baixa acurácia para as progênies. Por fim, constatou-se que, ao longo das gerações é melhor avaliar para selecionar progênies com base em médias ajustadas/BLUEs do que com estimativas de $m + a$, seja pela metodologia tradicional ou pela modificação proposta. Para resultados mais consistentes e complementares a respeito da metodologia, sugere-se o uso de dados oriundos de genotipagem para incorporação junto ao pedigree nas análises.

Palavras-chave: *Glycine max* [L.] Merrill. Capacidade preditiva. BLUEs. BLUPs. Matriz de parentesco.

ABSTRACT

High efforts are now being made at the stages of soybean crop improvement, which is constantly expanding in the country and in the world. Therefore, when generating segregating progenies, it is necessary a costly and time consuming phase of homozygosis increase to identify those with potential for extraction of new strains. The $m + a$ methodology, which estimates the concentration of favorable alleles in two consecutive generations, is an alternative for this, however, demands a lot of resources to perform. A modification of this methodology to circumvent these “barriers” can be very useful for breeding programs. In this context, the objective was to propose a modification to this methodology in order to make it more easily executable. To obtain the phenotypic data, simultaneous experiments were performed, required by the traditional methodology, and sequential, by the modified methodology, with three consecutive levels of homozygosity: F_2 , F_3 and F_4 . These experiments were carried out in the localities of Luiziana/PR and Campo Mourão/PR, in randomized blocks with three replications and as treatments the 20 progenies in the aforementioned homozygosity levels plus four commercial controls. To obtain the data used in the traditional methodology, progenies at the three homozygosity levels were evaluated simultaneously in the two mentioned locations. To generate data for the proposed modification of the methodology, sequential experiments with these levels were installed at different times, together with the checks in each one of them. Data regarding progenies kinship were used for matrix assembly and creation of predictive scenarios. The traits evaluated for both methodologies were: days to flowering (DF); plant height (PH); days to maturation (DM); 1000 seed weight (TSW) and yield (YLD). The data were analyzed by the R software. Using estimates of the best linear unbiased estimation (BLUEs) of progenies at homozygosity levels, superior correlations were obtained between F_2 and F_3 generations for most of the characters, except for PMS character in both methodologies. The correlations obtained with the estimates of $m + a$ were lower than those observed with the BLUEs in both methodologies. Contrasts with positive or negative correlations also coincided between the methodologies. Estimates of $m + a$ between F_2 and F_3 generations to predict progeny averages in F_4 did not show strong enough correlations to make accurate inferences about the latter generation based on those. The predictive scenarios elaborated with the incorporation of the kinship matrix for the generation of BLUPs were also not informative to predict the behavior of the generations. There was similarity between the results obtained by the traditional $m + a$ methodology and the proposal suggested in this study based on the magnitudes and directions

(signals) of the generated estimates. However, the traditional methodology generated higher predictive capacities for YLD and PH, while the modified one for DM and TSW, which were also the highest correlations obtained overall. It is very likely that neglected epistatic interactions in the design of the method were crucial for the expression of highly complex characters and, as a result, the predictions showed low accuracy for the progenies. Finally, it was found that over the generations it is better to evaluate to select progenies based on adjusted averages/BLUEs than with estimates of $m + a$, either by the traditional methodology or by the proposed modification. For more consistent and complementary results regarding the methodology, the use of genotyping data for pedigree incorporation in the analysis is suggested.

Keywords: Glycine max [L.] Merrill. Predictive capacity. BLUEs. BLUPs. Kinship matrix.

LISTA DE GRÁFICOS

- Gráfico 1 - Correlações de *Spearman* entre os BLUEs das gerações F₂, F₃ e F₄ nos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologias de *m + a* tradicional (1a) e modificada (1b) 51
- Gráfico 2 - Correlações de *Spearman* entre as estimativas de *m + a* (contrastes) das gerações consecutivas F₂/F₃ e F₃/F₄ dos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologias de *m + a* tradicional (2a) e modificada (2b) 54
- Gráfico 3 - Correlações de *Spearman* entre as estimativas de *m + a* das gerações F₂/F₃ e BLUEs médios (tradicional + modificada) da geração F₄ dos caracteres dias para floração (FLR), dias para maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD). Estimativa de *m + a* tradicional (3a) e modificada (3b). 57
- Gráfico 4 - Correlações de *Spearman* para os cenários propostos (BLUE F₂, BLUP F₃ e F₄; BLUE F₂ e F₃, BLUP F₄; BLUE F₂, F₃ e F₄; BLUP F₂, F₃ e F₄) dos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de *m + a*. 59
- Gráfico 5 - Correlações de *Spearman* para capacidade preditiva nos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologias de *m + a* tradicional e modificada. 60

LISTA DE TABELAS

Tabela 1 - Progênies segregantes de soja utilizadas em três níveis de homozigosidade (F ₂ , F ₃ e F ₄) para avaliação das metodologias tradicional e modificada.....	38
Tabela 2 - Testemunhas comerciais e seus respectivos grupos de maturação utilizadas nas metodologias tradicional e modificada	38
Tabela 3 - Características dos locais utilizados na condução dos experimentos	39
Tabela 4 - Caracteres avaliados a cada geração (F ₂ , F ₃ e F ₄) nas metodologias tradicional e modificada nos municípios de Campo Mourão/PR e Luiziana/PR.....	40
Tabela 5 - P-valores para os efeitos fixos de progênie e interação progênie x local e para os efeitos aleatórios de local e repetição/ambiente dos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ tradicional	47
Tabela 6 - P-valores para os efeitos fixos de progênie e interação progênie x local e para os efeitos aleatórios de local, repetição e testemunha/geração dos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ modificada	49

SUMÁRIO

1	INTRODUÇÃO GERAL	16
2	CAPÍTULO 1 – ASPECTOS GERAIS, ECONÔMICOS E RELACIONADOS AO MELHORAMENTO GENÉTICO DA SOJA – UMA REVISÃO DE LITERATURA	19
2.1	Aspectos gerais da soja (características botânicas, morfológicas e nutricionais)	19
2.2	Importância econômica	20
2.3	Origem e introdução da soja no Brasil	21
2.4	Melhoramento da soja no Brasil	23
2.5	Seleção precoce em autógamias	26
2.6	A metodologia de $m + a$ tradicional e uma proposta de modificação	30
3	CAPÍTULO 2 - PROPOSTA DE MODIFICAÇÃO DA METODOLOGIA DE “$m + a$” PARA SELEÇÃO PRECOCE DE PROGÊNIES DE SOJA ...	36
3.1	Introdução	36
3.2	Material e métodos	37
3.2.1	<i>Material genético</i>	37
3.2.2	<i>Condução experimental</i>	37
3.2.2.1	<i>Experimentos simultâneos da metodologia tradicional</i>	37
3.2.2.2	<i>Experimentos sequenciais da metodologia modificada</i>	39
3.2.3	<i>Mensuração de caracteres avaliados</i>	39
3.2.4	<i>Quadro de avaliações</i>	40
3.2.5	<i>Análises genético- estatísticas</i>	41
3.2.5.1	<i>Análises de deviance e estimativas dos componentes de variância</i>	41
3.2.5.2	<i>Obtenção das estimativas de $m + a$</i>	43
3.2.5.3	<i>Correlações dentro e entre os resultados obtidos em ambas as metodologias.</i>	44
3.2.5.4	<i>Cenários preditivos via BLUPs</i>	45
3.3	Resultados e discussão.....	46
3.3.1	<i>Análises de verossimilhança (ANADEVs) e parâmetros genéticos</i>	46
3.3.2	<i>Médias estimadas de valores fenotípicos via BLUEs e correlações</i>	50
3.3.3	<i>Estimativas da contribuição dos efeitos aditivos via $m + a$ e correlações</i>	53

3.3.4	<i>Capacidade preditiva via correlações entre estimativas de médias (BLUEs) e valores genotípicos (BLUPs)</i>	58
4	CONCLUSÃO	62
	REFERÊNCIAS	63
	APÊNDICE A – VALORES DE BLUES ESTIMADOS PARA AS 20 PROGÊNIES UTILIZADAS NAS TRÊS GERAÇÕES DE HOMOZIGOSIDADE AVALIADAS	71
	APÊNDICE B - ESTIMATIVAS DE $M + A$ PARA AS 20 PROGÊNIES UTILIZADAS NAS TRÊS GERAÇÕES DE HOMOZIGOSIDADE AVALIADAS	73
	APÊNDICE C – CORRELAÇÕES DE <i>SPEARMAN</i> ENTRE A TRÊS GERAÇÕES DE HOMOZIGOSIDADE AVALIADAS	75
	APÊNDICE D – CENÁRIOS PREDITIVOS	76

1 INTRODUÇÃO GERAL

A soja é uma cultura que ocupa lugar de destaque no agronegócio mundial, tendo os Estados Unidos como seu principal produtor. No Brasil, esta oleaginosa trata-se da cultura anual mais importante, colocando o país como o segundo maior produtor e exportador do grão (SZNITOWSKI, 2017). No entanto, este não foi sempre o desempenho brasileiro na produção deste *commodity*, pois, após ser introduzida no país no século XIX, só obteve destaque produtivo e comercial por volta de 1970 (GAZZONI; DALL'AGNOL, 2018).

Hoje, graças ao melhoramento genético, a soja é cultivada do Rio Grande do Sul à Roraima, com inúmeras cultivares adaptadas aos mais diversos ambientes, incrementando positivamente a cada ano o cenário agrícola do país. Com o passar dos anos, o surgimento de novas exigências quanto à qualidade do grão, fizeram com que os melhoristas desta cultura trabalhassem exaustivamente na obtenção de novas cultivares (CERA, 2016; OLIVEIRA, 2017). Para isso, foram necessários a criação e aplicação de técnicas que permitiram gerar novas cultivares de forma mais rápida e com o menor dispêndio físico e financeiro.

A biologia molecular, principalmente com técnicas de engenharia genética, foi uma das áreas do conhecimento que mais auxiliaram a incrementar resultados satisfatórios à pesquisa sojícola. No entanto, o melhoramento genético clássico ainda é essencial nesse processo, de forma a ser indispensável e complementar para o progresso genético da cultura. Nesse sentido, a seleção precoce vem ganhando cada vez mais espaço nos programas de melhoramento, por permitir a eliminação de genótipos indesejáveis logo nas primeiras gerações segregantes (ALMEIDA *et al.*, 1999; CARVALHO, 2010). Desta forma, pode-se concentrar recursos físicos, financeiros e humanos em progênies que apresentam elevado potencial genético para extração de novas linhagens.

A estimativa de $m + a$ consiste em uma metodologia para auxiliar no processo de seleção precoce, pois permite identificar progênies com maior potencial genético por meio de contrastes entre níveis consecutivos de homozigose, por exemplo, $m + a = 2F_3 - F_2$ (VENCOVSKY, 1987). O contraste obtido ($m + a$) exprime a contribuição dos *loci* fixados por homozigose nos genótipos, presentes em cada progênie a partir dos seus desempenhos *per se* e foi desenvolvido a partir de teorias já estabelecidas por Mather (1949) e Mather e Jinks (1982). Para obtenção dos dados necessários para se fazer esse contraste, a avaliação das progênies para os caracteres de interesse deve ser realizada de forma simultânea. Dessa forma, *a priori*, são

obtidas sementes de progênies em níveis diferenciados de homoziguidade, para a *posteriori*, constituírem o futuro experimento a ser avaliado de uma única vez. Isso requer um tempo considerável para obtenção das sementes e também alta demanda por mão de obra e área experimental uniforme, o que torna a metodologia pouco atrativa aos melhoristas, mesmo tendo sido obtidos resultados satisfatórios com o seu emprego (MATOS, 2016; RIBEIRO; TOLEDO; RAMALHO, 2009; ROCHA *et al.*, 2015).

Vários métodos usados na condução de progênies segregantes apresentam modificações que os tornam mais eficientes quanto ao uso de insumos necessários à sua execução, trazendo resultados semelhantes ou até superiores aos originais/tradicionais, a exemplo do *single pod descent (SPD)* como modificação do *single seed descent (SSD)* (TOLEDO *et al.*, 2009). Neste estudo, propõe-se uma modificação na metodologia tradicional da estimativa de $m + a$, que torna o processo de experimentação mais facilmente exequível. Para isto, avaliam-se as progênies na medida em que se obtém as sementes para cada nível de homoziguidade, ou seja, as avaliações são realizadas de forma sequencial/gradativa. Para ser possível corrigir o efeito diferenciado do ambiente ao longo dos experimentos com os níveis distintos, propõem-se o uso de testemunhas comuns como cultivares do tipo linhagens, com alto grau em homoziguidade. Utilizando o mesmo princípio de blocos aumentados de Federer (1955), essas testemunhas em comum possibilitarão identificar as variações ambientais ao longo dos experimentos e ajustar as médias fenotípicas das progênies em análise. Neste sentido, os esforços e insumos necessários para obtenção dos dados são diluídos ao longo dos experimentos (diferentes épocas).

Em meio a essas dificuldades, o próprio processo experimental está propenso a desbalanceamentos por erros ou perda de parcelas e dados (RESENDE, 2004). No entanto, de forma a contornar muitos desses problemas e obter contrastes de $m + a$ com melhor qualidade o uso de procedimentos estatísticos adequados é de suma importância nesse processo. Assim, acredita-se que os melhores estimadores (BLUEs) e preditores (BLUPs) não-viesados obtidos por modelos mistos geram estimativas mais fidedignas a essa investigação para se inferir a respeito do potencial das progênies, seja pela metodologia em sua forma tradicional ou pela modificação proposta.

Com base no exposto, objetivou-se com esse estudo propor uma modificação para a metodologia de $m + a$; comparar os resultados gerados aos obtidos pela metodologia

tradicional e verificar a eficiência do uso adicional do *pedigree* para prever o comportamento das gerações envolvidas na cultura da soja.

2 CAPÍTULO 1 – ASPECTOS GERAIS, ECONÔMICOS E RELACIONADOS AO MELHORAMENTO GENÉTICO DA SOJA – UMA REVISÃO DE LITERATURA

2.1 Aspectos gerais da soja (características botânicas, morfológicas e nutricionais)

A soja cultivada (*Glycine max* [L.]Merril) é uma leguminosa, pertencente à família Fabaceae, subfamília Faboideae, gênero *Glycine*, espécie *Glycine max* (SEDIYAMA, 2009). De acordo com Sedyama, Silva e Borém (2015), a soja é uma planta de cultivo anual, com ciclo de vida que varia entre 70 e 200 dias, germinação epígea e dicotiledônea. Possui caule herbáceo, ramificado, com direção de crescimento ortótopo cujo hábito de crescimento pode variar entre os tipos determinado, semideterminado e indeterminado e está associado aos tipos de porte ereto, semi-ereto ou prostrado. É uma planta de dias curtos, autógama, com inflorescências de cores brancas ou roxas. Seu fruto é uma vagem, que pode conter de uma a cinco sementes, sendo mais comum apresentar duas ou três sementes.

Quanto às exigências edafoclimáticas, além de dias curtos, se adapta melhor em ambientes com temperaturas que variam entre 20 e 30 °C. Por ser uma espécie que apresenta metabolismo C₃, a luz exerce relevante importância, principalmente sobre a emissão dos primórdios reprodutivos e no pegamento de vagens e grãos, influenciando também a expansão foliar, estatura da planta (alongação da haste principal) e o ciclo (ALAMBERT, 2010; MEOTTI *et al.*, 2012).

O fotoperíodo é um dos fatores que mais afeta o crescimento e desenvolvimento da soja, influenciando a estabilidade da cultura em ambientes diversos (AMORIM *et al.*, 2011). Em razão disto, foram criados 13 grupos de maturidade relativa, que consiste em predizer faixas de latitude onde uma determinada cultivar detém maior adaptação e estabilidade produtiva. No Brasil, cinco desses grupos estão presentes, em faixas que se distribuem do Rio Grande do Sul a regiões próximas à linha do Equador (SEDIYAMA; SILVA; BORÉM, 2015).

Em relação ao ciclo da cultura, o Brasil possui cultivares que maturam entre 90 (superprecoces) e 150 dias (tardias), variando de acordo com o local e a época de semeadura. As cultivares mais precoces geralmente são menos produtivas, mas apresentam a vantagem de possibilitar uma segunda safra na mesma área (SEDIYAMA, 2009).

No seu aspecto nutricional, o grão de soja possui elevado teor proteico, em média de 40,3%, sendo encontrado também nesses grãos 21,0% de óleo, 33,9% de carboidrato e 4,9%

de cinzas na base seca (ERICKSON, 1995). Esse elevado percentual de proteína é considerado ainda de alta qualidade, fazendo parte da complementação das mais variadas dietas, principalmente em substituição a proteína animal. No entanto, a soja ainda não é conhecida mundialmente como alimento básico para alimentação humana (SOUSA, 2015; ROSA FILHO *et al.*, 2009). Na alimentação animal, a complementação proteica também é bastante suprida por este grão, fazendo com que o farelo de soja seja o subproduto mais consumido por animais (PORTO *et al.*, 2009).

Recentemente, recorrentes pesquisas apontam a soja como fonte de isoflavonas, que são compostos bioativos com atividades biológicas benéficas de efeito estrogênico e anticancerígeno (FRIZON *et al.*, 2015).

2.2 Importância econômica

Nas últimas décadas, a produção mundial de soja é uma das atividades agrícolas que mais têm crescido. A demanda por essa crescente produção tem sido alavancada pelo comércio dos mais variados produtos do complexo soja, que é impulsionado principalmente pela produção de produtos de origem animal, assim como também pela expansão da exploração do grão nas mais diversas regiões do mundo (HIRAKURI; LAZZAROTTO, 2011).

A qualidade nutricional e organoléptica fazem dessa cultura uma importante matéria prima, com crescente expansão na produção de grãos e área cultivada também no Brasil. Atualmente, a soja detém 57% da área cultivada com grãos no país, com rendimento médio de 3394 kg ha⁻¹. Na safra 2017/2018 o Brasil produziu 119,282 milhões de toneladas do grão e exportou o equivalente a cerca de 70% desse total (CONAB, 2019). Essa produção coloca o Brasil atrás apenas dos Estados Unidos (120,07 milhões de toneladas na safra 2017/2018) no *ranking* dos maiores produtores. Esses dois países, juntamente com a Argentina, são responsáveis por cerca de 82% da soja produzida mundialmente (USDA, 2019).

Com a colheita finalizada nos principais estados produtores do Brasil, estima-se que a safra 2018/2019 seja de 115,018 milhões de toneladas. Os maiores produtores são os estados do Mato Grosso, Paraná, Rio Grande do Sul e Goiás, que produziram o equivalente a 27, 16, 14 e 10%, respectivamente, da produção nacional em 2018 (CONAB, 2019). O destaque para o principal centro de produção, a região Centro-Oeste, é associado às boas condições oferecidas, tais como: infra-estrutura, principalmente nas vias de acesso, comunicação e urbanização; estabelecimento de agroindústrias produtoras e processadoras de grãos e de carne; clima

favorável; extensas áreas cultivadas e elevado nível tecnológico, com modernas técnicas de produção (CAMPOS, 2011). De acordo com a projeção do MAPA (2019a), estima-se que em 10 anos (2028/2029) a produção nacional de soja será de 151,873 milhões de toneladas, o que representa um acréscimo de 32% em relação à estimativa para esse ano (2018/2019). Esse crescimento se dará principalmente em áreas de cerrado que compreendem os estados do Maranhão, Tocantins, Piauí e Bahia, conhecida atualmente por MATOPIBA, e que representa uma área de grande potencial produtivo sojícola brasileiro, também citada como “última fronteira agrícola do país”. A perspectiva é que, nos próximos anos, os demais maiores países produtores não terão a possibilidade de expandir a área de produção como no Brasil, por causa da acentuada exploração de outras culturas e da atividade pecuária.

