

MURILO VIOTTO DEL CONTE

**ESTRATÉGIAS DE SELEÇÃO EM ESTÁGIOS PRECOSES DO MELHORAMENTO
DE SOJA COM O USO DE MODELOS MISTOS**

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

Orientador: Felipe Lopes da Silva

Coorientadores: Marcos Deon Vilela de Resende
Pedro Crescêncio Souza Carneiro

**VIÇOSA - MINAS GERAIS
2021**

Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa

T

Del Conte, Murilo Viotto, 1992-
D331e Estratégias de seleção em estágio precoces do
2021 melhoramento de soja com o uso de modelos mistos / Murilo
Viotto Del Conte. – Viçosa, MG, 2021.
1 tese eletrônica (66 f.): il. (algumas color.).

Inclui apêndices.

Orientador: Felipe Lopes da Silva.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2021.101>

Modo de acesso: World Wide Web.

1. Soja - Seleção - Índices. 2. Soja - Melhoramento genético. 3. BLUP. 4. Análise combinatória. I. Universidade Federal de Viçosa. Departamento de Agronomia. Programa de Pós-Graduação em Genética e Melhoramento. II. Título.

CDD 22. ed. 633.342

MURILO VIOTTO DEL CONTE

**ESTRATÉGIAS DE SELEÇÃO EM ESTÁGIOS PRECOSES DO MELHORAMENTO
DE SOJA COM O USO DE MODELOS MISTOS**

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

APROVADA: 23 de julho de 2021

Assentimento:



Murilo Viotto Del Conte
Autor



Felipe Lopes da Silva
Orientador

A minha filha Livia Maciel Del Conte (in memoriam)

AGRADECIMENTOS

A Deus, pela graça da vida e por me guiar pelo caminho que me trouxeram até aqui.

Aos meus pais, Luzia Aparecida Viotto Del Conte e Anderson Del Conte; e ao meu irmão, Pedro Viotto Del Conte pelo amor incondicional, conselhos, apoio e confiança em todas as ocasiões.

À minha esposa Aline Biggi Maciel Del Conte pelo amor, companheirismo e cuidado investido em nossa relação.

À Universidade Federal de Viçosa, bem como ao Programa de Pós-Graduação em Genética e Melhoramento e ao Departamento de Fitotecnia pelo acolhimento e pela oportunidade da realização do título de doutorado em Genética e Melhoramento.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de estudos.

A Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelos recursos investidos nos projetos em que estive envolvido e que proporcionaram o meios necessários para desenvolver esta tese

Ao professor Felipe Lopes da Silva, pelo tempo dedicado às orientações, suporte para desenvolvimento dos trabalhos, conselhos e amizade.

Aos Professores Marcos Deon Vilela de Resende e Pedro Crescêncio Souza Carneiro pela coorientação e pelos ensinamentos.

Aos Professores Fabyano Fonseca e Silva, Guilherme da Silva Pereira, Luiz Alexandre Peternelli e Paulo Eduardo Teodoro pela participação na banca de defesa e contribuições para este trabalho.

Aos estagiários, funcionários e demais colegas de pós-graduação do Programa Soja pela ajuda na condução dos experimentos, ideias compartilhadas, discussões técnico-científicas e amizade.

*Conhecimento não é aquilo que você sabe, mas o
que você faz com aquilo que você sabe.*

(Aldous Huxley)

RESUMO

DEL CONTE, Murilo Viotto, D.Sc., Universidade Federal de Viçosa, julho de 2021. **Estratégias de seleção em estágios precoces do melhoramento de soja com o uso de modelos mistos.** Orientador: Felipe Lopes da Silva. Coorientadores: Marcos Deon Vilela de Resende e Pedro Crescêncio Souza Carneiro.