Grande parte da soja produzida é utilizada como farelo e óleo vegetal, considerados subprodutos da cadeia. A demanda por farelo está diretamente ligada à demanda de carnes, já que o uso deste como base da ração animal barateia os custos de produção. Assim, a produção de carnes no país cresce de acordo com a oferta do farelo (SEDIYAMA; SILVA; BORÉM, 2015). Em 2018, foram produzidos 33,185 milhões de toneladas de farelo de soja (76% da soja processada no país), sendo cerca de 51% destinado à exportação (ABIOVE, 2019).

A soja é a principal matéria prima para extração de óleo vegetal no mundo. No Brasil, entre os subprodutos da soja processada, cerca de 19% correspondem a óleo (8,833 milhões de toneladas de óleo em 2018). Essa produção é impulsionada essencialmente pela produção de biodiesel, um dos principais componentes da matriz energética do país, que tem o óleo de soja como a matéria-prima mais utilizada, seguido da gordura animal e do óleo de algodão (ABIOVE, 2019). Projeta-se que nos próximos anos o aumento do consumo interno de soja será o principal fator a impulsionar a produção de seu óleo (MAPA, 2019a).

2.3 Origem e introdução da soja no Brasil

A soja que hoje é cultivada teve ancestral originado no Leste da Ásia, com registro mais antigo datado de 2.838 a.C. Porém, a região asiática que se refere ao real centro de origem não é citada com precisão, havendo dessa forma discordância entre diversos autores (BONATO; BONATO, 1987). Alguns deles, no entanto, ao fazer estudo para divisão dos centros de diversidade por meio de comparações da origem geográfica de inúmeros acessos asiáticos, forneceram fortes evidências que o *pool* gênico da soja da Coreia do Sul e Japão

foram derivados de introduções da China. Coréia de Sul e Japão possuem acessos geneticamente similares entre eles e distintos de acessos chineses (LI; NELSON, 2001).

Permanecendo por milhares de anos pelo Oriente, a soja foi introduzida no Ocidente somente ao final do século XV e cultivada na Europa pela primeira vez em 1739, em Paris, como curiosidade, em jardins botânicos ali localizados (EMBRAPA, 2017; SEDIYAMA *et al.*, 1985).

Por volta do final do século XVIII e início do século XIX a soja chegou aos Estados Unidos, no estado da Geórgia, cidade de Savannah, com seus primeiros relatos de comportamento datados em 1804. No entanto, cultivos comerciais só foram iniciados por volta de 1920, primeiramente como forrageira e, posteriormente, como grão. A partir desse período, seu cultivo foi expandido mundialmente (COSTA; MANICA, 1996; SILVA *et al.*, 2017).

A introdução de soja no Brasil se deu em 1882, por Gustavo D'utra, no estado da Bahia. Nessa época, a soja despertava interesses apenas como planta de uso forrageiro. Como as sementes eram originárias dos Estados Unidos, não apresentaram boa adaptação, encontrando tal qualidade somente após ser introduzida em São Paulo por imigrantes japoneses em 1892, com a finalidade de gerar grãos para alimentação (BONATO; BONATO, 1987; SILVA *et al.*, 2017).

Por ser originária do continente asiático, a soja até então cultivada era adaptada apenas a locais com altas latitudes, o que facilitou o cultivo ao chegar oficialmente em 1914 no estado do Rio Grande do Sul, mais precisamente no município de Santa Rosa, introduzida por E.C. Craig, professor da Universidade Federal do Rio Grande do Sul, para estudos iniciais (BONATO; BONATO, 1987; MIYASAKA; MEDINA, 1981). Nessa localidade, a cultura encontrou condições edafoclimáticas favoráveis, semelhante as da região de origem, e foi cultivada com a finalidade de alimento para suínos (SANTOS, 1988). No entanto, mesmo com a chegada oficial datada em 1914, há relatos de plantios no estado anteriores a 1901 no município de Dom Pedrito, pelo agrônomo A.Welhauser, o que faz com que alguns historiadores ponham em dúvida que a primeira introdução da espécie tenha sido no estado da Bahia (MIYASAKA; MEDINA, 1981). Após o sucesso inicial da soja em todo o estado do Rio Grande do Sul, pesquisas puderam ser desenvolvidas com a cultura, assim como também o avanço para o Norte do país.

2.4 Melhoramento da soja no Brasil

A grande expansão da soja no Brasil até os dias de hoje é resultado principalmente dos estudos advindos do melhoramento genético para a cultura, realizado por programas de melhoramento instalados em diversas partes do país. Experimentos iniciais ocorreram na Estação Agronômica de Campinas, atual Instituto Agronômico de Campinas (IAC), em 1892 (BONATO; BONATO, 1987; ROMANATO, 2013). Esses pequenos ensaios foram o “pontapé” inicial para futuros trabalhos de melhoramento em soja no estado em 1921 (MIYASAKA; MEDINA, 1981).

De acordo com Bizari (2014), somente em 1930 iniciou-se oficialmente o primeiro programa de melhoramento de soja, na antiga Estação Experimental Fitotécnica das Colônias, no município de Veranópolis, Rio Grande do Sul. O programa possuía uma pequena coleção de cultivares americanas, onde eram avaliadas caracteres botânicos e agronômicos a partir de ensaios de competição.

Em 1946, o programa de melhoramento foi fortalecido com pesquisas em outras estações experimentais e convênios com instituições como o IAC e a Universidade Federal de Viçosa. Para obtenção das primeiras cultivares, foram iniciadas hibridações artificiais por volta de 1947. Segundo Miyasaka e Medina (1981) a primeira cultivar de soja foi lançada pela Secretaria de Agricultura do Rio Grande do Sul, em 1960, com o nome de Pioneira. Os resultados dessas pesquisas permitiram que as cultivares americanas fossem gradativamente substituídas por brasileiras. Isso resultou em incrementos na produtividade da cultura no Rio Grande do Sul e na expansão para regiões de menor latitude no país, onde as condições edafoclimáticas são diferentes da região Sul (BORÉM, 2013).

Também de grande expressão no Sul do país, principalmente nos dias atuais, o estado do Paraná deu início, em 1965, aos trabalhos de melhoramento da cultura, que teve como maior impulso a facilidade da comercialização da soja que era introduzida no estado para utilização como adubação verde em cafezais. Ao final da década de 1960 foi criado no Paraná a primeira empresa privada para pesquisa de melhoramento em soja do Brasil, FT – Pesquisa e Sementes na cidade de Ponta Grossa. Foi, junto a esse programa o desenvolvimento da conhecida cultivar Cristalina, recordista em produção, responsável por cerca de 52% da produção da soja no Brasil na safra 1993/94 (SILVA *et al.*, 2017). Em 1975, foi criado no estado o Centro Nacional de Pesquisa da Soja (CNPSO) da EMBRAPA, que ganhou dimensão nacional

e desenvolveu cultivares para diversos estados, além dos da região Sul do país (CUNHA; ESPINDOLA, 2016).

A região Sudeste também teve contribuições bastante expressivas no melhoramento da cultura da soja. No atual IAC foram realizados os experimentos iniciais da pesquisa em melhoramento de soja neste estado em 1921, que resultou nas cultivares Jomichel, Julieta e Joalo, muito precoces, com ciclo de 90 dias e estrutura de plantas muito baixas (MIYASAKA; MEDINA, 1981). Em 1957 foram identificadas cultivares sensíveis à dias curtos, o que se tornou o início da busca por cultivares mais bem adaptadas à latitudes próximas à linha do Equador. Logo depois, o IAC iniciou pesquisas para cultivares de período juvenil longo para indução floral (SILVA *et al.*, 2017). Entre outros programas criados para o melhoramento da soja no estado de São Paulo, destaca-se também o da Escola Superior de Agricultura Luiz de Queiroz (ESALQ), que desenvolve principalmente cultivares adequadas a alimentação humana.

De acordo com Zito *et al.* (2011), o início do melhoramento da cultura em Minas Gerais se deu por volta de 1960 pela Universidade Federal de Viçosa, tendo em 1969 o lançamento de suas primeiras cultivares. Em 1973 foi lançada a variedade UFV-1, derivada da cultivar Viçoja, que foi amplamente cultivada nas décadas de 70 e 80 em áreas de cerrado de Minas Gerais, Goiás, Mato Grosso do Sul e Mato Grosso.

O cultivo de soja no país foi impulsionado entre as décadas de 70 e 90 por pesquisas desenvolvidas em convênios entre órgãos privados como o Instituto Privado de Fomento à Soja – INTISOJA e instituições públicas, como o já citado IAC, universidades, a Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA, entre outros órgãos estaduais de pesquisa agropecuária (BORÉM, 2013; MIYASAKA; MEDINA, 1981;).

Na década de 70, a soja já era cultivada em regiões de menor latitude como as do cerrado brasileiro, por meio do desenvolvimento de cultivares mais adaptadas, a partir da introdução do gene de período juvenil longo, que atrasava o florescimento da espécie, mesmo em fotoperíodo indutor. Atualmente, para todas as faixas de latitude do país, existem cultivares adaptadas (OLIVEIRA, 2015; SEDIYAMA; SILVA; BORÉM, 2015).

A região Centro-Oeste, hoje maior produtora de soja do país, teve os primeiros experimentos com soja a partir de 1950, em Goiás, e apenas em 1973 o melhoramento da cultura foi fortalecido no estado, com início do primeiro programa, que era fruto de parcerias entre instituições públicas e privadas. Já nessa época foram encontradas grandes possibilidades de

desenvolvimento de pesquisa e cultivo comercial, já que o estado não apresentava problemas climáticos com geadas, possibilitando assim dois cultivos anuais. Diversas variedades foram desenvolvidas nesse programa não só para cultivo na região Centro-Oeste como também para localidades cada vez mais próximas da linha do Equador, como os estados nordestinos do Maranhão e Piauí (SPEHAR, 1994).

Hoje, não só os programas sediados no Centro-Oeste e em regiões de cerrado, mas também consolidados programas de melhoramento em todo o país, desenvolvem cultivares para região Nordeste do Brasil. Nessa região, as pesquisas com soja foram iniciadas em 1968, em Irecê, na Bahia, e em 1989 foi instalada uma base física da Embrapa-CNPSo com campo experimental para apoio ao seu programa de melhoramento em Balsas no Maranhão. (CARVALHO, 2002; EMBRAPA, 2016). A partir da metade da década de 1990, com elevação da atividade agrícola nordestina cresce o interesse de diversas empresas privadas em indicar cultivares para essa região. Anos mais tarde, parte dos estados do Nordeste junto ao Tocantins, estado do Norte do país, se tornam a região MATOPIBA, impulsionando assim novas pesquisas para as condições edafoclimáticas dessa localidade (ARAÚJO, 2018).

No Brasil, foram registradas 2114 cultivares de soja até julho de 2019 junto ao Ministério da Agricultura, Pecuária e Abastecimento (MAPA, 2019b). Atualmente, para escolha de uma cultivar de soja, o produtor observa os principais pontos a serem atendidos: região de indicação, grupo de maturidade, tipo de crescimento, ciclo médio, altura de planta, resistência ao acamamento, época de semeadura e reações a doenças (OLIVEIRA, 2015; SEDIYAMA; SILVA; BORÉM, 2015).

Além da introdução de germoplasma de outros países ou até mesmo de uma região para outra buscando adaptação dentro do próprio país, várias fases que normalmente são recomendados a espécies de reprodução autógama, são necessárias a se praticar no melhoramento genético da soja. Almeida *et al.* (1999) descreve que em fases iniciais há a necessidade do desenvolvimento de progênies segregantes, por hibridações artificiais, a fim de se aumentar a variabilidade genética e permitir recombinações gênicas. Em seguida, essas progênies são avançadas por sucessivas gerações para se obter um certo grau de homozigose genética. Geralmente, somente em gerações mais avançadas e uniformes são realizadas seleções de plantas com fenótipos desejáveis. Ensaio finais de produtividade e estabilidade são conduzidos com grande número de linhagens e repetidos em vários ambientes e anos/safras com o intuito de se observar a interação genótipos x ambientes.

Embora todo esse processo seja imensamente lento e oneroso, dificultando a fase clássica do melhoramento da cultura, o mesmo é necessário. Para avanço de gerações, fase que dispense mais tempo, os métodos mais comumente utilizados concentram-se em SSD (*single seed descent*) e *bulk*. Entretanto, modificações como o SPD (*single seed descente*) e até mesmo combinações entre métodos e procedimentos são usados como alternativas para se contornar entraves como este (ALMEIDA *et al.*, 1999).

Hoje, o melhoramento da soja é desenvolvido não só por meio de metodologias clássicas, mas em associação com técnicas advindas da biotecnologia, possibilitando o aumento da probabilidade de se obter genótipos superiores com fenótipos de interesse de cada programa (LEITE; MUNHOZ, 2013). Segundo Borém e Almeida (2011), entre os diversos métodos utilizados, os de maior sucesso são o da transformação mediada por *Agrobacterium* e a transformação direta por biobalística.

Com o auxílio dos avanços biotecnológicos, em 1998 foi lançada a primeira soja geneticamente modificada com evento transgênico, a soja RR (*Roundup Ready*). Obtida pela empresa Monsanto, essa é tolerante ao herbicida glifosato e permite o controle de plantas daninhas na cultura. Outras cultivares transgênicas já foram desenvolvidas com o caráter de tolerância a herbicidas de diferentes grupos químicos como também a insetos (BORÉM; ALMEIDA, 2011; MATEUS; SILVA, 2013).

O melhoramento de soja no Brasil hoje avança para o desenvolvimento de cultivares cada vez mais produtivas, com altos teores de óleo, proteína de qualidade, adaptadas as mais diferentes condições ambientais, principalmente em regiões com restrição hídrica. Pesquisas recentes também estão sendo direcionadas para o desenvolvimento de novas cultivares que sejam mais atrativas ao consumo humano, desde a palatabilidade e características físicas pós cozimento, assim como também para que sejam inseridas melhorias na composição nutricional do grão, que possui componentes benéficos a saúde como isoflavonas e Ômega 3 e 6 (CIABOTTI *et al.*, 2016).

2.5 Seleção precoce em autógamas

Como citado anteriormente, as seleções de plantas com fenótipos desejáveis são realizadas em gerações mais avançadas e uniformes quanto a segregação em determinados caracteres. Isso ocorre principalmente porque em diversos caracteres como produção de grãos,

o controle genético é mais complexo, por ser condicionado por elevada quantidade de genes sofrendo forte influência ambiental. Esses caracteres, são de herança quantitativa, geralmente de difícil avaliação em campo, e apresentam baixa herdabilidade (ALLARD, 1971). Sendo assim, a seleção neste nível deve ser postergada para gerações mais avançadas.

A seleção precoce é possível pela aplicação de metodologias de seleção em gerações iniciais, sendo um trabalho bastante dificultoso e passível de erros, já que a maioria dos caracteres de interesse agrônômico e econômico apresenta caráter quantitativo. No entanto, isso não impede a eficiência da seleção para caracteres governados por herança poligênica (CRUZIO, 2014). O que se sabe de fato é que em gerações iniciais, ao contrário dos caracteres mais complexos, os de alta herdabilidade podem ser facilmente selecionados por critérios ou parâmetros visuais, já que são representados por poucas classes fenotípicas. Para aumento de sucesso na seleção é necessário então que se trabalhe com a maior quantidade possível de indivíduos (desde que não permita redução na qualidade experimental e avaliativa). Em autógamias, durante os ciclos de avanço de gerações, genótipos com potencial superior podem ser reduzidos por perda em cada população, principalmente se esses tratar-se de caracteres de baixa herdabilidade (ROSAL *et al.*, 2000).

Para essa finalidade, geralmente são envolvidas na seleção a avaliação de progênies ou famílias derivadas da geração F₂ ou F₃ em experimentos com repetição (CRUZIO, 2014). Em autógamias, com base nessa seleção, os indivíduos e progênies considerados inferiores são descartados. Dessa forma, recursos e mão de obra são canalizados a partir de então para caracteres que só poderiam ser selecionados em gerações mais avançadas ou com maior potencial para produção de linhagens superiores (DONÇA, 2012).

O método genealógico, também conhecido como *pedigree*, consiste na seleção artificial e fenotípica em plantas a partir da geração F₂. Nessa, as melhores plantas, por meio de seleção visual, são colhidas para obtenção de famílias F_{2:3}, em que da mesma forma, são selecionadas plantas para se gerar famílias F_{3:4} e assim por diante até que seja alcançada a homozigose para a maioria dos *loci* e constituídas as linhagens para serem avaliadas em experimentos com repetições e comparações com testemunhas comerciais (RAMALHO; FERREIRA; OLIVEIRA, 2012).

Ao longo das gerações a variabilidade vai sendo reduzida dentro das famílias e aumentada entre famílias, já que dentro da família o grau de homoziguidade vai sendo

incrementado. Geralmente, por volta da geração F_4 é cessada a seleção, por conta do alcance da homozigose e iniciada as avaliações de progênies (SILVA, 2009). Esse método permite o descarte de indivíduos inferiores em gerações precoces e a utilização dos dados obtidos para estudos genéticos. No entanto, se torna bastante moroso, pois só permite a condução de uma única geração por ano, exige elevada demanda de mão de obra, campo experimental e requer equipe qualificada para que visualmente sejam selecionados os tipos desejáveis de plantas entre e dentro de famílias a cada geração (BORÉM; MIRANDA, 2013). Segundo Ramalho, Abreu e Santos (2001), tamanho trabalho, geralmente inviabiliza a utilização de muitas populações a cada experimento, além disso, a eficiência do método é posta em dúvida, principalmente quando caracteres que envolvam muitos genes são avaliados, já que a seleção é unicamente visual. Dessa forma, o método genealógico, ao menos até a geração F_4 , se torna ineficiente para caracteres de baixa herdabilidade, mas para caracteres de alta herdabilidade há maior previsibilidade de resultados, se seleção.

Outro método que permite seleção em gerações iniciais é o método populacional, mais conhecido como *bulk*, permitindo avançar várias populações independentemente da quantidade de indivíduos ao mesmo tempo. Na descrição do método, já em F_2 , as plantas são colhidas em conjunto, as sementes misturadas e então uma amostra é retirada para obtenção da geração F_3 e assim por diante até se atingir a homozigose para a maioria dos *loci* (RAMALHO *et al.*, 2001). Neste momento, chamado de “abertura do *bulk*”, progênies são obtidas com a escolha de plantas individuais, que serão avaliadas e selecionadas de acordo com os caracteres em estudo, com repetições até a obtenção de linhas puras (CARVALHO, 2008). Na teoria do método não há seleção até que a maioria dos *loci* estejam em homozigose. No entanto, na prática, caracteres de alta herdabilidade são comumente selecionados (BUENO; MENDES; CARVALHO, 2001).