A produtividade de grãos é o caractere de maior interesse para a cultura da soja. Por se tratar de um caractere quantitativo, recomenda-se o uso da seleção recorrente. Este processo tem o objetivo de promover o acúmulo gradativo de alelos favoráveis ao caractere melhorado nas linhagens obtidas após ciclos sucessivos de seleção e intercruzamento, visando o lançamento de novos cultivares ou a seleção de novos genitores. Para reduzir o tempo para alcançar estes objetivos, ou, reduzir o “ciclo de melhoramento” visando antecipar o acesso ao ganho genético proporcionado pelo programa, uma das alternativas é realizar a seleção em estágios precoces do programa. Para que esta estratégia seja efetiva, é preciso adotar metodologias que permitam a seleção acurada nestas fases. Neste sentido, a presente tese teve como objetivo: a) avaliar diferentes metodologias de condução de populações segregantes F_2 , incluindo estratégias guiadas pelo BLUP Individual Simulado Modificado (BLUPISM), a fim de avaliar qual destas metodologias promove melhores ganhos genéticos e se assemelha mais ao BLUPI que é entendida como a metodologia ótima para seleção de indivíduos em populações segregantes; b) investigar os benefícios do emprego das informações de genitores nos modelos estatísticos e do índice de seleção genitores + progênies para ranqueamento de progênies $F_{2:4}$. Para tanto, foram desenvolvidas 16 populações de soja em esquema de dialelo parcial com dois grupos de genitores. Na geração F_2 , foi realizado um ensaio em blocos casualizados com três repetições em Cajuri – MG na safra 2016/2017. Cada parcela foi constituída de quatro linhas de cinco metros, das quais as duas centrais foram colhidas em bulk para estimação da produtividade de grãos e as linhas laterais foram utilizadas para coleta de plantas individuais, nas quais foram avaliadas a massa de grãos por planta. Na geração $F_{2:4}$ foram realizados dois experimentos nas localidades de Capinópolis – MG e Sorriso – MT em delineamento de blocos aumentados com oito blocos e parcelas de duas linhas de cinco metro lineares, as quais foram todas colhidas para estimativa da produtividade de grãos das progênies. Os resultados obtidos na geração F_2 indicaram que o uso do BLUPISM com a adoção dos valores genéticos de CEC e CGA de ambos os grupos de pais é uma ferramenta útil e foi a que proporcionou ganhos genéticos e número de indivíduos selecionados mais semelhantes aos obtidos pelo BLUPI em populações segregantes F_2 de soja. Para a avaliação das progênies $F_{2:4}$, os resultados obtidos mostraram que

considerar as informações de genitores no modelo estatístico proporciona a possibilidade de se realizar a seleção de progênes $F_{2:4}$ de soja com maior acurácia e que estes efeitos são mais benéficos quando há variância significativa entre os genitores. Quanto ao índice de seleção genitores + progênes, este apresentou melhores estimativas de acurácia, quando comparadas à seleção realizada com base nos valores genotípicos (BLUPs) das progênes $F_{2:4}$.

Palavras-chave: BLUPI. BLUPISM. Índice de seleção. Capacidade combinatória. Seleção precoce.

ABSTRACT

DEL CONTE, Murilo Viotto, D.Sc., Universidade Federal de Viçosa, July, 2021. **Early selection strategies in soybean breeding using mixed models.** Advisor: Felipe Lopes da Silva. Co-advisors: Marcos Deon Vilela de Resende and Pedro Crescêncio Souza Carneiro.

Grain yield is of greatest interest to soybean crop. As it is a quantitative trait it is recommended to use recurrent selection. This process aims to promote the gradual accumulation of favorable alleles to the improved trait in the lines obtained after successive selection and intermating cycles, aiming to obtaining new cultivars or the selection of new parents. To reduce the time to meet these aims or reduce the “breeding cycle” in order to anticipate access to the genetic gain provided by the program, it is necessary to carry out selection in the early stages of the program. For this strategy to be effective, it is necessary to adopt methodologies that allow a precise selection in these phases. In this sense, this thesis aimed to: a) evaluate different methodologies for conducting segregating populations F_2 , including strategies guided by BLUPISM, in order to assess which of these methodologies promotes better genetic gains and is similar to BLUPI, which is understood as optimal methodology for selection in segregating groups; b) investigate the benefits of using information about parents in statistical models and the selection index parents + progenies for ranking progenies $F_{2:4}$. For this, 16 soybean populations were developed in a partial diallel scheme with two groups of parents. In the F_2 generation, a randomized block trial was carried out with three replications in Cajuri - MG in the 2016/2017 season. Each plot consisted of four rows of five meters, of which the two centrals were collected in bulk to estimate grain yield and the lateral rows were used to collect individual plants, in which the mass of grains per plant was evaluated. In the $F_{2:4}$ generation, two field trials were carried out in the Capinópolis - MG and Sorriso - MT in an augmented block design with eight blocks and plots of two lines of five linear meters, which were all harvested to estimate the grain yield of the progenies. The results obtained in the F_2 generation indicated that the use of BLUPISM with the adoption of genetic values of CEC and CGA of both groups of parents is a useful tool and provided the genetic gains and number of selected individuals most similar to BLUPI in soybean segregating populations F_2 . For the evaluation of $F_{2:4}$ progenies, the results showed that considering the information of the parents in the statistical model offers the possibility to carry out a more accurate selection of $F_{2:4}$ soybean progenies. Moreover, the results showed that parents effects are more beneficial when there is significant variance among parents. As

for the selection with parents + progenies index, it presented better accuracy, when compared to that selection based on the genotypic values (BLUPs) for the F_{2:4} progenies.

Palavras-chave: BLUPI. BLUPISM. Selection index. Combinatory ability. Early selection.

LISTA DE TABELAS

CAPÍTULO 1

Tabela 1 – Populações obtidas por meio de cruzamentos entre dois grupos de genitores em esquema de dialelo parcial.	24
Tabela 2 – Valores preditos utilizados e descrição das modalidades de seleção entre e dentro de populações segregantes F ₂ com base valores genotípicos (VG) ou amostragem simulada (AS) adotadas no trabalho.	28
Tabela 3 – Estimativas de parâmetros genéticos e fenotípicos para o caractere massa de grãos por planta em plantas individuais de populações F ₂ de soja.....	30
Tabela 4 – Ganhos genéticos percentuais obtidos para massa de grãos por planta em todas as modalidades adotadas para a seleção em populações segregantes F ₂	32
Tabela 5 – Número de indivíduos selecionados (IS = 15%) dentro de cada população F ₂ sob a adoção das modalidades E2, E5, E6, E7 e E8 e respectivos valores genotípicos das populações em modalidades que contemplaram a seleção entre.	33
Tabela 6 – Número de indivíduos selecionados em cada população de acordo com as estratégias E1 (BLUPI) e E8 (BLUPISM com o emprego de $g_{cec} + g_{cgc1} + g_{cgc2}$) e correlação para as intensidades de seleção de 15%, 10% e 5%.	34
Tabela 7 – Estimativas de parâmetros genéticos e fenotípicos para o caractere produtividade de grãos (kg ha ⁻¹) em progênies F _{2:4} de soja.	36
Tabela 8 – Valores dos efeitos genotípicos das populações nas gerações F ₂ e F _{2:4} para as análises individuais e conjunta.	37

CAPÍTULO 2

Tabela 1 – Populações obtidas por meio de cruzamentos entre dois grupos de genitores em esquema de dialelo parcial.	46
Tabela 2 – Estimativas de parâmetros genéticos e fenotípicos para produtividade de grãos (kg ha ⁻¹) das progênies F _{2:4} para os ensaios conduzidos em Sorriso – MT e Capinópolis – MG sob ajuste do modelo reduzido (Eq. 1).	50

Tabela 3 - Estimativas de parâmetros genéticos e fenotípicos para produtividade de grãos (kg ha ⁻¹) das progênies F _{2:4} para os ensaios conduzidos em Sorriso – MT e Capinópolis – MG sob ajuste do modelo completo (Eq. 2).	52
Tabela 4 – Estimativas de acurácia mínima, média e máxima para o modelo completo e o índice de seleção genitores + progênies.....	55
Tabela 5 – Índice de coincidência entre as progênies selecionadas entre as predições obtidas sob ajuste do modelo completo e do índice de seleção genitores + progênies para os ensaios conduzidos e Sorriso – MT e Capinópolis – MG, considerando as intensidades de seleção de 10, 15 e 20%.	55

LISTA DE FIGURAS

CAPÍTULO 1

Figura 1 – Fluxograma de coleta e análise de dados do ensaio de campo das populações F₂..25

CAPÍTULO 2

Figura 1 – Distribuição de dos valores genotípicos (BLUP) das progênes agrupados por seus respectivos genitores do grupo 1 para as localidades de Sorriso – MT (A) e Capinópolis – MG54