De acordo com a espécie em questão, a abertura do *bulk* para seleção precoce pode já ser satisfatória na geração F_4 (SILVA, 2009). É vantagem desse método a condução de grande número de populações com maior facilidade. Ademais, as populações constituídas por indivíduos mais adaptados e competitivos são as que farão parte das gerações seguintes com a maior proporção de indivíduos. Uma das desvantagens é que há impossibilidade de condução em casa de vegetação, em ambiente divergentes a cada geração e até mesmo a execução de mais de uma geração por ano, ou seja, as gerações devem ser conduzidas em condições ambientais semelhantes à de início da seleção (BORÉM; MIRANDA, 2013). Problemas de amostragem também são considerados uma desvantagem do método, pois ao longo dos avanços de gerações

podem ocorrer perdas de combinações genotípicas por conta de deficiência na amostra, principalmente nas gerações mais avançadas (SILVA, 2009).

Frey, em 1954, propôs a modificação do método *bulk* para *bulk* dentro de progênies, que tem como propósito a colheita de plantas individuais em gerações iniciais em que derivará famílias que serão semeadas em linhas. Em cada família, as sementes são colhidas em *bulk* e utilizadas para se obter novas famílias. Em gerações posteriores são selecionadas plantas superiores dentro de cada família, onde serão abertas linhas na geração seguinte. A partir dessa etapa será feito o processo tradicional para avaliação de linhagens. De acordo com Carvalho (2008), esse método nada mais é que uma otimização do método genealógico com a junção do uso de *bulks*, em que há, dessa forma, maior representatividade da amostra em relação ao *bulk* tradicional, assim como também abertura das progênies em gerações precoces.

No *bulk* dentro de progênies os ciclos avaliados fazem com que a seleção seja feita com base no desempenho médio ao longo dos ciclos de avaliação, de forma a atenuar os efeitos da interação genótipo x ambiente. Como desvantagem, pode-se apontar que mesmo com maior representatividade da amostra inicial ainda há grandes perdas de genótipos que poderiam ser favoráveis a seleção em gerações avançadas, principalmente para caracteres de baixa herdabilidade (RAMALHO *et al.*, 1993).

A fim de se fazer seleção precoce e descartar progênies inferiores dentro de uma mesma população em gerações iniciais, foi proposto o teste de geração precoce por Immer (1941). Nesse método, as progênies eliminadas são pouco promissoras e não possuem potencial genético para dar origem a linhagens superiores, ao contrário das que serão avançadas, que possibilitarão aumentar a porcentagem de linhagens advindas de famílias com o caráter em questão, favorável. A cada geração avançada com seleção são formadas novas famílias que são inseridas em ensaios para se testar o caráter em questão. As famílias desta nova geração, oriunda de teste de competição, são selecionadas e novamente cultivadas da mesma forma até o momento de serem colhidas em *bulk* e destinadas a ensaios preliminares e finais (FEHR, 1978).

O propósito principal da metodologia desse teste é que desde gerações iniciais já sejam descartadas famílias inferiores, com desempenho pouco satisfatório para os caracteres avaliados, reduzindo o tamanho da população e permitindo alocar recursos apenas para genótipos promissores. No entanto, os entraves relacionados a esse teste são que mesmo alocando recursos nas gerações posteriores apenas para os melhores genótipos, até que sejam

selecionados, há limitações devido a necessidade de avaliação das progênies que antecedem o descarte e o maior tempo gasto (BERNARDO, 2003; BORÉM; MIRANDA, 2013).

É de responsabilidade do melhorista em cada programa, discernir a respeito de qual o melhor para método conduzir a sua população, não só entre os presentes na literatura, mas também realizando alguma modificação nos mesmos a fim de corrigir algum aspecto desfavorável. A estratégia escolhida deve ser a que proporcionará maiores ganhos genéticos com a seleção em menor espaço de tempo e com redução de custos (VENCOVSKY, 1987). O método deve também levar em consideração o objetivo do programa de melhoramento, variabilidade genética disponível a ser explorada ao longo das gerações, disponibilidade de infraestrutura do programa, assim como também a habilidade da equipe de melhoramento (MILADINOVIC *et al.*, 2011).

Inúmeras empresas dividem o mercado da soja no Brasil e no mundo, lançando várias cultivares a cada ano. Dessa forma, a rapidez na geração dessas cultivares é imprescindível para competir neste aspecto, o que faz com que metodologias do melhoramento clássico, como as usadas para seleção em gerações precoces exijam cada vez mais adaptações e melhorias para torna-las mais práticas e úteis aos programas de melhoramento.

2.6 A metodologia de $m + a$ tradicional e uma proposta de modificação

A seleção de progênies segregantes com potencial genético superior é, sem dúvidas, uma das etapas mais importantes em um programa de melhoramento. No entanto, altamente complicada quando feita precocemente por conta da complexidade dos caracteres de maior importância econômica. Assim, pelo processo de seleção, progênies de baixo potencial podem ser mantidas, enquanto aquelas superiores podem ser descartadas, por conta, entre outros fatores, de variações proporcionadas pelo ambiente e de efeitos gênicos não-aditivos não presentes em linhagens com altos níveis de homozigosidade (CARVALHO; GERALDI; SANTOS, 2009; VENCOVSKY; BARRIGA, 1992).

Diante dessas limitações, é interessante que a seleção precoce seja eficiente para identificar progênies com elevado potencial genético para extração de linhagens. Entre os procedimentos que são utilizados para alcançar esse propósito, está a estimativa de $m + a$ e d , ou simplesmente, $m + a$, proposta por Vencovsky (1987). Trata-se de uma extensão das teorias de Mather (1949) e Mather e Jinks (1982), que forneceram a base para predizer a média de um caráter quantitativo em gerações segregantes. O procedimento tem como propósito fornecer a

estimativa das médias das n progênes que venham a ser utilizadas, por meio da contribuição dos *loci* fixados pela homozigose na geração F_{∞} . Dessa forma, essa metodologia possibilita prever o potencial genético das progênes segregantes até mesmo em gerações iniciais.

Para obtenção das estimativas de $m + a$ e d , são consideradas as frequências genotípicas determinadas por $(p + q)^2$, em que p e q representam, respectivamente, as frequências do alelo efetivo (favorável) e do alelo não-efetivo (desfavorável) para um determinado loco, como descrito por Ramalho *et al.* (2012). Nessa abordagem, a média populacional (\bar{P}), é fornecida por:

$$\bar{P} = m + (p - q)a + 2pqd$$

em que, m é a média fenotípica para o loco em questão; a representa o efeito aditivo e; d os desvios de dominância. Contudo, quando são considerados os vários *loci* responsáveis pelo controle genético de um caráter quantitativo, a média da progênie será obtida pela expressão:

$$\bar{P} = m + \sum (p - q)a_k + \sum 2pqd_k$$

em que, a_k corresponde ao somatório dos efeitos aditivos, em relação à média, dos *loci* fixados nos indivíduos pela homozigose e d_k representa o somatório dos efeitos, em relação à média, dos *loci* em heterozigose proporcionado pelos desvios de dominância (VIANA *et al.*, 2009).

Embora Vencovsky (1987) tenha adaptado a metodologia para plantas alógamas, ela pode muito bem ser usada na predição do potencial genético de populações autógamias. Por exemplo, supondo que haja informações de n populações em duas gerações consecutivas de homozigosidade (F_2 e F_3), a predição seria calculada pelo seguinte contraste: $m + a = 2F_3 - F_2$, em que F_2 e F_3 são as médias das populações nas respectivas gerações. Vale salientar que poderia ser estabelecido contraste entre gerações consecutivas em outros níveis de homozigosidade. Contudo, pela abordagem tradicional, as gerações distintas das progênes devem ser avaliadas (conduzidas) na mesma condição ambiental (época), ou seja, em conjunto. Isso, na maioria das vezes, demanda muita área experimental, mão de obra e riscos inerentes até mesmo a uma perda dos ensaios em conjunto, já que todas as gerações a serem avaliadas são inseridas em campo simultaneamente.

Com o avanço das gerações sem seleção, os desvios dos heterozigotos reduzem-se pela metade a cada geração. Dessa forma, uma geração resultante do cruzamento entre parentais

teria média $F_1 = m + a + d$ e, em consequência, as gerações seguintes teriam médias $F_2 = m + a + 1/2d$, $F_3 = m + a + 1/4d$, e assim sucessivamente. Portanto, quando houver apenas *loci* em homozigose para a determinação desse contraste, a média das linhagens será apenas $P = m + a$ e diferenças serão ocasionadas apenas pelos *loci* fixados (BORÉM; MIRANDA, 2013; VIANA, 2007).

A estimativa de $m + a$ depende do desempenho *per se* dos indivíduos. Nesse sentido, para que uma população seja eleita como boa fonte para extração de linhagens, deve apresentar elevada estimativa de $m + a$. Ao comparar diferentes populações, a que contribui com maior frequência de alelos favoráveis certamente será classificada como a de maior potencial genético na geração avaliada (ROCHA *et al.*, 2015).

Matos (2016) utilizou a estimativa de $m + a$ para identificar se a geração mais precoce ($F_{3:4}$) gerava estimativas tão confiáveis quanto aquelas obtidas em $F_{3:5}$ na seleção de progênies de feijão-caupi. Com estimativas de elevada acurácia, a autora verificou que a seleção praticada na geração mais precoce proporcionaria resultados semelhantes àqueles gerados com mais um avanço de geração. Portanto, essa metodologia mostrou-se viável para identificação de progênies geneticamente superiores e, com isso, a possibilidade de descartar aquelas progênies com menor potencial genético para extração de linhagens.

Ao avaliar diversas estratégias para obtenção de linhagens com resistência a ferrugem asiática da soja, Ribeiro, Toledo e Ramalho (2009) aplicaram a metodologia de $m + a$ para conhecer o controle genético da doença com dados das gerações F_2 e F_3 de 10 populações de soja. As menores estimativas de $m + a$ implicaram em populações com menor susceptibilidade ao patógeno. Essa metodologia gerou resultados concordantes com todos os outros procedimentos utilizados para escolha das melhores populações.

A metodologia de $m + a$ trata-se, portanto, de um bom preditor de comportamento de progênies em gerações precoces. Ainda assim, como toda metodologia aplicável para fins de resultados no melhoramento genético, modificações e/ou adaptações podem ser adicionadas com o intuito de corrigir aspectos desfavoráveis ao trabalho do melhorista, principalmente no que diz respeito a parte experimental (CRUZ; REGAZZI; CARNEIRO, 2012).

A metodologia de $m + a$ atualmente aplicada oferece alguns riscos inerentes a experimentação, além da morosa execução. Por um lado, como a metodologia propõe que as gerações consecutivas de homozigose sejam conduzidas na mesma condição ambiental, ou seja,

em conjunto (VENCOVSKY, 1987), por outro, o risco de perda concentrada em maior parte dos ensaios é elevado. Por exemplo, essa perda pode ser resultado não só por erro de avaliação propriamente dita, mas também por morte das plantas e incidência de pragas ou doenças, que dessa forma poderia concentrar-se em uma quantidade bem maior de parcelas relativo ao tamanho do experimento em campo (LÚCIO; STORCK, 1999; STORCK, 2000).

Em extensos experimentos que utilizam grande área experimental é comum a perda de parcelas por erros de avaliação. E este problema acentua-se em decorrência da demanda quantitativa e qualitativa de mão de obra que se exige para condução dos mesmos, que muitas vezes não podem ser supridas por conta do dispêndio financeiro para esta finalidade (KLEINPAUL *et al.*, 2017). Diante desses “entraves”, o ideal seria fracionar os experimentos, de forma que avaliações de campo pudessem ser feitas sequencialmente, de preferência, à medida que as sementes das progênies em níveis variados de homozigidade estão sendo obtidas.

Outro problema inerente à experimentação da metodologia de $m + a$ trata-se da necessidade de se ter todas as sementes das gerações a serem avaliadas para semeadura simultânea, segundo proposta de Vencovsky (1987). Já é enorme o montante de progênies obtidas pelos programas de melhoramento que visam gerar cultivar, daí quando estas necessitam fazer parte de avaliações para metodologias como de $m + a$, ainda se acrescenta a necessidade de repetições e mais de uma geração a cada progênie.

Para a obtenção dessas sementes, avanços de gerações são necessários de serem feitos anteriormente à montagem dos experimentos que serão avaliados, de forma que somente após reunidas todas as gerações, os experimentos possam ser iniciados (MATOS, 2016). Isso requer um tempo considerável para obtenção e reunião dessas sementes, se tornando mais um entrave para uso desta metodologia, além do que, até reunir as sementes necessárias, ao iniciar os experimentos propriamente ditos as sementes da primeira geração obtida já não terão mais o mesmo vigor da seguinte, ao serem reunidas e utilizadas.

Para contornar esses problemas experimentais, uma modificação na metodologia de $m + a$ pode ser realizada, tornando o processo de experimentação mais facilmente exequível. O processo, consiste em avaliar progênies na medida em que se obtém as sementes para cada nível de homozigidade. Para isto, as avaliações das gerações utilizadas são realizadas de forma sequencial/gradativa, permitindo assim que cada geração seja avaliada em uma época

diferente, mas em um mesmo local. Esse formato de experimentação para as avaliações reduz os esforços e insumos necessários para obtenção dos dados que antes eram feitas de forma aglutinada segundo proposta de Vencovsky (1987) adaptada das teorias de Mather (1949) e Mather e Jinks (1982).

Esse formato de experimentação para as avaliações reduz os esforços por meio do fracionamento da equipe para as avaliações que antes teriam que ser de número maior de pessoas, já que o experimento com as gerações sequenciais estava simultaneamente em campo. Parte dos insumos necessários à experimentação também são reduzidos, como é o caso do sistema de irrigação que pode ser reaproveitado a cada experimento sequencial.

No entanto, o efeito de épocas que antes não estava presente na metodologia tradicional, aumenta a possibilidade de erros ao longo da experimentação. Para minimiza-los, testemunhas comuns com alto grau de homozigidade, como cultivares comerciais, são utilizadas junto às progênes a cada geração a ser avaliada, para que assim seja possível corrigir o efeito diferenciado do ambiente (épocas) ao longo dos experimentos com as gerações distintas e ajustar as médias fenotípicas das progênes, em análise semelhante ao proposto por Federer (1955).

Os experimentos, agora gradativos, se tornam menos propensos às perdas experimentais acima mencionadas. Por exemplo, as perdas referentes à incidência de uma praga em algum dos estádios de avaliação provavelmente terão menor intensidade (no sentido de tamanho de área experimental afetada), o que fará com que apenas a geração a ser avaliada no momento seja prejudicada. Mesmo em meio à redução de perdas experimentais, há que se considerar que os resultados ainda podem ter sua qualidade reduzida. Dessa forma, com a posse de todos os resultados pós avaliações dos experimentos pela nova proposta de metodologia, análises são procedidas de forma semelhante à metodologia tradicional (obtenção de contrastes), utilizando ferramentas estatísticas de qualidade para a obtenção de estimativas acuradas.

Para a obtenção dos contrastes, acredita-se que os melhores estimadores (BLUEs) e preditores (BLUPs) não-viesados, obtidos por modelos mistos, geram estimativas mais fidedignas a essa investigação para se inferir a respeito do potencial genético das progênes, seja pela metodologia na forma tradicional ou modificada (FREITAS, 2013).

Validar essa nova metodologia não é uma tarefa fácil, haja ser necessário conduzir experimentos que utilizem progênies, testemunhas e locais semelhantes para a montagem de ambas as metodologias para fins de comparação e cruzamentos de dados que permitam demonstrar a equivalência os mesmos. A nova proposta, se validada, torna-se uma ferramenta a mais para auxílio no trabalho do melhorista, contornando a grande maioria dos entraves da proposta tradicionalmente usada.

3 CAPÍTULO 2 - PROPOSTA DE MODIFICAÇÃO DA METODOLOGIA DE “ $m + a$ ” PARA SELEÇÃO PRECOCE DE PROGÊNIES DE SOJA

3.1 Introdução

A soja (*Glycine max* [L.]Merril) é uma oleaginosa de elevada importância econômica para o Brasil, sendo utilizada tanto na alimentação humana e animal, quanto em inúmeras aplicações industriais (CARRÃO-PANIZZI; SILVA, 2011). A cada ano a demanda pelo grão é crescente e, em razão disso, o cenário futuro para a soja é de uma atividade ainda mais lucrativa. Contudo, mudanças ambientais constantes fazem com que cultivares que hoje são empregadas com excelente produtividade, tenham reduções significativas de rendimento. Assim, para que esse problema seja contornado, novas cultivares mais adaptadas devem ser lançadas constantemente, atendendo às exigências dos produtores e consumidores da cultura (CERA, 2016).

Nesse contexto, o melhoramento genético busca elevar as frequências de alelos desejáveis para que seja possível gerar cultivares superiores às atuais. Todavia, as atividades de pesquisa no melhoramento clássico relativas a caracteres quantitativos, são dificultosas por demandar bastante tempo, recursos e área para execução de experimentos. Nesse sentido, a seleção precoce pode proporcionar economia de tempo e de recursos físicos e financeiros, pois permite descartar rapidamente aquelas progênies de baixo potencial genético para a geração de novas linhagens, concentrando esforços na avaliação nas demais (ROCHA *et al.*, 2015).

A estimativa de $m + a$ é uma metodologia que permite auxiliar o processo de seleção precoce, pois permite inferir a respeito do potencial genético destas, por meio de contrastes entre gerações consecutivas de homozigose ($m + a = 2F_3 - F_2$) (MATHER, 1949; MATHER E JINKS, 1982; VENCONVSKY, 1987). Os resultados desses contrastes nada mais é que a contribuição dos *loci* fixados por homozigose dos genótipos avaliados. Porém, para esta aplicação, a experimentação exige que experimentos das gerações consecutivas de homozigose sejam realizados de forma simultânea como nos trabalhos de Ribeiro, Toledo e Ramalho (2009) com soja e Rocha *et al.* (2015) com feijão, tornando a metodologia morosa e com demanda de grandes áreas experimentais uniformes e mão de obra qualificada.

No presente estudo propõe-se modificar a metodologia tradicional da estimativa de $m + a$ de forma a tornar o processo menos dificultoso. Para tanto, as progênies foram avaliadas na medida em que se obteve as sementes para cada geração, permitindo avaliações sequenciais

e menos concentradas temporalmente, reduzindo a possibilidade de perda de dados. Para correção do efeito ambiental causado por épocas diferenciadas de avaliação, propõe-se o uso de testemunhas em comum para todas as gerações.

Em geral, o processo experimental da metodologia de $m + a$ tende a desbalanceamentos por erros ou perdas de dados (RESENDE, 2004). Dessa forma, tanto para metodologia tradicional como para a modificação proposta é ideal o uso de procedimentos estatísticos que permitam obter resultados acurados, evitando resultados viciados. O uso de melhores estimadores (BLUEs) e preditores (BLUPs) não-viesados, obtidos por modelos mistos, geram estimativas mais confiáveis e de qualidade para que assim se possa ter maior confiabilidade a respeito do potencial de progênies em quaisquer níveis de homozigidade.

Com base no exposto, objetivou-se com esse estudo propor uma modificação para a metodologia de $m + a$; comparar os resultados gerados aos obtidos pela metodologia tradicional e verificar a eficiência do uso adicional do *pedigree* para predizer o comportamento das gerações envolvidas na cultura da soja.

3.2 Material e métodos

3.2.1 Material genético

Foram utilizadas 20 progênies segregantes (Tabela 1) derivadas de hibridações realizadas em 2016 pelo Programa de Melhoramento de Soja do Centro Universitário Integrado de Campo Mourão-PR. Adicionalmente, foram usadas quatro testemunhas comerciais (Tabela 2), todas cultivares de ampla aceitação na região Sul do Brasil.