SUMÁRIO

INTRODUÇÃO GERAL.....	14
REFERÊNCIAS	18
CAPÍTULO 1.....	20
RESUMO.....	21
1. INTRODUÇÃO.....	22
2. MATERIAL E MÉTODOS.....	23
2.1 Obtenção do material genético.....	23
2.2 Condução dos experimentos a campo e análises dos dados	24
2.2.1 Populações F ₂	24
2.2.2 Progênies F _{2:4}	27
3. RESULTADOS E DISCUSSÃO	29
3.1 Modalidades de condução e seleção das populações segregantes	29
3.2 Comparação entre as populações indicadas em F ₂ e F _{2:4}	35
4. CONCLUSÕES	38
5. REFERÊNCIAS BIBLIOGRÁFICAS	39
CAPÍTULO 2.....	42
RESUMO	43
1 INTRODUÇÃO.....	44
2 MATERIAL E MÉTODOS.....	46
2.1 Obtenção do material genético.....	46
2.2 Condução dos experimentos a campo e avaliação das progênies F _{2:4} a campo	47
2.3 Análises estatísticas.....	47
3. RESULTADOS E DISCUSSÃO	50
3.1 Análises estatísticas para o modelo reduzido.....	50
3.2 Análises estatísticas para o modelo completo.....	51
3.3 Modelo reduzido vs modelo completo.....	53
3.1 Modelo completo vs Índice de seleção genitores + progênies.....	54
4. CONCLUSÕES	57
5 REFERÊNCIAS BIBLIOGRÁFICAS	57
CONCLUSÕES GERAIS	61
APÊNDICES	62
Apêndice 1	62
Apêndice 2	64

INTRODUÇÃO GERAL

A soja [*Glycine Max* (L.) Merr.] é uma das culturas mais cultivadas no mundo e uma das principais fontes de óleo e proteína vegetal. Os principais interesses sobre a cultura é o emprego dos grãos da leguminosa na alimentação animal e humana. O grão de soja apresenta uma composição favorável à alimentação e é formado por aproximadamente 40% de proteína e 20% de óleo, além de outras porções, tais como: algumas vitaminas e compostos polifenólicos (Silva et al., 2017; Sedyama et al., 2015).

Nos últimos 45 anos a produtividade média de grãos da cultura no Brasil evoluiu de 1748 kg ha⁻¹ para 3517 kg ha⁻¹. (CONAB, 2021). O melhoramento genético que visa a busca por maiores produtividades, figura como um dos grandes responsáveis por esta evolução (Todeschini et al., 2019; Mauad et al., 2010).

A produtividade de grãos é um caractere governado por muitos genes e por isso é considerado de natureza quantitativa (Cruz, 2005). Para que os melhoristas possam alcançar sucesso no melhoramento de caracteres desta natureza, o programa deve ser orientado para o acúmulo gradativo de alelos favoráveis, o que pode ser alcançado com o uso da seleção recorrente (Borém et al, 2017; Silva et al, 2017 e Ramalho et al, 2012). No melhoramento de espécies autógamas, como no caso da soja, a seleção recorrente é realizada em três etapas que compõe um ciclo: condução das populações segregantes por sucessivas gerações de endogamia, escolha das melhores linhagens e a recombinação destas (Borém et al, 2017; Silva et al, 2017, Ramalho et al, 2012; Hallauer e Darrah, 1985). Ao longo destes ciclos que ocorrem no programa, o melhorista tem dois diferentes objetivos: o lançamento de novas variedades e a escolha de genitores para a formação das novas populações segregantes.

Um ciclo dos programas de melhoramento de soja clássico pode levar até 12 anos e pode ser compreendido como o tempo entre o cruzamento entre genitores e o lançamento de uma nova variedade ou a adoção de uma linhagem em novos blocos de cruzamento (Borém et al., 2017). Uma das maneiras conhecidas para melhorar a eficiências de programas de melhoramento é reduzir o ciclo de melhoramento, que é o tempo compreendido entre os primeiros cruzamentos e o emprego dos cultivares resultantes destes cruzamentos ou ainda o tempo de “cruzamento a cruzamento” (Coob et al., 2019; Ceccarelli, 2015).