3.2.2 Condução experimental

3.2.2.1 Experimentos simultâneos da metodologia tradicional

As 20 progênies segregantes foram avançadas até a geração F_4 em fileiras de 15 metros de comprimento com 300 indivíduos, usando as premissas do método SSD (*single seed descent*). Dessa forma, foi permitido que os efeitos proporcionados pela deriva genética e seleção fossem mínimos, e que cada indivíduo F_2 pudesse ser representado em F_4 .

Tabela 1. Progenies segregantes de soja utilizadas em três níveis de homoziguidade (F₂, F₃ e F₄) para avaliação das metodologias tradicional e modificada.

Número	Código	Número	Código
1	16INT1237	11	16INT1444
2	16INT1298	12	16INT1447
3	16INT1323	13	16INT1449
4	16INT1347	14	16INT1460
5	16INT1348	15	16INT1491
6	16INT1349.1	16	16INT1501
7	16INT1412	17	16INT1510
8	16INT1428	18	16INT1515
9	16INT1432	19	16INT1545
10	16INT1253	20	16INT1835

Tabela 2. Testemunhas comerciais e seus respectivos grupos de maturação utilizadas nas metodologias tradicional e modificada.

Testemunha	Nome da testemunha	Grupo de Maturação
21	BMX Potência RR	6.7
22	TMG 7062 IPRO	6.2
23	DM 6563 IPRO	6.3
24	M 6210 IPRO	6.2

De posse das sementes, foram conduzidos experimentos em uma única época (simultâneos) com as gerações F₂, F₃ e F₄ em duas localidades contrastantes no estado do Paraná (Tabela 3). Os experimentos foram conduzidos em blocos casualizados com 24 tratamentos (20 progenies segregantes + 4 testemunhas comerciais) e três repetições. As parcelas foram constituídas por quatro fileiras de cinco metros de comprimento, em que as duas fileiras centrais foram consideradas como área útil experimental. As demais fileiras foram constituídas por uma cultivar comercial, representando a bordadura de cada parcela.

O espaçamento entre fileiras foi de 0,5 metros com 12 plantas por metro linear para os experimentos de Luiziana/PR e de 15 plantas por metro linear para os experimentos de Campo Mourão/PR. A densidade foi escolhida segundo recomendações de plantio pré-estabelecidas pela empresa cedente das sementes para cada localidade.

Tabela 3. Características dos locais utilizados na condução dos experimentos.

Local	Município/ Estado	Coordenadas	Altitude	Classificação do solo
Fazenda Record	Luiziana/PR	24° 12' 21.29" S 52° 16' 36.74" O	720 m	Latossolo vermelho distroférico
Centro Universitário Integrado	Campo Mourão/PR	23° 59' 24.81" S 52° 21' 50.15" O	560 m	Latossolo vermelho distroférico

Os experimentos, nas duas localidades, com as gerações F₂, F₃ e F₄, foram instalados na safra 2018/2019 em única época a cada local. Todos os tratamentos culturais, da semeadura até a colheita, foram realizados conforme as recomendações para o plantio de soja nas condições do estado do Paraná.

3.2.2.2 Experimentos sequenciais da metodologia modificada

Com o intuito de gerar as estimativas de $m + a$ pela metodologia modificada, as 20 progênies (Tabela 1) das três gerações (F₂, F₃ e F₄) foram conduzidas em épocas distintas (sequenciais) em conjunto às mesmas testemunhas comerciais (Tabela 2) da metodologia tradicional, nas duas localidades mencionadas na Tabela 3. Os experimentos com as três gerações foram instalados antes dos experimentos do item 3.2.2.1, na safra 2017/2018, em épocas gradativas após a segunda quinzena de setembro/2017. Ao final de setembro de cada ano, o estado do Paraná apresenta mudanças constantes de fotoperíodo a cada sete dias corridos. Assim, experimentos espaçados em um mês, apresentaram claramente condições de fotoperíodo e temperatura bem diferentes.

Em partes, os experimentos foram conduzidos de forma semelhante aos experimentos simultâneos da metodologia tradicional. Portanto, informações sobre delineamento experimental, tamanho das parcelas, espaçamento e tratamentos culturais foram as mesmas descritas no item 3.2.2.1.

3.2.3 Mensuração de caracteres avaliados

Nas avaliações das gerações, tanto nos experimentos simultâneos quanto nos graduais, foram obtidas informações nas áreas úteis para os seguintes caracteres:

- ALT (altura de planta) – medida em centímetros da superfície do solo (colo da planta) ao ápice da haste principal no estádio R8.

- FLR (dias para a floração) – período em dias da data de semeadura até a data em que aproximadamente 50% das plantas da fileira possuíam pelo menos uma flor aberta. Considerada data de R1.
- MAT (dias para a maturação) - período em dias da data de semeadura até a data em que aproximadamente 50% das vagens de cada progênie apresentaram-se maduras. Considerada data de R7.
- PMS (peso de mil sementes) - massa em gramas de uma amostra aleatória de 1000 sementes obtida por meio de uma balança de precisão.
- PROD (produtividade) – massa total das sementes colhidas na área útil e extrapoladas para um hectare.

3.2.4 Quadro de avaliações

No total, somaram-se 12 experimentos (seis experimentos em cada localidade sendo três de cada metodologia). Em cada experimento foram avaliados todos os caracteres com exceção de dias para a floração (FLR) em algumas gerações, por causa de perdas experimentais (Tabela 4).

Tabela 4. Caracteres avaliados a cada geração (F₂, F₃ e F₄) nas metodologias tradicional e modificada nos municípios de Campo Mourão/PR e Luiziana/PR.

Caracteres	Local	Metodologias e Gerações Avaliadas					
		Tradicional			Modificada		
		F ₂	F ₃	F ₄	F ₂	F ₃	F ₄
ALT	Campo Mourão/PR	x	x	x	x	x	x
FLR					x	x	
MAT		x	x	x	x	x	x
PMS		x	x	x	x	x	x
PROD		x	x	x	x	x	x
ALT	Luiziana/PR	x	x	x	x	x	x
FLR		x	x	x	x	x	x
MAT		x	x	x	x	x	x
PMS		x	x	x	x	x	x
PROD		x	x	x	x	x	x

3.2.5 Análises genético-estatísticas

3.2.5.1 Análises de deviance e estimativas dos componentes de variância

Com o intuito de estimar os componentes de variância para determinação dos principais parâmetros genéticos os dados fenotípicos foram submetidos à análise de modelos mistos pelo procedimento REML/BLUP (Método da Máxima Verossimilhança Restrita/Melhor Preditor Linear Não Viesado) e BLUE (Melhor Estimador Linear Não Viesado). Também foram possíveis prever os valores genotípicos médios e as estimativas de médias ajustadas das progênies.

Para isso, os dados dos caracteres obtidos das avaliações dos experimentos simultâneos da metodologia tradicional (*item 3.2.2.1*) foram analisados *a priori* considerando o seguinte modelo genético-estatístico:

$$\mathbf{y} = \mathbf{Xr} + \mathbf{Wp} + \mathbf{Yl} + \mathbf{Zpl} + \mathbf{e} \quad (1)$$

em que: \mathbf{y} é o vetor das médias fenotípicas das progênies; \mathbf{r} é o vetor do efeito de repetição somados à média geral; \mathbf{p} é o vetor do efeito de progênies, em que, $p \sim N(0, G)$, sendo; \mathbf{l} é o efeito de locais; \mathbf{pl} é o vetor do efeito da interação P x L; \mathbf{e} é o vetor de erros, sendo que $e \sim N(0, R)$ e $R = I\sigma_e^2$; e \mathbf{X} , \mathbf{W} , \mathbf{Y} e \mathbf{Z} são matrizes de incidência que relacionam, respectivamente, os efeitos de \mathbf{r} , \mathbf{p} , \mathbf{l} e \mathbf{pl} ao vetor \mathbf{y} . Nessa primeira análise, todos os efeitos foram considerados aleatórios. Nesse modelo o efeito de épocas não é considerado, já que todos os experimentos foram montados em uma só época nos dois locais.

Posteriormente, os dados das avaliações dos experimentos simultâneos foram submetidos às análises por esse mesmo modelo, no entanto, com os efeitos de progênies (\mathbf{p}) e de interação progênie x local (\mathbf{pl}) considerados como fixo para fins de estimativas de médias ajustadas (BLUEs) e previsões genotípicas (BLUPs).

Os dados dos caracteres obtidos das avaliações dos experimentos sequenciais para a metodologia modificada (*item 3.2.2.2*), foram analisados *a priori* considerando o seguinte modelo genético-estatístico:

$$\mathbf{y} = \mathbf{Vs} + \mathbf{Xr} + \mathbf{Wp} + \mathbf{Ya} + \mathbf{Zpa} + \mathbf{e} \quad (2)$$

em que: \mathbf{y} é o vetor das médias fenotípicas das progênies; \mathbf{s} é o vetor do efeito aninhado de testemunhas dentro de gerações; \mathbf{r} é o vetor do efeito de repetição dentro de ambiente somados

à média geral; \mathbf{p} é o vetor do efeito de progênes, em que, $p \sim N(0, G)$, sendo $G = I\sigma_g^2$; \mathbf{l} é o efeito de local; \mathbf{pl} é o vetor do efeito da interação PxL; \mathbf{e} é o vetor de erros, sendo que $e \sim N(0, R)$ e $R = I\sigma_e^2$; e \mathbf{V} , \mathbf{X} , \mathbf{W} , \mathbf{Y} e \mathbf{Z} são matrizes de incidência que relacionam, respectivamente, os efeitos de \mathbf{g} , \mathbf{r} , \mathbf{p} , \mathbf{a} e \mathbf{pa} ao vetor \mathbf{y} . Nessa primeira análise, todos os efeitos foram considerados aleatórios.

Posteriormente, os dados das avaliações dos experimentos sequenciais foram submetidos a análises por esse mesmo modelo, no entanto, com os efeitos somente de progênes (\mathbf{p}) e de interação progênie x local (\mathbf{pl}) considerados como fixo para fins de estimações de médias ajustadas (BLUEs) e previsões genótípicas (BLUPs), da mesma forma como procedeu nas análises dos experimentos conjuntos da metodologia tradicional.

Utilizando os modelos genético-estatístico 1 e 2 foram realizadas as análises de deviance (ANADEVs) para as metodologias tradicional e modificada, respectivamente. O efeito entre épocas distintas foi estimado e ajustado pelo desempenho das testemunhas (cultivares), que estiveram presentes em todas as avaliações dos experimentos da metodologia modificada.

Os efeitos considerados como aleatórios foram testados pela razão de verossimilhança (LRT) via análises de deviance (ANADEV) (RESENDE, 2007) e os fixos pelo teste F de Wald.

A herdabilidade em nível de médias (h_m^2), foi obtida para todos os caracteres a cada metodologia pela seguinte expressão:

$$h_m^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \frac{\hat{\sigma}_e^2}{rl}}$$

em que, $\hat{\sigma}_g^2$ = variância genotípica, $\hat{\sigma}_e^2$ = variância residual, l = número de locais, r = número de repetições.

A partir dos valores de herdabilidades foram estimadas as acurácias (\hat{r}_g), pela expressão:

$$\hat{r}_g = \sqrt{h_m^2}$$

As análises dessa etapa (3.2.5.1) foram realizadas com o auxílio do pacote breedR (MUÑOZ; SANCHEZ, 2014) do software R (R CORE TEAM, 2015).

3.2.5.2 Obtenção das estimativas de $m + a$

As estimativas de $m + a$ foram obtidas por duas formas distintas:

a) *Contrastes entre os BLUEs das gerações F_2 , F_3 e F_4 das avaliações dos experimentos simultâneos da metodologia tradicional*

Os BLUEs das progênes segregantes obtidos com os dados das gerações conduzidas simultaneamente foram confrontados por meio dos contrastes dos tipos:

$$(i) \quad m + a = 2F_3 - F_2$$

$$(ii) \quad m + a = 2F_4 - F_3$$

A avaliação simultânea das gerações foi realizada conforme a metodologia de $m + a$ tradicional, de forma a minimizar a interferência ambiental (mesma época de semeadura) e em duas localidades (Campo Mourão/PR e Luiziana/PR). Os dados coletados nessa avaliação foram usados como “controle”. Assim, foram analisados *a priori* pelo modelo 1, descrito no item 3.2.5.1 e, *a posteriori*, os BLUEs foram contrastados conforme descrito em (i) e (ii).

b) *Contrastes entre as gerações F_2 , F_3 e F_4 das avaliações dos experimentos sequenciais da metodologia modificada*

Os BLUEs das progênes segregantes obtidos com os dados das gerações conduzidas sequencialmente foram confrontados por meio de contrastes como no item “a” deste tópico.

Os dados coletados nessa avaliação foram analisados, *a priori*, pelo modelo II descrito no item 3.2.5.1 e *a posteriori* os BLUEs foram contrastados conforme descrito em “a”.

As análises dessa etapa (3.2.5.2) foram realizadas com o auxílio do pacote breedR (MUÑOZ; SANCHEZ, 2014) do software R (R CORE TEAM, 2015).

3.2.5.3 Correlações dentro e entre os resultados obtidos em ambas as metodologias

Para verificar a equivalência entre os resultados obtidos de cada metodologia (entre gerações a cada metodologia) e entre metodologias, foram correlacionados os BLUEs para as seguintes situações:

a) Correlação entre os BLUEs ao longo das gerações dentro de cada metodologia

A cada caráter avaliado, os BLUEs obtidos das progênies foram correlacionados entre eles para que fosse possível se observar o comportamento entre as estimativas de médias ajustadas entre uma geração e outra de cada metodologia.

b) Correlação entre os contrastes ao longo das gerações dentro de cada metodologia

A cada caráter avaliado, os contrastes ortogonais ($m + a$) oriundos das gerações consecutivas F_2 e F_3 foram correlacionados com os de F_3 e F_4 , pertencentes a cada metodologia. Inicialmente, a correlação estabelecida foi apenas dentro de cada metodologia com a finalidade de comparar a coincidência entre a estimativa de $m + a$ de cada geração.

c) Correlação entre os contrastes das gerações F_2 e F_3 com os BLUEs da geração F_4

Os contrastes das gerações F_2 e F_3 foram correlacionados para todos os caracteres de cada metodologia com os BLUEs da geração F_4 , com a finalidade de prever o comportamento das estimativas das médias das progênies em F_4 .

Todas as correlações supracitadas foram obtidas por meio de pares de resultados obtidos nas diversas situações entre as duas metodologias (BLUE, BLUE'), ($m + a, m + a'$) ou ($m + a, BLUE$) em cada metodologia analisada utilizando conjuntamente os dois locais avaliados para cada uma. Para isso, foi utilizado o coeficiente de correlação de *Spearman*, determinado pela seguinte expressão:

$$r_{a,a'} = \frac{1 - 6 \sum d_i^2}{n(n^2 - 1)}$$

em que, d_i^2 é a diferença entre as posições de cada valor correspondente de a e a' e n é o número de pares de valores.

Para melhor visualização dos resultados, foram elaborados gráficos do tipo coluna a partir de todas as correlações obtidas.

As análises dessa etapa (3.2.5.3) foram realizadas com o auxílio do pacote `breedR` (MUÑOZ; SANCHEZ, 2014) do software R (R CORE TEAM, 2015). Os gráficos foram elaborados com o auxílio dos pacotes `ggplot2` (WICKHAM, 2016) e `reshape2` (WICKHAM, 2007).

3.2.5.4 Cenários preditivos via BLUPs

Com o intuito de se prever o comportamento das gerações via BLUPs (valores genotípicos) a partir dos BLUEs (estimativas de médias ajustadas) das gerações, e uma matriz de parentesco elaborada para as 20 progênies, foram propostos os seguintes cenários:

a) Capacidade preditiva para as gerações F_3 e F_4 a partir das estimativas de médias obtidas na geração F_2 , com matriz de parentesco das progênies

Para isso foram excluídos os BLUEs obtidos nas gerações F_3 e F_4 de ambas as metodologias e conservados os BLUEs da geração F_2 de forma a se prever os valores genotípicos de F_3 e F_4 . Em seguida, verificou-se a eficiência dessa capacidade preditiva por meio de correlação.

b) Capacidade preditiva para a geração F_4 a partir das estimativas de médias obtidas nas gerações F_2 e F_3 com matriz de parentesco das progênies

Para isso foram excluídos os BLUEs obtidos na geração F_4 de ambas as metodologias e conservados os BLUEs das gerações F_2 e F_3 , de forma a se prever os valores genotípicos de F_4 . Em seguida, verificou-se a eficiência desta capacidade preditiva por meio de correlação.

c) Capacidade preditiva para todas as gerações a partir das estimativas de médias obtidas nas próprias gerações F_2 , F_3 e F_4 , com matriz de parentesco das progênies

Para isso foram conservados os BLUEs obtidos em todas as gerações de ambas as metodologias de forma a se prever os valores genotípicos de todas a partir das suas próprias estimativas de média ajustadas. Em seguida, verificou-se mais uma vez a eficiência dessa capacidade preditiva por meio de correlação.

Para todos os cenários propostos foram observadas as capacidades preditivas para cada caráter avaliado, assim como também para as metodologias tradicional e modificada, com o propósito de realizar uma validação.

As correlações foram calculadas por meio do coeficiente de correlação de Spearman ($r_{a,a'}$), idem ao item 3.2.5.3. Os gráficos foram elaborados para melhor visualização dos resultados obtidos com os cenários propostos.

Para a geração da matriz de parentesco a ser usada nos cenários preditivos foram utilizadas a genealogia (pai e mãe) das 20 progênies utilizadas no trabalho. Não foi possível obter os dados relacionados à genealogia das testemunhas pois as cultivares eram oriundas de outras empresas. A partir desses dados criou-se uma matriz de dimensões 64 x 64, com as respectivas genealogias a serem incluídas para cálculos de predições junto aos BLUPs nos cenários propostos.

As análises dessa etapa (3.2.5.4) foram realizadas com o auxílio do pacote pedigreeemm (VAZQUEZ *et al.*, 2010) do software R (R CORE TEAM, 2015). Os gráficos foram elaborados com o auxílio dos pacotes ggplot2 (WICKHAM, 2016) e reshape2 (WICKHAM, 2007).

3.3 RESULTADOS E DISCUSSÃO

3.3.1 Análises de verossimilhança (ANADEVs) e parâmetros genéticos

Pela metodologia tradicional (Tabela 5) houve diferenças significativas ($p < 0,01$) para o efeito de progênie para todos os caracteres avaliados, com exceção de dias para a floração, o que indica que para a maioria desses caracteres há variabilidade genética, tornando possível a obtenção de ganhos mediante seleção. Essas diferenças podem ter sido influenciadas pelos diferentes níveis de homozigosidade, já que a cada avanço de geração uma mesma progênie gera níveis de segregação distintos, tornando-as diferentes geneticamente (BARBOSA-PRESTES *et al.*, 2008). O caráter dias para a floração, que não dispôs de avaliações para um dos locais avaliados (Tabela 4) pode não ter dado contribuição de variabilidade para torná-lo significativo a nível de progênies.

Tabela 5. P-valores para os efeitos fixos de progênie e interação progênie x local e para os efeitos aleatórios de local e repetição/ambiente dos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ tradicional.