Atualmente, existem diversas técnicas modernas que podem ser utilizadas para reduzir o ciclo de melhoramento da soja para 4 a 7 anos e acelera a obtenção de ganhos genéticos. Tal redução pode ser alcançada pela adoção de berçários em regiões de baixas latitudes ou no

hemisfério oposto, conduzindo dois ciclos da cultura em situações de chuvas em distribuição bimodal; uso de *single-seeds descent* (SSD) (Goulden, 1939); colheita antecipada de sementes (Hickey, 2019); emprego de ferramentas biotecnológicas tais como a introgressão de QTLs (Ribaut e Hoisington, 1998); e condução de gerações de cruzamento e avanços de geração em ambientes com umidade, temperatura, fotoperíodo e suplementação de CO₂ controlados e o emprego da edição gênica (Nagatoshi e Fujita, 2019). De modo geral, estas estratégias utilizadas para acelerar o melhoramento são conhecidas como “*speed breeding*” e tem sido assuntos de muitos estudos na atualidade.

O emprego da seleção precoce também é uma ferramenta para aumento da eficiência do programa pela redução do ciclo de melhoramento. Esta estratégia, tende a oferecer melhor custo-benefício ao melhorista quando comparado a situações em que o programas de melhoramento executam as primeiras etapas de seleção em gerações avançadas (Ceccarelli, 2015). Com a seleção precoce é possível que programa avance um menor número de populações ou progênies e concentre recursos e esforços na avaliação de progênies promissoras (Silva, et al., 2017, Bernardo, 2010).

A literatura sugere que devido à predominância dos efeitos aditivos apresentados pelos caracteres de interesse agrônômico na cultura da soja, a seleção de indivíduos dentro de populações segregantes F₂ deve derivar progênies mais produtivas nas próximas gerações (Cooper, 1990; Gate et al., 1960). Neste sentido, é comum em programas de melhoramento de espécies autógamas a seleção de indivíduos dentro de populações segregantes nas gerações F₂ ou F₃ (Silva et al., 2017, Bernardo, 2010), sendo que em muitos casos, é realizada a seleção visual entre e dentro de populações segregantes. Tal seleção pode ser realizada com base em aspectos fenotípicos avaliados à campo ou pela tomada de dados e posterior análise estatística, seguindo os procedimentos de preferência do melhorista. Nesta ocasião pode-se escolher por fazer a seleção de um número diferente de plantas em cada população de acordo com seu mérito. Em outras situações o que ocorre é a amostragem das populações, na qual são avançados o mesmo números de plantas em todas as populações.

Mesmo assim, para contornar os possíveis efeitos perturbadores à seleção que podem ser ocasionados pelos loci em heterozigose em fases precoces do programa de melhoramento, se faz necessário lançar mão de metodologias que permitam alcançar boas acurácias de seleção na avaliação das progênies, além de conduzir os ensaios com boa qualidade experimental, sobretudo em fases de populações segregantes, onde ainda se considera cada indivíduo como um genótipo único e diferente dos demais. Para fins de seleção de indivíduos, Resende (2002) afirma que a predição de valores genotípicos utilizando o melhor preditor linear não viesado

(*best linear unbiased predictor* - BLUP) a nível de indivíduo (BLUPI) é a estratégia ótima para a seleção entre e dentro de populações. Tal procedimento considera simultaneamente informações das populações e indivíduos dentro das populações.

Diante da alta demanda operacional requerida para o emprego do BLUPI, cujo qual requer a avaliação de plantas individuais, Resende e Barbosa (2006) propuseram o uso do BLUP individual simulado (BLUPIS). Esta estratégia consiste na seleção baseada apenas nos valores genotípicos das populações ou famílias testadas a campo, sem a necessidade da avaliação em nível de indivíduos. Mais tarde, tendo notado que o BLUPIS tendia a guiar a seleção de número grande de indivíduos dentro de poucas e promissoras famílias, Castro et al. (2016) propuseram modificações na metodologia BLUPIS (BLUPISM - BLUP individual simulado modificado) que distribui de forma mais homogênea o número de indivíduos selecionados dentro das famílias mais promissoras, preservando melhor a variabilidade. Estas metodologias podem auxiliar o melhorista na condução das populações e derivação das progênies.

Outro aspecto a ser considerado na seleção precoce, é a natureza e a qualidade dos dados a serem utilizados para os ajustes dos modelos estatísticos. Nesse contexto, Resende (2015) propôs o uso de um índice de seleção que considera as informações de genitores e progênies (Índice Genitores + progênies). O objetivo deste índice é proporcionar o ranqueamento mais confiável das progênies, mesmo em fases precoces do programa, de acordo com o seu mérito genético ponderado pelas suas acurácias e pelos méritos genéticos e acurácias de seus genitores.

A seleção efetuada contemplando as informações de genitores ou de pedigree tende a ser mais eficiente que a efetuada considerando apenas a informação das progênies (Bauer e León, 2008). Para Duarte e Vencovsky (2001) e Volpato et al. (2020) desconsiderar informações de parentais ou parentesco pode comprometer as estimativas dos componentes de variância, as predições dos valores genéticos e por consequência, o ranqueamento das progênies, conduzindo uma seleção viesada e reduzindo as chances da obtenção de ganhos genéticos.

Diante disso, esta tese teve como objetivo avaliar diferentes metodologias de seleção entre e dentro de populações segregantes F_2 e investigar os benefícios do emprego das informações de genitores nas análises para ranqueamento de progênies $F_{2:4}$ por meio de índice de seleção. Seus objetivos específicos foram:

- Propor e aplicar BLUPISM com modificações na metodologia original apresentada por Castro et al. (2016);

- Comparar os ganhos genéticos previstos pela adoção do BLUPISM na seleção entre e dentro de populações segregantes F_2 frente a outras modalidades empregadas;
- Comparar o número de indivíduos indicados para seleção dentro de cada população pelo BLUPI e BLUPISM em populações F_2 ;
- Comparar as populações indicadas como promissoras na geração F_2 com aquelas que derivaram as melhores progênies para produtividade de grãos na geração $F_{2:4}$.
- Avaliar os efeitos do emprego da informação de genitores na avaliação de progênies $F_{2:4}$ de soja sobre a qualidade experimental;
- Avaliar os efeitos do emprego do índice de seleção genitores + progênies proposto por Resende (2015) sobre a acurácia de seleção de progênies frente à seleção baseada nos melhores preditores lineares não viesados (BLUP – *Best Linear Unbiased Prediction*) para progênies $F_{2:4}$.

Para tanto, esta tese foi dividida em dois capítulos, sendo que no capítulo 1 são apresentados os resultados das comparações entre as diferentes modalidades de condução de populações segregantes F_2 de soja, entre elas o BLUPI e o BLUPISM. No capítulo 2 é apresentado o estudo realizado sobre a aplicação do índice de seleção genitores + progênies na seleção de progênies $F_{2:4}$ de soja frente à seleção tradicional, baseada apenas nos valores genotípicos das progênies.

REFERÊNCIAS

- BAUER, A. M.; LÉON, J. Multiple-trait breeding values for parental selection in self-pollinating crops. **Theoretical and Applied Genetics**, v. 116, n. 2, p. 235–242, jan. 2008.
- BERNARDO, R. N. **Breeding for quantitative traits in plants**. Woodbury, Minn.: Stemma Press, 2010.
- BORÉM, A.; MIRANDA, G. V.; FRITSCHÉ-NETO, R. **Melhoramento de plantas**. 7. ed. Viçosa, MG: Editora UFV, 2017.
- CASTRO, R. D. et al. Selection between and within full-sib sugarcane families using the modified BLUPIS method (BLUPISM). **Genetics and Molecular Research**, v. 15, n. 1, 2016.
- CECCARELLI, S. Efficiency of Plant Breeding. **Crop Science**, v. 55, n. 1, p. 87–97, jan. 2015.
- CONAB. **Série Histórica**. Disponível em: <<http://www.conab.gov.br/info-agro/safras/serie-historica-das-safras/item/7666-soja/>>. Acesso em: 4 jul. 2021.
- COOPER, R. L. Modified Early Generation Testing Procedure for Yield Selection in Soybean. **Crop Science**, v. 30, n. 2, p. 417, 1990.
- DUARTE, J. B.; VENCOVSKY, R. Estimaco e predico por modelo linear misto com ênfase na ordenaco de mdias de tratamentos genticos. **Scientia Agricola**, v. 58, n. 1, p. 109–117, mar. 2001.
- GATES, C. E.; WEBER, C. R.; HORNER, T. W. A Linkage Study of Quantitative Characters in a Soybean Cross ¹. **Agronomy Journal**, v. 52, n. 1, p. 45–49, jan. 1960.
- GOULDEN, C. H. **Problems in plant selection**. Proceedings of the Seventh International Genetics Congress. **Anais...** In: SEVENTH INTERNATIONAL GENETICS CONGRESS. Cambridge: Cambridge Univ. Press, 1939.
- HALLAUER, A. R.; DARRAH, L. L. Compendium of recurrent selection methods and their application. **Critical Reviews in Plant Sciences**, v. 3, n. 1, p. 1–33, jan. 1985.
- HICKEY, L. T. et al. Breeding crops to feed 10 billion. **Nature Biotechnology**, v. 37, n. 7, p. 744–754, jul. 2019.
- MAUAD, M. et al. Influncia da densidade de semeadura sobre caractersticas agronmicas na cultura da soja. **Agrarian**, v. 3, n. 9, p. 175–181, 2010.