Efeitos	p-valores				
	FLR	MAT	ALT	PMS	PROD
Repetição/Ambiente ²	0,2534	0,2869	0,2099	0,3524	0,0134
Progênie(P) ¹	0,9416	<0,0001	<0,0001	<0,0001	0,0003
Local(L) ²	-	0,0349	0,4538	0,0166	0,1090
Interação P x L ¹	-	0,0021	<0,0001	<0,0001	<0,0001
Resíduo	3,9241	9,3154	0,1524	17,7357	342960
Médias	51,82	120,53	0,92	165,79	4350,89
CV (%)	3,82	2,53	6,03	2,54	13,46
h_m^2	0,90	0,94	0,86	0,85	0,74
Acurácia	0,95	0,97	0,93	0,92	0,86

¹Teste de razão de verossimilhança (LRT), testado via qui-quadrado com 1 grau de liberdade.

²Teste F de Wald.

Valores em negrito são significativos ($p < 0,01$) ou ($p < 0,05$) para ambos os testes

Não foram observadas diferenças significativas para o efeito de repetição/ambiente (Tabela 5) para a grande maioria dos caracteres estudados, com exceção da produtividade ($p < 0,05$), o que na prática significa que houve homogeneidade nos blocos dos experimentos conduzidos nas duas localidades. Esse efeito está aninhado ao local, que por sua vez não apresentou diferenças suficientemente contrastantes para altura de planta e produtividade. Na prática, os dois locais avaliados possuem influência ambiental das suas diferentes altitudes no que diz respeito ao desenvolvimento da cultura. Para cada local avaliado, há recomendações de densidades diferentes de plantio (que variam por causa da altitude), um dos fatos que comprova a distinção ambiental entre esses para a cultura da soja. O caráter dias para a floração, pela falta de avaliações para um dos locais (Tabela 4) foi isento de inferências relacionadas ao fator de variação local nessa metodologia.

Houve diferenças significativas ($p < 0,01$) para o fator interação P x L (Tabela 5) para todos os caracteres avaliados, o que demonstra que as progênies se comportaram de forma diferente em cada local avaliado. Pelo fato de não haver avaliações do caráter dias para a floração em um dos locais, também não se pôde verificar a respeito da interação P x L.

Todos os caracteres apresentaram baixos valores para o coeficiente de variação CV (%), comprovando elevada confiabilidade das estimativas obtidas neste estudo. Em adição aos coeficientes de variação para verificação da qualidade experimental, observou-se elevados valores de acurácia seletiva, decorrentes das altas magnitudes das herdabilidades médias, mesmo para os caracteres em questão, que são de natureza quantitativa e herança complexa. Altos coeficientes de herdabilidade torna possível a seleção de progênes com elevada acurácia em gerações iniciais de homozigose (ROSADO *et al.*, 2012). A acurácia demonstra a precisão da inferência acerca dos valores genotípicos, ou seja, altos valores de acurácia remetem a estimações ou predições de vício nulo ou pequeno. Sendo assim, avaliar a qualidade experimental por meio desta não retrata o uso de uma perspectiva puramente estatística como o coeficiente de variação (RESENDE; DUARTE, 2007).

É importante ressaltar que os valores de herdabilidade e acurácia seletiva, mesmo sendo altos, decrescem no sentido da complexidade de cada caráter avaliado, o que já era esperado, pois caracteres de maior complexidade tanto envolvem um maior número de genes para determinação assim quanto sofrem uma maior influência do efeito ambiental. Dentre estes, os menores valores de herdabilidade e acurácia foram de peso de mil sementes, 0,85 e 0,92, respectivamente e de produtividade, 0,74 e 0,86, respectivamente.

Para a metodologia de $m + a$ modificada (Tabela 6) houve diferenças significativas para o efeito de progênes para quase todos os caracteres avaliados ($p < 0,01$), com exceção da produtividade. Portanto, para a grande maioria dos caracteres há variabilidade genética presente, o que possibilita ganhos com a seleção.

Nesta metodologia não foram observadas diferenças significativas para o efeito de repetição. O efeito de testemunha/geração diferiu estatisticamente para todos os caracteres, o que mostra que as três gerações avaliadas se comportam de forma diferente. Para essa metodologia, o efeito de geração tem o efeito de testemunha aninhado a ele, pois, a inclusão de tratamentos regulares (testemunhas) em comum a cada geração, permite que seja deduzida (descontada) a variação inerente as diferenças ambientais entre as gerações (níveis diferenciados de homozigosidade). O experimento de cada geração que corresponde ao conjunto de progênes e testemunhas, mas em níveis de homozigosidade diferentes, foi inserido em campo em épocas distintas, mas analisado de forma conjunta como na metodologia tradicional. Para tanto, foi inserido o efeito de testemunha/geração no modelo estatístico dessa modificação.

Tabela 6. P-valores para os efeitos fixos de progênie e interação progênie x local e para os efeitos aleatórios de local, repetição e testemunha/geração dos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ modificada.

Efeitos	p-valores				
	FLR	MAT	ALT	PMS	PROD
Progênie (P) ¹	<0,0001	<0,0001	<0,0001	0,0020	0,2276
Local (L) ²	0,0186	0,0281	0,2721	0,0210	0,0271
Interação P x L	0,0355	<0,0001	0,0077	0,0013	<0,0001
Repetição ²	0,4345	0,1894	0,3105	0,2154	0,1664
Testemunha/Geração ²	0,0027	0,0010	0,0032	0,0126	0,0030
Resíduo	35448	6,43030	28,704	11791000	2,72E11
Médias	53,02	122,67	0,93	166,31	4323,02
CV (%)	3,55	2,07	40,79	6,53	69,22
h_m^2	0,93	0,93	0,88	0,68	0,54
Acurácia	0,97	0,97	0,94	0,83	0,73

¹Teste de razão de verossimilhança (LRT), testado via qui-quadrado com 1 grau de liberdade.

²Teste F de Wald.

Valores em negrito são significativos ($p < 0,01$) ou ($p < 0,05$) para ambos os testes

O efeito de local diferiu estatisticamente ($p < 0,05$) para quase todos os caracteres avaliados, com exceção de altura de planta, o que mostra que estatisticamente há diferenças entre o comportamento dos locais de avaliação, o que na prática já é sabido como mencionado anteriormente. Já a interação progênie x local diferiu para quase todos os caracteres avaliados, com exceção de dias para floração, o que demonstra que as progênies se comportam de forma diferente em cada local avaliado. Assim como na metodologia tradicional também houve ausência das avaliações de dias para a floração nessa metodologia, porém com menores perdas, tendo ausência de valores na avaliação F₄ de um dos locais (Campo Mourão-PR) (Tabela 4). Esse detalhe, a se observar com mais cautela, pode ter contribuído para que esse caráter não apresentasse diferenças significativas.

Os coeficientes de variação foram baixos para dias para a floração, dias para a maturação e peso de mil sementes o que denota alta precisão experimental na avaliação desses caracteres e alta confiabilidade das estimativas. Os valores de CV para altura de planta e

produtividade foram altos e bem maiores que aqueles observados para os mesmos caracteres na metodologia tradicional. Assim, entende-se que houve bastante variação ambiental entre um experimento e outro (sequencial). Vale salientar que essa variação é capturada pelas testemunhas comuns a todos os experimentos, o que possibilita corrigir/ajustar as médias fenotípicas das progênes em cada nível de homoziguidade.

Elevados valores de acurácia foram observados juntos a herdabilidade. Semelhantemente aos valores da Tabela 5, as magnitudes decresceram no sentido da complexidade de cada caráter avaliado. Dentre estes, os menores valores de herdabilidade e acurácia foram para o peso de mil sementes e para produtividade. Contudo, é necessário ressaltar que a avaliação em Campo Mourão-PR teve leve infestação de nematóides na área experimental, que afetou parte das parcelas principalmente no último quarto do seu ciclo, cujos caracteres ligados à produtividade são mais afetados.

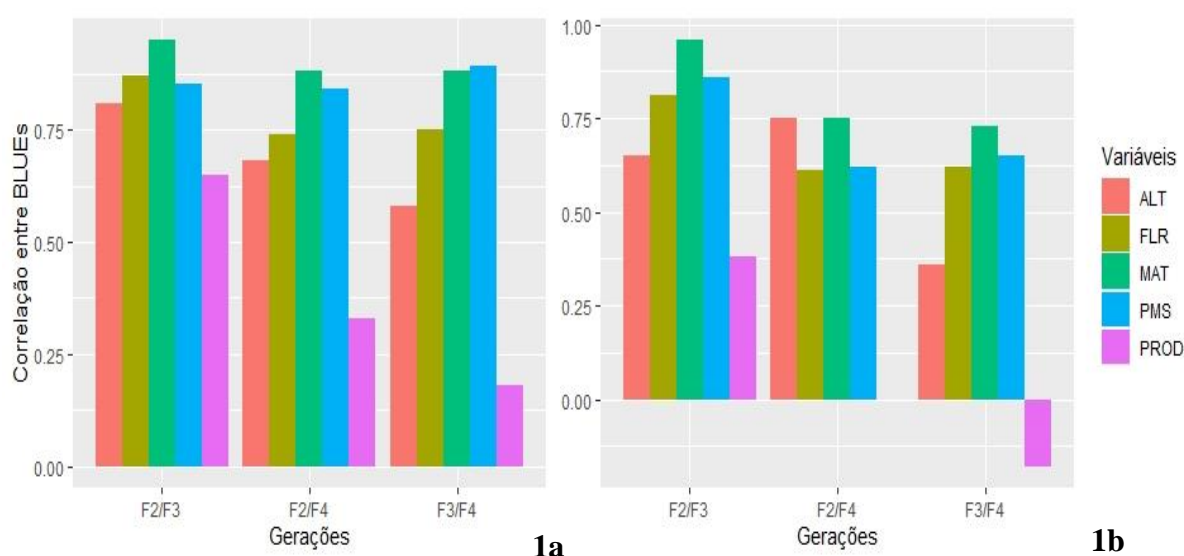
3.3.2 Médias estimadas de valores fenotípicos via BLUEs e correlações

Ao se avaliar as progênes nos três níveis de homoziguidade propostos, os valores fenotípicos obtidos permitiram que fossem preditos valores genotípicos que correspondem a resultados não observáveis de natureza aleatória via BLUP. No entanto, ao obter as estimativas de $m + a$ contrastando duas gerações consecutivas ocorreria uma dupla penalização por herdabilidade, relativa à penalização no BLUP de cada geração. Dessa forma, valores de BLUEs (Apêndice A) obtidos quando se considera o efeito de progênie como fixo, e que não faz em uso dessa penalização, foram utilizados para a obtenção das estimativas de $m + a$ (Apêndice B) e para o uso de correlações entre as próprias médias (Apêndice C). Segundo Pastina (2010), quando o objetivo é a determinação de diferenças entre pares de genótipos, o método de predição via BLUP é inadequado por conta da obtenção de um resultado, neste caso, subestimado. O efeito utilizado no contraste deve então ser considerado como fixo.

As correlações de *Spearman* entre as gerações F₂ e F₃ da metodologia tradicional foram positivas e mais fortes para os caracteres altura, dias para a floração, dias para a maturação e produtividade com valores de 0,81, 0,87, 0,95, 0,65, respectivamente (Gráfico 1). Em termos gerais, as maiores correlações foram entre as gerações F₂ e F₃ para a maioria dos caracteres avaliados, exceto para o peso de mil sementes entre as gerações F₃ e F₄ (0,89). Porém, ainda com um valor muito próximo de sua correlação entre F₂ e F₃ (0,85). Isso mostra que o uso de médias ajustadas via BLUEs é bastante informativo para inferência de caracteres poligênicos na avaliação de progênes segregantes de soja. Portanto, em teoria, avaliar gerações

consecutivas F_3 e F_4 para discernir quais são aquelas superiores via contraste de $m + a$ seria mais útil, já que entre F_2 e F_3 há bastante coincidência das médias e comportamento, o que mostra menor alteração no comportamento de progênies da geração F_2 para F_3 do que de F_3 para F_4 , onde provavelmente mais *loci* fixaram-se por homozigose.

Gráfico 1. Correlações de *Spearman* entre os BLUEs das gerações F_2 , F_3 e F_4 nos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologias de $m + a$ tradicional (1a) e modificada (1b).



De acordo com Bernardo (1991), uma das formas de se avaliar a eficiência da seleção precoce é por meio da correlação estabelecida entre o fenótipo de um indivíduo ou família em geração inicial e qualquer outra que se deseje observar o comportamento e promover a seleção. No caso dos resultados obtidos, as correlações não foram estabelecidas apenas entre médias fenotípicas, mas entre médias ajustadas oriundas de metodologia de modelos mistos, dando ainda mais qualidade aos resultados obtidos.

Neste estudo, os BLUEs obtidos com a finalidade de gerar estimativas de $m + a$ para confronto de metodologias já mostram indícios de equivalência entre as mesmas, pois no geral as correlações seguiram uma mesma tendência em ambas as metodologias (1a e 1b). Observou-se que os BLUEs oriundos da metodologia modificada (1b) continuam a ter, por exemplo, valores de correlação F_2 com F_3 mais acentuados para a maioria dos caracteres quando comparadas aos resultados das demais correlações. Na metodologia proposta, apenas altura de plantas apresentou correlação entre F_2 e F_3 inferior (0,65) a aquela observada entre F_2 e F_4

(0,75). Tais resultados possibilitam possível validação da modificação proposta à metodologia que Vencosvsky (1987) sugeriu a partir da extensão de teorias já estabelecidas por Mather (1949) e Mather e Jinks (1982).

Resultados em maior parte semelhantes entre as duas metodologia apresentadas, que possa vir a validar a modificação proposta, não ocorreram puramente ao acaso, pois há que se relembrar que para provar a equivalência existem muitos pontos a contar positivamente como: os dados são de experimentos totalmente independentes, ora conduzidos conforme exigências da metodologia de $m + a$ tradicional (com três gerações avaliadas simultaneamente), e em um outro momento, com experimentos conduzidos conforme proposta da modificação da metodologia (três gerações conduzidas em épocas distintas - sequencialmente); os experimentos foram repetidos em duas localidades com condições climáticas bastante diferentes; os resultados das avaliações são para cinco caracteres de herança poligênica avaliados ao longo do ciclo da cultura; foi adicionada uma geração a mais de homozigose do que propõe o método tradicional para desse modo permitir o resultado de dois contrastes a cada local ao invés de um e uso de 20 progênies segregantes geradas a partir de alguns cruzamentos entre pais de comportamento bastante divergente. Todos esses argumentos comprovam que o conjunto de dados usado neste estudo é bem informativo, representativo e permite tirar conclusões bem amplas a respeito da temática contemplada.

Ao se comparar as correlações entre F_3 com F_4 e F_2 com F_4 , observa-se que em F_2 com F_4 as magnitudes das correlações são mais elevadas para maior parte dos caracteres (Gráfico 1), independentemente da metodologia aplicada. Isso mostra que as médias das gerações F_2 coincidem muito mais com F_4 , que de F_3 com F_4 , ou seja, os BLUEs reduzem-se de F_2 para F_3 e elevam-se novamente de F_3 para F_4 . Esse comportamento não era esperado, já que há perdas de desvios de dominância de uma geração para outra, principalmente de F_2 para F_3 . Dessa forma, seria esperada maior correlação entre os dois últimos níveis de homozigosidade.

Até mesmo neste caso de comportamento não esperado entre os resultados, se comparadas as duas metodologias, encontram-se novamente grandes coincidências, agora entre esses dois últimos pares de correlações analisados (F_3/F_4 e F_2/F_4). Isso prova mais uma vez a equivalência entre os métodos utilizados. A cada par de gerações correlacionadas por qualquer um dos métodos, o caráter produtividade que tem a maior quantidade de genes envolvidos na expressão, apresenta sempre a menor correlação entre os demais.

A partir de F_2 a variância genética é composta não só por variância aditiva, mas também por desvios de dominância e epistasia (não-aditiva), de modo que quando a última está presente pode haver influência às próprias estimativas de componentes genéticos aditivos e dominantes, reduzindo a precisão nas estimativas de parâmetros que fazem uso destes (ABREU; RAMALHO; SANTOS, 2002; BARONA *et al.*, 2012). Pela complexidade dos caracteres avaliados, é muito provável que efeitos epistáticos, do tipo aditivo x aditivo, aditivo x dominante e dominante x dominante estejam envolvidos na expressão dos mesmos, já que a cada avanço de geração reduzem-se as frequências de heterozigotos, ocorrendo novos rearranjos entre os alelos nos diferentes *loci*. Uma grande fração desses efeitos são negligenciados pela metodologia em estudo e, com base nestes resultados, acredita-se que esses efeitos sejam cruciais para o processo de estimação e/ou predição. Os valores fenotípicos também podem variar ao longo de anos e locais em decorrência de interações epistáticas por ambiente (BARONA, 2007). Em soja, estudos acerca da produtividade mostraram que interações de efeitos epistáticos com os ambientes dependem do número de *loci* envolvidos na herança de um dado caráter (BARONA *et al.*, 2012).

3.3.3 Estimativas da contribuição dos efeitos aditivos via $m + a$ e correlações

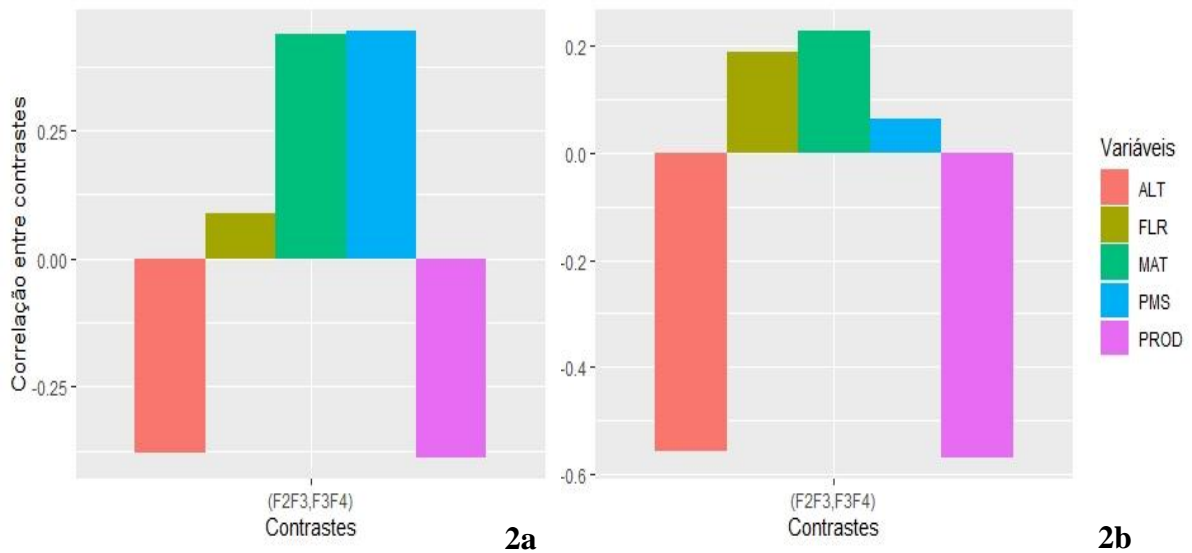
A fim de observar não apenas o comportamento das estimativas de médias ajustadas das progênies segregantes nas três gerações avaliadas, mas também a contribuição dos *loci* fixados por homozigose nessas gerações, foram estimadas médias por contrastes entre as 20 progênies e as estimativas de $m + a$ (Apêndice B).

Para as correlações entre os contrastes (Apêndice C) foi observado que em geral, foram mais baixas ($< 0,50$) para contrastes do que para os BLUEs em ambas as metodologias. Contudo, são por meio das correlações entre contrastes que se pode fazer uma comparação utilizando apenas os *loci* fixados por homozigose (Gráfico 2).

Por meio da metodologia tradicional (2a), os caracteres peso de mil sementes e dias para a maturação apresentaram correlações mais fortes entre os contrastes, 0,44 e 0,43, respectivamente. Mesmo sendo as maiores correlações positivas entre todos os caracteres avaliados, ainda é considerada baixa e exclui a possibilidade de que a geração F_3 seja confiável para se fazer seleção para esses caracteres. Essas baixas correlações indicam que provavelmente os valores de $m + a$ obtidos das progênies F_2/F_3 pouco coincidem com os do contraste F_3/F_4 ao se fazer um *ranqueamento* entre as progênies. Grande parte dos valores estão crescendo de um

contraste para o outro, mas não necessariamente em uma taxa constante. Desse modo, supõem-se que os *loci* fixados por homozigose aumentem de F_3 para F_4 , de modo a ser necessário mais uma geração de avaliação, F_5 . Assim, é possível estabelecer mais um contraste de forma que o nível de coincidência entre os contrastes F_3/F_4 com F_4/F_5 seja forte e positivo e, com isso, F_4 poderia ser considerada uma geração confiável para seleção de progênes segregantes.

Gráfico 2. Correlações de *Spearman* entre as estimativas de $m + a$ (contrastes) das gerações consecutivas F_2/F_3 e F_3/F_4 dos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologias de $m + a$ tradicional (2a) e modificada (2b).



Rocha *et al.* (2015) avaliaram o caráter produtividade em progênes precoces de feijão e estabeleceram contrastes do tipo F_2/F_3 e F_3/F_4 , cujas gerações para cada contraste foram conduzidas em estações distintas (inverno e verão) e em um mesmo local. Os autores observaram baixa correlação (0,13) entre os contrastes e concluíram que as estimativas sofreram grande efeito do ambiente para produtividade e que a identificação de progênes de feijão com base nas estimativas de $m + a$, obtidas em uma única safra, não é precisa. Pouco se encontra na literatura a respeito de correlação entre médias de gerações seguidas, assim como também sobre a observação deste comportamento a respeito de estimativas de $m + a$, até porque a proposta da metodologia já estabelecida é que apenas duas gerações de homozigose sejam necessárias para se estimar os *loci* fixados por homozigose e identificar precocemente progênes com potencial genético para se selecionar linhagens.

É necessário salientar que para obtenção de estimativas de $m + a$, efeitos epistáticos são negligenciados no contraste, mas estão presentes, pois com o aumento da homozigose os efeitos de dominância são dissipados restando somente os efeitos aditivos e epistáticos do tipo aditivo x aditivo (MELO, 2006; VENCONVSKY; RAMALHO; TOLEDO, 2012). Ao longo das gerações em estudo, a presença de epistasia deve estar provavelmente mascarando os reais valores das médias junto aos *loci* fixados por homozigose.

A correlação para dias para a floração foi 0,09, isto é, próxima de zero, e mostra que a associação entre as colocações do *ranking* dos contrastes de F_2/F_3 com F_3/F_4 para esse caráter são quase nulas. Assim, não há confiança ao se afirmar sobre o comportamento das progênes para esse caráter em gerações iniciais de endogamia. Contudo, é válido lembrar que não houve avaliação de dias para a floração para a metodologia tradicional na geração F_4 . Assim, só foi possível estabelecer um contraste para esse caráter.

Ainda na metodologia tradicional (2a) os caracteres altura de planta e produtividade tiveram correlações de -0,38 e -0,39, respectivamente. Isso mostra que grande parte das progênes não mantiveram (pelos contrastes) uma tendência de crescimento de F_2/F_3 para F_3/F_4 , nestes caracteres. Dessa forma, valores de $m + a$ que antes eram altos em F_2/F_3 se tornaram baixos em F_3/F_4 , o que alterou o posto que cada contraste tomou nos dois *rankeamentos*.

Para a metodologia modificada (Gráfico 2b), o comportamento das correlações para cada caráter foi bastante similar a metodologia tradicional. As correlações quando positivas também foram baixas e refletem as mesmas considerações que foram feitas para metodologia tradicional. Em 2b, a maior correlação foi apenas de 0,23 para dias para a maturação, mostrando que o *rankeamento* dos contrastes em F_3/F_4 pouco cresceu quando em associação com F_2/F_3 . Mesmo sendo positiva essa correlação é considerada muito fraca. Os demais caracteres atingiram correlações ainda menores que essa. Peso de mil sementes teve correlação próxima de zero, demonstrando falta de associação do crescimento no *rankeamento* de contrastes em F_2/F_3 com o de F_3/F_4 , em que o peso de mil sementes pouco se alterou ao longo das gerações iniciais avaliadas, como aconteceu para dias para a floração na metodologia tradicional. Altura de plantas e produtividade tiveram correlações negativas de -0,56 e -0,57, respectivamente. Para estes caracteres a falta de associação na tendência de crescimento destas tem o mesmo comportamento que na metodologia tradicional.

Em resumo, fazendo-se um paralelo entre as duas metodologias, baixas correlações entre os dois contrastes foram observadas em ambas. As correlações que se mantiveram positivas, mesmo fracas, são dos mesmos caracteres quando se correlacionou os BLUEs (1a e 1b), da mesma forma que aqueles caracteres que apresentaram correlações negativas.

Como já mencionado, foram feitas algumas etapas a mais do que é solicitado na metodologia tradicional de modo a assegurar resultados equivalentes a modificação. Sendo assim, diferenças encontradas entre uma metodologia e outra podem ser resultado da ineficácia de correção dos efeitos pelas testemunhas. Em linhas gerais, tanto os comportamentos que eram esperados como os não esperados, observados na metodologia tradicional, se mantiveram ao ser feita a modificação do método. Com o uso de todas as etapas requeridas e as adicionais desde a metodologia tradicional e o cumprimento das premissas experimentais necessárias, aproxima-se a nova metodologia proposta de uma possível validação.

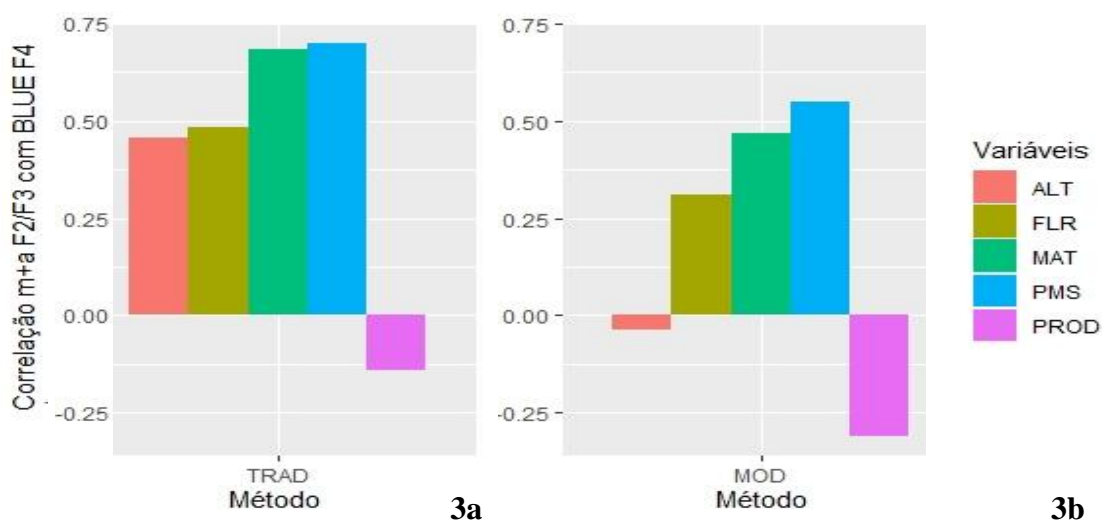
Segundo informações da empresa responsável por ceder as progênies, muitas delas foram oriundas de pais distantes geneticamente, o que resulta em segregação acentuada presente em gerações iniciais para alguns caracteres. Isso faz com que demore mais para que os *loci* sejam fixados a cada autofecundação (MUNIZ, 2007). Diante disso, reitera-se que em programas de melhoramento, a identificação de progênies segregantes com potencial genético superior é uma das etapas mais importantes, porém altamente complicada. Entre as várias metodologias já existentes para essa finalidade, o que é proposto neste trabalho trata-se de uma alternativa que facilite a seleção precoce, principalmente com relação a experimentação, obtendo junto a esta modificação resultados confiáveis.

Após correlacionar estimativas de média ajustadas via BLUEs entre gerações e estimativas de $m + a$ via contrastes entre gerações, foi proposto correlacionar o contraste das duas primeiras gerações (F_2/F_3) a cada metodologia com os BLUEs médios (tradicional + modificada) da última avaliada (F_4) (Apêndice C) a fim de se observar uma possível predição de comportamento dos caracteres em F_4 .

No Gráfico 3a observa-se que nas correlações dos contrastes F_2/F_3 com os BLUEs de F_4 da metodologia tradicional os valores superiores e positivos foram para dias para a maturação e peso de mil sementes, 0,69 e 0,70, respectivamente. Nesse comparativo nota-se que parte dos valores dos contrastes F_2/F_3 crescem positivamente em sentido e intensidade dos BLUEs das progênies de F_4 , no entanto, esta correlação não é forte o suficiente para se afirmar

a respeito do comportamento desta última geração. Para a predição da geração F_4 dos dois caracteres em questão, a metodologia tradicional foi mais informativa via correlação de BLUEs entre gerações (Gráfico 1a).

Gráfico 3. Correlações de *Spearman* entre as estimativas de $m + a$ das gerações F_2/F_3 e BLUEs médios (tradicional + modificada) da geração F_4 dos caracteres dias para floração (FLR), dias para maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD). Estimativa de $m + a$ tradicional (3a) e modificada (3b).



Os demais caracteres com correlações positivas, altura de planta (0,46) e dias para a floração (0,49), apresentaram valores que também foram inferiores a comparação estabelecida entre os BLUEs de F_2 ou F_3 com F_4 (Gráfico 3). Produtividade de grãos apresentou correlação negativa (-0,14) e acompanha a mesma linha de raciocínio. Neste sentido, é possível afirmar que os contrastes entre F_2 e F_3 não são bons preditores do comportamento de estimativas de média em F_4 para metodologia tradicional. De certa forma esse comportamento já era esperado, pois não se observou altas correlações entre os BLUEs ao longo das gerações.

No geral, as correlações dos contrastes F_2/F_3 com as estimativas de média das progênes em F_4 para metodologia modificada (Gráfico 3b) apresentaram magnitudes que também não têm consistência informativa para prever o comportamento dos BLUEs em F_4 . Da mesma forma, para a proposta da modificação da metodologia tradicional, correlações entre os BLUEs (Gráfico 1b) mostram resultados mais informativos. Assim, mesmo não conseguindo resultados satisfatórios de correlação, as metodologias demonstraram concordância de

comportamento para quase todos os caracteres, com exceção de altura de planta que teve a correlação positiva de 0,46 na metodologia tradicional e de -0,04 na metodologia modificada.

3.3.4 Capacidade preditiva via correlações entre estimativas de médias (BLUEs) e valores genotípico (BLUPs)

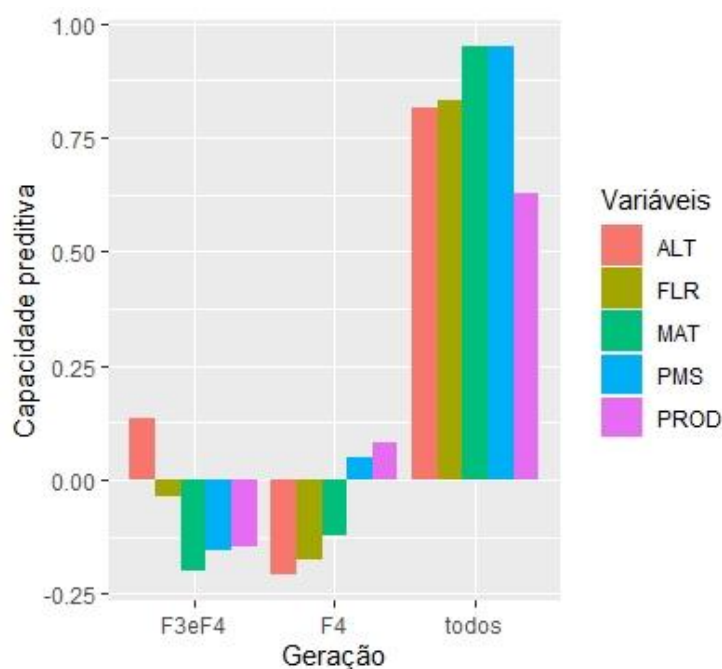
Por fim, simularam-se cenários em que ao se excluir propositalmente os BLUEs de uma ou duas gerações, foram realizadas predições por meio de BLUPs destas gerações excluídas, utilizando os BLUEs da(s) geração(ões) não excluída(s) ponderados por uma matriz de parentesco (informações referentes a genealogia de cada progênie). Essa matriz foi usada na tentativa de aumentar a acurácia das estimativas, ou capacidade preditiva, por meio da inserção de efeitos aditivos assim como captar parte das interações epistáticas do tipo aditivo por aditivo.

Predições via BLUP proporcionam informações relevantes, imparciais, acuradas e não viesadas dos valores genotípicos para serem utilizados na identificação de genótipos superiores. O uso de BLUPs em gerações precoces permite a maximização da probabilidade de seleção dos melhores indivíduos (SILVA; VIDAL NETO; DOVALE, 2017). A partir dos cenários propostos foram gerados os valores de correlações médias (metodologia tradicional + modificada) presentes no Apêndice D e representadas graficamente na Gráfico 4. Para o primeiro cenário, com base nos BLUEs de F₂, foram gerados BLUPs para F₃ e F₄, cuja correlação foi positiva apenas para o caráter altura de planta (0,14) e negativa para os demais, mostrando que os dados fenotípicos de F₂ não são bons preditores para F₃ e F₄. Em um segundo cenário, foram utilizados os BLUEs de F₂ e F₃ e gerados os BLUPs para F₄, em que se obtiveram resultados positivos, mas muito baixos para peso de mil sementes e produtividade de 0,05 e 0,08, respectivamente. Os caracteres altura de planta, dias para a floração e dias para a maturação obtiveram correlações negativas de -0,21, -0,18 e -0,12, respectivamente. Isso isenta dessa vez a capacidade preditiva das médias estimadas de F₂ e F₃ para valores genotípicos da geração F₄.

Por fim, para o último cenário proposto utilizaram-se todas as gerações e matriz de parentesco para prever os próprios BLUPs. As correlações BLUEs/BLUPs apresentaram uma capacidade preditiva elevada para dias para a maturação e peso de mil sementes, com valores de 0,95, mostrando que as observações fenotípicas de cada geração junto a uma matriz de parentesco das progênies utilizadas são boas preditores dos valores genotípicos das mesmas. O comportamento desse cenário com alta magnitude de correlação já era esperado, pois para esse

foi fornecido o máximo possível de informações para predição. Segundo Resende (2007) e Cruz, Regazzi e Carneiro (2012) o procedimento maximiza a correlação entre os valores preditos e os verdadeiros valores (reais) a partir de dados fornecidos.

Gráfico 4. Correlações de *Spearman* para os cenários propostos (BLUE F₂, BLUP F₃ e F₄; BLUE F₂ e F₃, BLUP F₄; BLUE F₂, F₃ e F₄; BLUP F₂, F₃ e F₄) dos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$.



Os caracteres altura de planta e dias para a floração apresentaram estimativas um pouco inferiores de 0,82 e 0,83, respectivamente. Mesmo com o uso de todo o conjunto de observações fenotípicas juntamente com a matriz de parentesco, o coeficiente de correlação BLUEs/BLUPs para produtividade de grãos foi de 0,63.

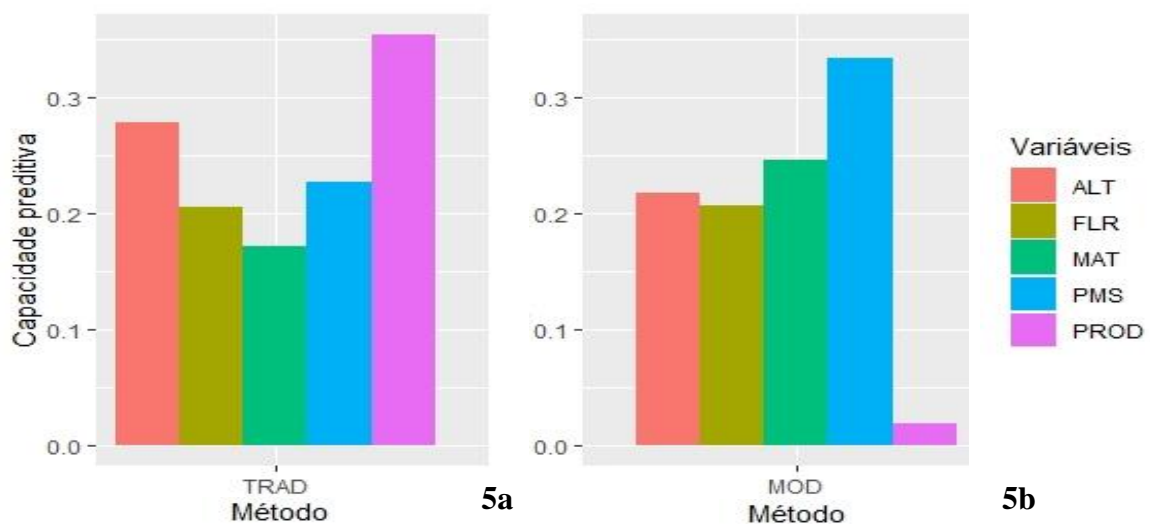
Assim, entende-se que a complexidade para predição desse caráter se deva a presença de efeitos gênicos não-aditivos cruciais e que são muitas vezes negligenciados ou capturados com menor nível de acurácia pelos modelos genéticos, inclusive os usados aqui neste estudo. Interações epistáticas do tipo aditivo x dominante e dominante x dominante não foram absorvidas pelo modelo mesmo com o uso da matriz de parentesco e devem apresentar papel preponderante nas predições desses caracteres de alta complexidade. Em geral, para estas predições, mesmo com o uso da matriz de parentesco, os resultados não são bons, o efeito aditivo parece ajudar um pouco, mesmo assim a correlação não é linear, ainda há presença de

efeitos não aditivos como desvios de dominância que ainda estão sendo perdido ao longo das autofecundações.

Estudos dessa natureza poderão ser repetidos no futuro com uso de progênies de pais genotipados, de preferência com plataformas de alto rendimento. Assim, será possível construir matrizes (*kernels*) de efeitos aditivos e não-aditivos e, possivelmente, obter contrastes e acurácias maiores as observadas no presente estudo. Isso porque estudos têm mostrado que com o uso dessas matrizes é possível absorver nas estimativas partes da variância genotípica normalmente não captadas com o uso exclusivo de dados fenotípicos (LYRA *et al.*, 2018; MATIAS *et al.*, 2019).