- NAGATOSHI, Y.; FUJITA, Y. Accelerating Soybean Breeding in a CO₂-Supplemented Growth Chamber. **Plant and Cell Physiology**, v. 60, n. 1, p. 77–84, 1 jan. 2019.
- RAMALHO, M. A. P. et al. **Aplicações da Genética Quantitativa no Melhoramento de Plantas Autógamas**. Lavras, MG: UFLA, 2012.
- RESENDE, M. D. V. DE. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília, DF: Embrapa Informação Tecnológica, Colombo: Embrapa Florestas, 2002.
- RESENDE, M. D. V. DE; BARBOSA, M. H. P. Selection via simulated individual BLUP based on family genotypic effects in sugarcane. **Pesquisa Agropecuária Brasileira**, v. 41, n. 3, p. 421–429, mar. 2006.
- RIBAUT, J.-M.; HOISINGTON, D. Marker-assisted selection: new tools and strategies. **Trends in Plant Science**, v. 3, n. 6, p. 236–239, jun. 1998.
- SEDIYAMA, T.; SILVA, F. L.; BORÉM, A. **Livro Soja do Plantio à Colheita - UFV**. Viçosa, MG: Editora UFV, 2015.
- SILVA, F. L. DA et al. **Soybean breeding**. New York, NY: Springer Science+Business Media, 2017.
- TODESCHINI, M. H. et al. Soybean genetic progress in South Brazil: physiological, phenological and agronomic traits. **Euphytica**, v. 215, n. 7, p. 124, jul. 2019.
- VOLPATO, L. et al. Selection of inbred soybean progeny (*Glycine max*): an approach with population effect. **Plant Breeding**, v. 137, n. 6, p. 865–872, dez. 2018.
- VOLPATO, L. et al. Inference of population effect and progeny selection via a multi-trait index in soybean breeding. **Acta Scientiarum. Agronomy**, v. 43, p. e44623, 14 ago. 2020.
- WANGA, M. A. et al. Opportunities and challenges of speed breeding: A review. **Plant Breeding**, v. 140, n. 2, p. 185–194, abr. 2021.