Analisando o enfoque do método tradicional e sua modificação, os resultados obtidos mostraram que, em média, as correlações dos valores fenotípicos com os genotípicos preditos, alcançados na análise anterior (Gráfico 4) nos três cenários, apresentaram baixíssimos valores como representados nos gráficos da Gráfico 5. No entanto, na modificação proposta, para essas capacidades preditivas os caracteres se comportaram de forma diferente. Dias para a maturação e peso de mil sementes mostraram melhor capacidade preditiva, 0,25 e 0,33, respectivamente. Talvez isso tenha ocorrido porque apresentam fixação alélica mais rápida que os outros que se destacaram na metodologia tradicional. Nessa comparação, dias para a floração teve correlação semelhante nas duas metodologias, de 0,21.

Gráfico 5. Correlações de *Spearman* para capacidade preditiva nos caracteres dias para a floração (FLR), dias para a maturação (MAT), altura de planta (ALT), peso de mil sementes (PMS) e produtividade (PROD) via metodologias de $m + a$ tradicional e modificada.



Por todos os resultados e discussão anteriores nota-se que é válida a aplicação da modificação do método, o que facilitaria extremamente a experimentação para a mesma. Porém, não se trata apenas de modificar uma metodologia laboriosa, mas também se alcançar bons resultados, permitindo que o objetivo seja cumprido: prever o potencial genético de progênes em gerações precoces de homozigose para se praticar seleção.

Em linhas gerais, os resultados para o formato modificado da metodologia não foram satisfatórios, o que é comprovado com as correlações superiores apresentadas nos BLUEs em relação aos contrastes, ou seja, os BLUEs comportaram-se como bons “preditores” da estimação de outras médias de forma menos inconsistente que os contrastes de $m + a$. No entanto, os resultados para o formato tradicional da metodologia também não foram satisfatórios, o que faz acreditar que houve uma certa falha na concepção do método, por negligenciar efeitos gênicos de suma importância para a predição, que podem estar inflacionando ou até mesmo capitalizando os resultados das estimativas de $m + a$.

4 CONCLUSÃO

De forma geral, as estimativas de metodologia de $m + a$ não apresentam resultados convergentes ao longo do processo de homozigosidade, seja pela sua forma tradicional ou pela modificação proposta. Por sinal, ambas apresentam resultados semelhantes para a grande maioria dos caracteres estudados. Entre estes, a seleção precoce funcionou melhor para dias para a maturação e peso de mil sementes.

Médias ajustadas via BLUEs são mais informativas em gerações precoces de soja que os contrastes obtidos pelas estimativas de $m + a$, seja pela metodologia tradicional ou modificada.

A incorporação da matriz de parentesco não foi suficiente para elevar a capacidade preditiva entre as gerações.

REFERÊNCIAS

- ABIOVE. ASSOCIAÇÃO BRASILEIRA DAS INDÚSTRIAS DE ÓLEOS VEGETAIS. **Cadeia produtiva de oleaginosas e biodiesel (Estatística)**. ABIOVE: São Paulo. 2019. Disponível em: <http://abiove.org.br/estatisticas/>. Acesso em: 29 jul. 2019.
- ABREU, A. F. B.; RAMALHO, M. A. P.; SANTOS, J. B. Prediction of seed-yield potential of common bean populations. **Genetics and Molecular Biology**, Ribeirão Preto, v. 25, n. 3, p. 323-327, 2002.
- ALAMBERT, Marcelo Rodrigues. **Estimação estocástica de parâmetros produtivos da soja: uso do modelo PPDSO em um estudo de caso em Piracicaba/SP**. 2010. Dissertação (Mestrado em Agroenergia) – Escola de Economia de São Paulo, Fundação Getúlio Vargas, São Paulo, 2010.
- ALLARD, R.W. **Princípios do melhoramento genético de plantas**. São Paulo: Edgard Blucher, 381p., 1971.
- ALMEIDA, L. A.; KIIHL, R. A. S.; MIRANDA, M. A. C.; CAMPELO, G. J. A. 1999. **Melhoramento da soja para regiões de baixas latitudes**. In: QUEIRÓZ, M. A. de; GOEDERT, C. O.; RAMOS, S. R. R. (Orgs.). Recursos genéticos e melhoramento de plantas para o Nordeste brasileiro. Petrolina: EMBRAPA Semi-Árido, 1999. cap. 5, p. 73-88.
- AMORIM, F. A.; HAMAWAKI, O. T.; SOUSA, L. B.; LANA, R. M. Q.; HAMAWAKI, C. D. L. Época de semeadura no potencial produtivo de soja em Uberlândia-MG. **Semina Ciências Agrárias**, Londrina, v. 32, n. 4, p. 1793-1802, 2011.
- ARAÚJO, Mayara Lucyanne Santos. **Dinâmica espaço-temporal da cultura de soja na região do Mapitoba, Brasil (1990-2015)**. 2018. Dissertação (Mestrado em Geociências Aplicadas e Geodinâmica) – Instituto de Geociências, Universidade de Brasília, Brasília, 2018.
- BARBOSA-PRESTES, M. M.; FEDERIZZI, L. C.; MILACH, S. C. K.; MARTINELLI, J. A. Genetic control of partial resistance of crown rust in oat (*Avena sativa* L.). **Ciência Rural**, Santa Maria, v. 38, n. 2, p. 308-314, 2008.
- BARONA, Marco Antonio Azevedo. **Epistasia e interação epistasia por locais para a produção de grãos em soja**. 2007. Tese (Doutorado em Agronomia). Escola Superior de Agricultura Luiz de Queiroz – Universidade de São Paulo, Piracicaba, 2007.
- BARONA, M. A. A., COLOMBARI FILHO, J. M., SANTOS, V. D. S., GERALDI, I. O. Epistatic effects on grain yield of soybean [*Glycine max* (L.) Merrill]. **Crop Breeding and Applied Biotechnology**, Viçosa, v. 12, n. 4, p. 231-236, 2012.
- BERNARDO, R. Correlation between tests cross performance of lines at early and last selfing generation. **Theoretical Applied Genetics**, Berlin, v. 82, p.17-21, 1991.
- BERNARDO, R. On the effectiveness of early generation selection in self-pollinated crops. **Crop Science**, Madison, v. 43, n. 4, p. 1558-1560, 2003.

- BIZARI, Eduardo Henrique. **Índices de seleção para caracteres agronômicos em populações segregantes de soja**. 2014. Dissertação (Mestrado em Agronomia) - Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, 2014.
- BONATO, E. R.; BONATO, A. L. V. **A soja no Brasil: história e estatística**. Londrina: Embrapa CNPSo, 1987. 61p. (EMBRAPA-CNPSo. Documentos, 21).
- BORÉM, A.; ALMEIDA, G. D. **Plantas Geneticamente Modificadas**. 1. Visconde do Rio Branco: Suprema, 2011. 390p
- BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: Editora UFV, 2013. 969p.
- BORÉM, A.; MIRANDA, G. V. M. **Melhoramento de plantas**. Viçosa: Editora UFV, 3 ed. 2013. 523p.
- BUENO, L. C. S.; MENDES, A. N. G.; CARVALHO, S. P. **Melhoramento de plantas: princípios e procedimentos**, Lavras: UFLA, 2001, 282p.
- CAMPOS, M. C. Modernização da agricultura, expansão da soja no Brasil e as transformações sócio espaciais no Paraná. **Revista Geografar**, Curitiba, v. 6, n. 1, p. 161-191, 2011.
- CARRÃO-PANIZZI, M. C.; SILVA, J. B. **Soja na alimentação humana: qualidade na produção de grãos com valor agregado**. In: CONGRESO DE LA SOJA DEL MERCOSUR - MERCOSOJA, 5., 2011, Rosário. ... Resumos Rosário: Asociación de la Cadena de la Soja Argentina, 2011. p. 1-3
- CARVALHO, Agnaldo Donizeti Ferreira. **Avaliação da eficiência do teste precoce no melhoramento genético de soja**. 2008. Tese (Doutorado em Agronomia), Escola Superior de Agricultura Luiz de Queiroz – Universidade de São Paulo, Piracicaba, 2008.
- CARVALHO, A. D. F.; GERALDI, I. O.; SANTOS, V. S. Avaliação de progênies F_{2:4} e F_{4:6} de soja e perspectivas do uso de teste precoce para a produção de grãos. **Bragantia**, Campinas, v. 68, n. 4, p. 857-861, 2009.
- CARVALHO, B. C. R. A soja na Bahia. **Revista Bahia Agrícola**, Salvador, v. 5, n. 2, p. 14-15, 2002.
- CARVALHO, Izabela Mendes. **Evolução dos processos de registro e de proteção de cultivares no Brasil**. 2010. Dissertação (Mestrado em Ciência e Tecnologia de Sementes), Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas, Pelotas, 2010.
- CERA, Jossana Ceolin. **Simulação do desenvolvimento, crescimento e produtividade de soja em condições climáticas atuais e futuras para o Rio Grande do Sul**. 2016. Tese (Doutorado em Engenharia Agrícola) - Centro de Ciências Rurais, Universidade Federal de Santa Maria, Santa Maria, 2016.
- CIABOTTI, S.; SILVA, A. C. B. B.; JUHASZ, A. C. P.; MENDONÇA, C. D.; TAVANO, O. L.; MANDARINO, J. M. G.; GONÇALVES, C. A. A. Chemical composition, protein profile, and isoflavones content in soybean genotypes with different seed coat colors. **International Food Research Journal**, Selangor, v. 23, n. 2, p. 621-629, 2016.

CONAB. COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira de grãos**. Safra 2018/19 – Décimo levantamento. v. 6. 2019. 113p. Disponível em: <https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos>. Acesso em: 29 jul. 2019.

COSTA, J. A.; MANICA, I. **Cultura da soja**. Porto Alegre: Editora Evangraf, 1996. 233p.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. 4. ed. Viçosa, MG: UFV, 2012. v. 1, 514 p

CRUZIO, Akemi Suzuki. **Eficiência da seleção precoce para tamanho de grão e seu efeito em outros caracteres de feijão-caupi**. 2014. Dissertação (Mestrado em Genética e Melhoramento), Universidade Federal do Piauí, Teresina, 2014.

CUNHA, R. C. C.; ESPÍNDOLA, C. J. A Relevância do progresso técnico na consolidação da cadeia produtiva da soja no Sul do estado do Maranhão (Brasil). **Geografia**. Londrina, v. 25, n. 1, p. 87-106, 2016.

DONÇA, Marques Cachisso Bambo. **Seleção precoce para caracteres dos grãos no melhoramento do feijão caupi**. 2012. Dissertação (Mestrado em Genética e Melhoramento de Plantas), Universidade Federal de Lavras, Lavras, 2012.

EMBRAPA. EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. EMBRAPA **Unidade de Execução de Pesquisa – UEP/Balsas**. EMBRAPA Cocais: São Luis. 2016. Disponível em: <https://www.embrapa.br/cocais/infraestrutura/unidade-de-execucao-de-pesquisa>. Acesso em: 18 maio. 2019.

EMBRAPA. EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. EMBRAPA **Soja, cultivos (História da soja)**. EMBRAPA CNPSO: Londrina. 2017. Disponível em: <https://www.embrapa.br/web/portal/soja/cultivos/soja1/historia>. Acesso em: 18 jun. 2017.

ERICKSON, D. **Practical handbook of soybean processing and utilization**. Champaign: AOCS Press, 1995. 584p.

FEDERER, W. T. **Experimental design: theory and application**. New York: MacMillan, 1955. 544 p.

FEHR, W. R. 1978. **Breeding**. In: soybean physiology, agronomy and utilization. Norman, A. G. (ed). Academic Press, New York, USA. 120-155 pp.

FREITAS, Edjane Gonçalves de. **Uso de informações de parentesco e modelos mistos para avaliação e seleção de genótipos de cana-de-açúcar**. 2013. Tese de Doutorado (Doutorado em Ciências) -Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2019.

FREY, K. J. The Use of F₂ lines in predicting the performance of F₃ selections in two barley crosses. **Agronomy Journal**, [s.l.] v. 46, n. 12, p. 541-544, 1954.

FRIZON, C. N. T.; PERUSSELLO, C. A.; STURION, J. A.; FRACASSO, A. F.; HOFFMANN-RIBANI, R. Stability of beverages of yerba mate (*Ilex paraguariensis*) with soy. **Nutrition & Food Science**, [s.l.] v. 45, n. 3, p. 467-478, 2015.

GAZZONI, D. L.; DALL'AGNOL, A. **Paralelo entre a soja no mundo e no Brasil**. In: SOJA: quebrando recordes. p. 37-59. Sorocaba: CESB, 2018.

HIRAKURI, M. H.; LAZZAROTTO, J. J. **Evolução e perspectivas de desempenho econômico associadas com a produção de soja nos contextos mundial e brasileiro**. Londrina: Embrapa-Soja, 2011. 68p.

IMMER, F. R. Relation between yielding ability and homozygosity in barley crosses. **Journal of the American Society the Agronomy**, New York, v. 33, n. 3, p. 200–206, 1941.

KLEINPAUL, J. A.; FILHO, A. C.; ALVES, B. M.; BURIN, C.; NEU, I. M. M.; SILVEIRA, D. L.; SIMÕES, F. M. Tamanho de amostra para estimação da média de caracteres de milheto em épocas de avaliação. **Revista Brasileira de Milho e Sorgo**, Sete Lagoas, v. 16, n. 2, p. 251-262, 2017.

LEITE, D. S.; MUNHOZ, L. L. Biotecnologia e melhoramento das variedades vegetais: cultivares e transgênicos. **Veredas do Direito**, Belo Horizonte, v. 10, n. 19, p. 23-44, 2013.

LI, Z.; NELSON, R. L. Genetic diversity among soybean accessions from three countries measured by RAPDs. **Crop Science**, Madison, v. 41, n. 4, p. 1337-1347, 2001.

LÚCIO, A. D.; STORCK, LINDOLFO. O manejo das culturas interfere no erro experimental. **Pesquisa Agropecuária Gaúcha**, Porto Alegre, v. 5, n. 2, 1999.

LYRA, D. H.; GRANATO, Í. S. C.; MORAIS, P. P. P.; ALVES, F. C.; DOS SANTOS, A. R. M.; YU, X.; GUO, T, Yu. J; FRITSCHÉ-NETO, R. Controlling population structure in the genomic prediction of tropical maize hybrids. **Molecular breeding**, [s.l.], v. 38, n. 10, p. 126, 2018.

MAPA. MINISTÉRIO DA AGRICULTURA PECUÁRIA E ABASTECIMENTO. **Projeções do Agronegócio**. Brasil 2018/19 a 2028/29 Projeções de longo prazo. 2019. 126p. Disponível em: <http://www.agricultura.gov.br/assuntos/politica-agricola/todas-publicacoes-de-politica-agricola/projecoes-do-agronegocio/projecoes-do-agronegocio-2018-2019-2028-2029/view>. Acesso em: 29 jul. 2019a.

MAPA. MINISTÉRIO DA AGRICULTURA PECUÁRIA E ABASTECIMENTO. **Registro nacional de cultivares – RNC**. MAPA: Brasília. 2019. Disponível em: http://sistemas.agricultura.gov.br/snpc/cultivarweb/cultivares_registradas.php. Acesso em: 31 jul. 2019b.

MATEUS, R. P. G.; DA SILVA, C. M. Avanços biotecnológicos na cultura da soja. **Campo Digital**, Campo Mourão, v. 8, n. 2, p. 23-27, 2013.

MATHER, Kenneth; JINKS, J. L. **Introduction to Biometrical Genetics**. 3rd edition. 1982.

MATHER K (1949) **Biometrical Genetics: The Study of Continuous Variation**. London UK: Methuen.

MATOS, Renata Fernandes. **Potencial genético de progênies de feijão-caupi para a obtenção de genótipos de porte ereto e ciclo precoce**. 2016. Dissertação (Mestrado em Fitotecnia) - Departamento de Fitotecnia, Universidade Federal do Ceará, Fortaleza, 2016.

MATIAS, F. I.; ALVES, F. C.; MEIRELES, K. G. X.; BARRIOS, S. C. L.; DO VALLE, C. B.; ENDELMAN, J. B.; FRITSCHÉ-NETO, R. On the accuracy of genomic prediction models considering multi-trait and allele dosage in *Urochloa* spp. interspecific tetraploid hybrids. **Molecular Breeding**, [s.l.], v. 39, n. 7, p. 100, 2019.

MELO, Carlos Lázaro Pereira. **Melhoramento de feijão do tipo carioca: avaliação de populações segregantes e uso de marcadores moleculares visando resistência a patógenos**. 2006. Tese (Doutorado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa, 2006.

MEOTTI, G. V.; BENIN, G.; SILVA, R. R.; BECHE, E.; MUNARO, L. B. Épocas de semeadura e desempenho agrônômico de cultivares de soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 47, n. 1, p. 14-21, 2012.

MILADINOVIC, J.; BURTON, J. W.; TUBIC, S. B.; MILADINOVIC, D.; DJORDJEVIC, V.; DJUKIC, V. Soybean breeding: comparison of the efficiency of different selection methods. **Turkish Journal of Agriculture and Forestry**, Ankara, v. 35, n. 5, p. 469-480, 2011.

MIYASAKA, S.; MEDINA, J. C. **A soja no Brasil**. Campinas: Instituto de Tecnologia de Alimentos, 1981. 1062p.

MUÑOZ F.; SANCHEZ L. **breedR: statistical methods for forest genetic resources analysts. R package version 0.7–16** 2014. Disponível em: <https://github.com/famuvie/breedR>. Acesso em: 18 ago. 2019.

MUNIZ, Franco Romero Silva. **Análise da variabilidade genética em populações segregantes de soja**. 2007. Tese (Doutorado em Agronomia) - Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual de São Paulo, Jaboticabal, 2007.

OLIVEIRA, Antonio Edson Rocha. **Perfil de expressão e análise filogenética dos genes da proteína desacopladora mitocondrial durante o desenvolvimento e estresse em soja [Glycine max (L.) Merr.]**. 2015. Tese (Doutorado em Bioquímica) – Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, 2015.

OLIVEIRA, Renata de Souza. **Adsorção de contaminantes de biodiesel sobre fibras de bagaço modificadas em sua superfície**. 2017. Dissertação (Mestrado em Biocombustíveis) – Universidade Federal de Uberlândia, Uberlândia, 2017.

PASTINA, Maria Marta. **Mapeamento de QTLs e estudo da interação entre QTLs, ambientes e cortes em cana-de-açúcar, usando a abordagem de modelos mistos**. 2010. Tese (Doutorado em Ciências) – Universidade de São Paulo - Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2010.

PORTO, M. O.; PAULINO, M. F.; VALADARES FILHO, S. D. C.; SALES, M. F. L.; LEÃO, M. I.; COUTO, V. R. M. Fontes suplementares de proteína para novilhos mestiços em recria em pastagens de capim-braquiária no período das águas: desempenho produtivo e econômico. **Revista Brasileira de Zootecnia**, Viçosa, v. 38, n. 8, p. 1553-1560, 2009.

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing; 2015. Disponível em: <http://www.R-project.org/>. Acesso em: 20 ago. 2019

RAMALHO, A. P. R.; SANTOS, J. B.; ABREU, A. F. B.; NUNES, J. A. R. **Aplicações da genética quantitativa no melhoramento de plantas autógamas**. 1 ed. Lavras: Ed. UFLA, 2012.

RAMALHO, M. A. P.; SANTOS, J. B. dos; ZIMMERMANN, M. J. de O. **Genética quantitativa em plantas autógamas: aplicações ao melhoramento do feijoeiro**. Goiânia: UFG, 1993. 271p.

RAMALHO, M. A. P.; ABREU, A. F. B.; SANTOS, J. B. Melhoramento de plantas autógamas. In Nass, L. L.; Valois, A. C. C.; Melo, I. S.; Valadares-Inglis, M. C. (eds.) **Recursos genéticos e melhoramento de plantas**. Fundação MT, Rondonópolis, 2001, p. 201-230.

RAMALHO, M. A. P.; FERREIRA, D. F.; OLIVEIRA, A. C. de. **Experimentação em genética e melhoramento de plantas**. 3.ed. Lavras: UFLA, 2012. 305p.

RESENDE, M. D. V. de. **Métodos estatísticos ótimos na análise de experimentos de campo**. Colombo: Embrapa Florestas, 57 p. (Documentos 100). 2004.

RESENDE, M. D. V. **SELEGEN-REML/BLUP: sistema estatístico e seleção genética computadorizada via modelos lineares mistos**. Colombo: Embrapa Florestas, 2007.

RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, Goiânia, v. 37, n. 3, p. 182-194, 2007.

RIBEIRO, A. S.; TOLEDO, J. F. F.; RAMALHO, M. A. P. Selection strategies of segregant soybean populations for resistance to Asian rust. **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, n. 11, p. 1452-1459, 2009.

ROCHA, G. S.; CARNEIRO, J. E. S.; CARNEIRO, P. C. S.; POERSCH, N. L.; LIMA, M. S.; SILVA, L. C. Estratégias de predição e efeitos de ambientes na avaliação de populações segregantes de feijão. **Revista Ceres**, Viçosa, v. 62, n. 5, p. 438-445, 2015.

ROMANATO, Fernanda Neves. **Correlações fenotípicas e genotípicas, adaptabilidade e estabilidade em genótipos de soja**. 2013. Dissertação (Mestrado em Agronomia) – Universidade Federal de Uberlândia, Uberlândia, 2013.

ROSADO, A. M.; ROSADO, T. B.; ALVES, A. A.; LAVIOLA, B. G.; BHERING, L. L. Seleção simultânea de clones de eucalipto de acordo com produtividade, estabilidade e adaptabilidade. **Pesquisa Agropecuária Brasileira**, Brasília, v. 47, n. 7, p. 966-973, 2012.

ROSA FILHO, G.; CARVALHO, M. D. P.; ANDREOTTI, M.; MONTANARI, R.; BINOTTI, F. F. D. S.; GIOIA, M. T Variabilidade da produtividade da soja em função de atributos físicos de um Latossolo Vermelho distroférrico sob plantio direto. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 33, n. 2, p. 283-293, 2009.

ROSAL, C. J. D. S.; RAMALHO, M. A. P.; GONÇALVES, F. M. A.; ABREU, A. D. F. B. Seleção precoce para a produtividade de grãos no feijoeiro. **Bragantia**, Campinas, v. 59, n. 2, p. 189-195, 2000.

- SANTOS, O. S. **A cultura da soja-1: Rio Grande do Sul, Santa Catarina, Paraná.** Rio de Janeiro: Globo, 1988. 299p.
- SEDIYAMA, T.; PEREIRA, M. G.; SEDIYAMA, C. S.; GOMES, J. L. L. **Cultura da soja: Parte I.** Viçosa: Imprensa Universitária, 1985. 96p.
- SEDIYAMA, T.; SILVA, F.; BORÉM, A. **Soja: do plantio a colheita.** Viçosa: Editora UFV, 2015. 333p.
- SEDIYAMA, T. **Tecnologias de produção e usos da soja.** Londrina: Editora Mecenaz, 2009. 314p.
- SILVA, Lelizangela Carvalho. **Estratégias de condução de populações segregantes no melhoramento genético do feijoeiro.** 2009. Tese (Doutorado em Genética e Melhoramento de Plantas), Universidade Federal de Viçosa, Viçosa, 2009.
- SILVA, T. P.; VIDAL NETO, F. C.; DOVALE, J. C. Prediction of genetic gains with selection between and within S2 progenies of papaya using the REML/Blup analysis. **Pesquisa Agropecuária Brasileira**, Brasília, v. 52, n. 12, p. 1167-1177, 2017.
- SILVA, F. L.; BORÉM, A.; SEDIYAMA, T.; LUDKE, W. H. **Melhoramento da soja.** Viçosa: Ed. UFV, 2017, p. 563.
- SOUSA, Camila Campelo. **Mapeamento associativo para tolerância a altas temperaturas em germoplasma exótico de soja (*Glycine max*).** 2015. Tese (Doutorado em Ciências) - Escola Superior de Agricultura "Luiz De Queiroz, Universidade de São Paulo, Piracicaba, 2015.
- SPEHAR, C. R. Melhoramento da soja para as baixas latitudes dos Cerrados. **Pesquisa Agropecuária Brasileira**, Brasília, v. 29, n. 8, p. 1167-1180, 1994.
- STORCK, L. et al. **Experimentação vegetal.** Santa Maria : UFSM, 2000. 198p
- SZNITOWSKI, A. M. Fontes de conhecimento/tecnologia para o agronegócio da soja em Mato Grosso. **Revista UNEMAT de Contabilidade**, Nova Mutum, v. 6, n. 11, p. 236-259, 2017.
- TOLEDO, M. R.; TANCREDI, F. D.; SEDIYAMA, T.; RIBEIRO JÚNIOR, J. I.; REIS, M. S. Remoção do meristema apical e adensamento em plantas de soja visando sua utilização no método descendente de uma única semente. **Acta Scientiarum Agronomy**, Maringá, v. 31, n. 1, p. 113-119, 2009.
- USDA. UNITED STATES DEPARTMENT OF AGRICULTURE. **World Agricultural Supply and Demand Estimates (WASDE-590).** USDA: WASHINGTON, DC. 2019. 40p. Disponível em: <https://usda.library.cornell.edu/concern/publications/3t945q76s?locale=en>. Acesso em: 29 jul. 2019.
- VAZQUEZ, A. I.; BATES, D. M.; ROSA, G. J. M.; GIANOLA, D.; WEIGEL, K. A. Technical note: an R package for fitting generalized linear mixed models in animal breeding. **Journal of Animal Science**, [s.l.], v. 88, n. 2, p. 497-504, 2010.

VENCOVSKY, R. Herança quantitativa. In: PATERNIANI, E. (Ed.). **Melhoramento e produção do milho no Brasil**. Piracicaba: Escola Superior de Agricultura Luiz de Queiroz, 1987. p. 122-201.

VENCOVSKY, R.; BARRIGA, P. Genética biométrica no fitomelhoramento. Ribeirão Preto. **Revista Brasileira de Genética**, 496p, 1992.

VENCOVSKY, R.; RAMALHO, M. A. P.; TOLEDO, F. H. R. B. Contribution and perspectives of quantitative genetics to plant breeding in Brazil. **Crop Breeding and Applied Biotechnology**, Viçosa, v. 12, n. SPE, p. 07-14, 2012.

VIANA, Laryssa Ferreira. **Estimativas de $m+a'$ e d como indicadores do potencial de híbridos de milho para extração de linhagens**. 2007. Dissertação (Mestrado em Genética e Melhoramento de Plantas), Universidade Federal de Lavras, Lavras, 2007.

VIANA, L. F.; DE SOUZA, J. C.; MACHADO, J. C.; LIMA, J. L. Predição de médias de linhagens obtidas de híbridos simples de milho (*Zea mays* L.). **Ciência e agrotecnologia (Impressa)**, Lavras, v. 33, n. especial, p. 1999-2004, 2009.

WICKHAM, H. “Reshaping Data with the reshape Package.” **Journal of Statistical Software** v. 21, n. 12, p. 1-20, 2007.

WICKHAM, H. ggplot2: **Elegant Graphics for Data Analysis**. New York: Springer, 2016, p.256.

ZITO, R. K.; ARANTES, N. E.; FRONZA, V.; SÁ, M. E. L.; PÁDUA, G. P.; ZANETTI, A. L.; JUHÁSZ, A. C. P. Soja em Minas Gerais. **Informe Agropecuário**, Belo Horizonte, v.3, n.25, p.16-21, 2011.

APÊNDICE A – VALORES DE BLUES ESTIMADOS PARA AS 20 PROGÊNIES UTILIZADAS NAS TRÊS GERAÇÕES DE HOMOZIGOSIDADE AVALIADAS.

i) Valores de BLUEs dos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ tradicional.

Progênie	ALT			FLR			MAT			PMS			PROD		
	F ₂	F ₃	F ₄	F ₂	F ₃	F ₄	F ₂	F ₃	F ₄	F ₂	F ₃	F ₄	F ₂	F ₃	F ₄
1	0,96	0,99	0,91	53	55	54	119	122	120	167	174	165	4408	3879	3646
2	1,01	1,10	0,98	50	52	53	128	125	129	147	146	149	3553	3655	3221
3	1,07	1,04	0,95	55	54	56	127	128	126	146	144	142	3915	2983	2949
4	0,91	0,92	0,84	55	52	54	117	117	115	169	167	158	3511	3320	2929
5	0,91	0,97	0,88	56	55	54	116	119	118	168	172	170	3251	4011	3156
6	0,90	0,94	0,88	45	50	52	116	117	117	160	166	160	3541	3836	3074
7	0,80	0,88	0,87	40	44	42	109	112	116	145	149	144	2063	2751	2531
8	0,91	0,90	0,88	47	49	51	108	110	111	146	152	145	2846	3396	2452
9	0,94	0,97	0,95	51	50	49	123	122	126	161	151	152	2883	3391	3004
10	0,97	0,92	0,90	56	54	54	117	117	116	154	148	145	3925	3964	3247
11	1,01	1,09	0,89	44	43	52	125	121	123	162	164	154	3608	3810	3172
12	0,88	0,88	0,85	55	53	55	115	115	113	143	138	137	3562	3712	2961
13	0,86	0,90	0,82	55	53	55	118	119	121	154	148	150	3749	3781	2805
14	0,91	0,91	0,84	50	54	53	118	116	116	149	144	144	3369	3747	2296
15	1,04	1,02	0,97	43	46	49	126	123	124	151	156	149	4154	3831	3610
16	0,95	0,96	0,89	52	53	54	119	119	118	164	163	159	4329	3889	2433
17	0,95	1,00	0,87	53	53	55	131	130	126	153	148	145	4251	4059	2457
18	0,92	0,95	0,87	54	55	52	116	118	118	166	157	152	3537	3650	3480
19	0,86	0,85	0,92	51	49	48	113	114	118	149	150	135	3108	2964	3127
20	1,00	0,94	0,93	54	51	53	118	117	116	158	156	147	3969	3612	3458

ii) Valores de BLUEs dos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ modificada

Progênie	ALT			FLR			MAT			PMS			PROD		
	F ₂	F ₃	F ₄	F ₂	F ₃	F ₄	F ₂	F ₃	F ₄	F ₂	F ₃	F ₄	F ₂	F ₃	F ₄
1	1,01	0,95	1,00	47	52	51	118	120	118	157	157	161	4662	3791	3844
2	1,13	1,01	1,12	50	54	50	125	122	126	135	138	150	3477	3046	4262
3	1,18	0,96	1,11	47	54	52	125	121	127	137	136	141	3696	3055	3747
4	0,96	0,91	0,91	48	53	52	115	115	116	155	156	158	3908	3574	3495
5	1,00	0,91	0,99	48	52	50	114	116	117	153	166	164	4182	3379	3599
6	0,94	0,93	0,97	45	51	50	115	116	118	154	164	162	4191	3386	3812
7	0,92	0,84	0,98	41	49	40	102	108	116	139	139	143	3382	2637	3559
8	1,05	0,95	0,97	44	49	48	106	110	110	155	161	141	4493	4243	2901
9	0,96	0,91	1,06	47	51	44	119	119	127	149	153	159	4028	3142	2966
10	1,00	1,00	1,00	47	52	52	113	116	117	143	145	147	4144	3944	4175
11	1,07	0,99	1,01	43	49	52	116	118	124	151	149	150	3746	3248	4092
12	0,98	0,93	0,94	48	54	52	114	115	113	143	142	141	3661	4130	3762
13	0,92	0,85	0,93	50	52	52	116	116	118	144	143	149	3979	3930	3723
14	0,96	0,94	0,96	47	51	52	115	115	116	141	145	138	3675	4068	2983
15	1,08	1,00	1,08	43	49	44	118	119	124	148	160	150	4299	3625	4304
16	1,02	0,99	0,95	46	52	51	116	118	116	149	158	159	4215	4048	3800
17	0,98	0,97	0,94	48	51	52	124	122	127	150	163	153	3704	4350	3347
18	0,98	0,95	0,97	47	54	53	112	116	119	156	157	160	4026	3363	3849
19	0,91	0,84	0,98	47	50	43	111	113	122	152	154	144	3992	3535	3650
20	0,98	1,01	0,97	49	53	51	116	117	117	144	150	149	4030	3862	4423

**APÊNDICE B – ESTIMATIVAS DE $M + A$ PARA AS 20 PROGÊNIOS UTILIZADAS NAS TRÊS GERAÇÕES DE
HOMOZIGOSIDADE AVALIADAS.**

i) Valores de contrastes dos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ tradicional.

Progênie	ALT		FLR		MAT		PMS		PROD	
	F ₂ /F ₃	F ₃ /F ₄	F ₂ /F ₃	F ₃ /F ₄	F ₂ /F ₃	F ₃ /F ₄	F ₂ /F ₃	F ₃ /F ₄	F ₂ /F ₃	F ₃ /F ₄
1	1,02	0,84	57	53	125	118	182	155	3349	3414
2	1,19	0,86	53	55	122	133	145	152	3758	2787
3	1,01	0,85	52	59	128	125	141	141	2051	2914
4	0,93	0,76	48	56	118	113	164	149	3130	2537
5	1,04	0,78	54	54	123	116	176	167	4770	2302
6	0,98	0,82	55	54	118	116	172	155	4131	2313
7	0,96	0,86	47	40	116	121	154	139	3440	2311
8	0,89	0,85	50	53	112	112	158	138	3946	1508
9	1,00	0,94	49	48	121	129	141	153	3899	2617
10	0,88	0,88	53	53	117	115	142	142	4003	2530
11	1,17	0,69	43	60	117	124	166	145	4012	2534
12	0,88	0,81	50	57	116	110	132	136	3862	2211
13	0,95	0,74	51	56	119	123	142	152	3814	1829
14	0,92	0,76	57	52	114	116	139	144	4124	845
15	1,01	0,91	49	52	120	125	161	142	3509	3388
16	0,97	0,82	54	54	118	118	161	156	3449	977
17	1,04	0,75	53	58	129	123	143	143	3866	855
18	0,98	0,78	56	50	120	119	149	146	3764	3309
19	0,83	0,98	47	46	115	122	152	120	2821	3289
20	0,89	0,91	48	55	116	115	155	138	3255	3303

ii) Valores de contrastes dos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ modificada.

Progênie	ALT		FLR		MAT		PMS		PROD	
	F ₂ /F ₃	F ₃ /F ₄	F ₂ /F ₃	F ₃ /F ₄	F ₂ /F ₃	F ₃ /F ₄	F ₂ /F ₃	F ₃ /F ₄	F ₂ /F ₃	F ₃ /F ₄
1	0,89	1,06	57	50	123	116	157	166	2921	3897
2	0,89	1,23	57	47	118	130	140	163	2615	5478
3	0,73	1,27	60	51	118	132	134	146	2413	4440
4	0,85	0,91	57	51	116	117	157	159	3241	3415
5	0,83	1,07	56	49	119	117	180	161	2577	3818
6	0,92	1,01	57	49	117	119	175	159	2582	4238
7	0,75	1,11	56	30	114	123	140	147	1891	4482
8	0,85	0,99	53	47	114	109	166	122	3994	1559
9	0,86	1,20	55	37	120	134	158	165	2256	2790
10	1,01	0,99	57	51	119	118	148	149	3743	4406
11	0,91	1,02	54	54	120	131	147	150	2750	4936
12	0,88	0,96	60	49	117	111	141	140	4600	3393
13	0,78	1,02	54	51	116	120	142	154	3881	3516
14	0,92	0,97	55	53	116	116	149	131	4461	1898
15	0,92	1,16	56	39	120	129	172	141	2952	4983
16	0,96	0,91	57	50	120	114	167	160	3882	3552
17	0,96	0,92	54	54	121	132	176	144	4996	2344
18	0,92	1,00	60	52	121	121	158	163	2699	4335
19	0,76	1,13	54	36	115	131	157	133	3077	3765
20	1,05	0,93	57	50	119	116	156	147	3694	4984

APÊNDICE C – CORRELAÇÕES DE *SPEARMAN* ENTRE A TRÊS GERAÇÕES DE HOMOZIGOSIDADE AVALIADAS.

i) Correlações de *Spearman* entre os BLUEs das gerações F₂, F₃ e F₄ dos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ tradicional.

Correlação	ALT	FLR	MAT	PMS	PROD
F ₂ /F ₃	0,81	0,87	0,95	0,85	0,65
F ₂ /F ₄	0,68	0,74	0,88	0,84	0,33
F ₃ /F ₄	0,58	0,75	0,88	0,89	0,18

ii) Correlações de *Spearman* entre os BLUEs das gerações F₂, F₃ e F₄ dos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$ modificada.

Correlação	ALT	FLR	MAT	PMS	PROD
F ₂ /F ₃	0,65	0,81	0,96	0,86	0,38
F ₂ /F ₄	0,75	0,61	0,75	0,62	0,00
F ₃ /F ₄	0,36	0,62	0,73	0,65	-0,18

iii) Correlações de *Spearman* entre os contrastes das gerações F₂, F₃ e F₄ dos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD) via metodologias de $m + a$ tradicional e modificada.

(F₂/F₃, F₃/F₄)	ALT	FLR	MAT	PMS	PROD
Tradicional	-0,38	0,09	0,43	0,44	-0,39
Modificada	-0,56	0,19	0,23	0,06	-0,57

iv) Correlações de *Spearman* entre as estimativas de $m + a$ das gerações F₂/F₃ e BLUEs médio (tradicional + modificada) da geração F₄ dos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD). Estimativas de $m + a$ tradicional e modificada.

	ALT	FLR	MAT	PMS	PROD
Tradicional	0,45	0,68	0,49	0,70	-0,14
Modificada	-0,03	0,47	0,31	0,55	-0,31

APÊNDICE D – CENÁRIOS PREDITIVOS

i) Correlações de *Spearman* para os cenários propostos (BLUE F₂, BLUP F₃ e F₄; BLUE F₂ e F₃, BLUP F₄; BLUE F₂, F₃ e F₄; BLUP F₂, F₃ e F₄) nos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia de $m + a$.

PREDIÇÕES	ALT	FLR	MAT	PMS	PROD
F ₃ e F ₄	0,14	-0,04	-0,20	-0,15	-0,15
F ₄	-0,21	-0,18	-0,12	0,05	0,08
TODOS	0,82	-0,83	0,95	0,95	0,63

ii) Correlações de *Spearman* para capacidade preditiva nos caracteres altura de planta (ALT), dias para a floração (FLR), dias para a maturação (MAT), peso de mil sementes (PMS) e produtividade (PROD) via metodologia $m + a$ tradicional e modificada.

PREDIÇÕES	ALT	FLR	MAT	PMS	PROD
Tradicional	0,28	0,21	0,17	0,23	0,36
Modificada	0,22	0,21	0,25	0,33	0,02