**EFICIÊNCIA DO BLUPISM NA SELEÇÃO DE POPULAÇÕES F2 DE SOJA
FRENTE A OUTRAS ESTRATÉGIAS DE SELEÇÃO**

CAPÍTULO 1

RESUMO

A produtividade de grãos é o caractere de maior interesse para a cultura da soja. Devido à predominância dos efeitos aditivos para os caracteres de interesse agrônômico no melhoramento de soja é comum que os melhoristas realizem a fase de seleção em estádios precoces do programa. O BLUP Individual (BLUPI) é considerada a metodologia ótima para seleção nestas situações, porém requer a avaliação de indivíduos um a um e por isso é impraticável em programas de melhoramento, o que pode ser contornado pelo BLUPI simulado modificado (BLUPISM) que considera o mérito genético das populações para guiar a seleção de plantas dentro delas sem a necessidade da avaliação de indivíduo a indivíduos. Diante disto, os objetivos do trabalho foram: a) comparar os ganhos genéticos previstos pela adoção do BLUPISM na seleção em populações segregantes frente a outras sete estratégias de condução das populações; b) comparar o número de indivíduos indicados para seleção dentro de cada população pelo BLUPI e BLUPISM; e c) comparar as populações indicadas como promissoras na geração F_2 com aquelas que derivaram as melhores progênies para produtividade de grãos na geração $F_{2:4}$. Foram desenvolvidas 16 populações de soja em esquema de dialelo parcial com dois grupos de genitores. Na geração F_2 , foi realizado um ensaio em blocos casualizados com três repetições em Cajuri – MG na safra 2016/2017 para avaliação da produtividade de grãos das parcelas e de plantas individuais. Na geração $F_{2:4}$ foram realizados dois experimentos nas localidades de Capinópolis – MG e Sorriso – MT em delineamento de blocos aumentados com oito blocos para avaliação da produtividade de grãos das progênies. Os resultados obtidos indicaram que o uso do BLUPISM com a adoção dos valores genotípicos de CEC e CGC ($g_{cec} + g_{cgc1} + g_{cgc2}$) (estratégia E8) de ambos os grupos de pais é uma ferramenta útil e foi a que proporcionou ganhos genéticos mais semelhantes aos obtidos pelo BLUPI em populações segregantes F_2 de soja. O número de indivíduos selecionados pelo uso de BLUPISM e BLUPI foram semelhantes. Houve alta concordância entre as populações selecionadas em F_2 pelo BLUPISM na estratégia E8 e aquelas indicadas como promissoras na geração $F_{2:4}$.

1. INTRODUÇÃO

A soja [*Glycine max* (L.) Merrill] é uma das principais oleaginosas cultivadas em território nacional e mundial. A série histórica da produtividade de grãos da cultura no Brasil mostra que foram obtidos ganhos lineares ao longo dos anos, devido às evoluções dos implementos agrícolas, melhoras das técnicas de manejo e, sobretudo ao melhoramento genético (CONAB, 2021 e MAUAD et al., 2010).

A produtividade de grãos é um caractere governado por um grande número de genes e por isso é considerado um caractere quantitativo. Para melhorar caracteres desta natureza, o programa deve ter como objetivo o acúmulo gradativo de alelos favoráveis via sucessivos ciclos de seleção e recombinação, o que pode ser alcançado com o uso da seleção recorrente (BORÉM et al, 2017; SILVA et al, 2017 e RAMALHO et al, 2012). No melhoramento de espécies autógamas, a seleção recorrente é realizada em etapas de condução das populações segregantes por sucessivas gerações de endogamia, escolha das melhores linhagens e a recombinação destas. Para estas espécies, como no caso da soja, tem sido comum a adoção da seleção recorrente com o objetivo de realizar a seleção em gerações precoces, o que permite concentrar recursos na condução de populações ou famílias mais promissoras, para posteriormente recombinar as de melhor desempenho para o caractere de interesse. (SILVA et al, 2017)

Alguns estudos sugerem que devido à predominância dos efeitos aditivos apresentados pelos caracteres de interesse agrônômico na cultura da soja, a seleção de indivíduos dentro de populações segregantes F_2 deve derivar progênieis mais produtivas nas próximas gerações (COOPER, 1990; GATE et al., 1960). Neste sentido, é comum em programas de melhoramento de espécies autógamas a seleção de indivíduos dentro de populações segregantes nas gerações F_2 ou F_3 (SILVA et al., 2017, BERNARDO, 2010), sendo que em muitos casos, é realizada a seleção visual entre e dentro, ou apenas dentro de populações segregantes. Tal seleção pode ser realizada com base em aspectos fenotípicos avaliados à campo ou pela tomada de dados e posterior análise estatística, seguindo os procedimentos de preferência do melhorista. Nesta ocasião pode-se escolher por fazer a seleção de um número diferente de plantas em cada população de acordo com seu mérito. Em outras situações o que ocorre é a amostragem das populações, na qual são avançados o mesmo números de plantas em todas as populações.

Para fins de seleção, Resende (2002) afirma que a predição de valores genotípicos utilizando o melhor preditor linear não viesado (*best linear unbiased predictor* - BLUP) a nível de indivíduo (BLUPI) é a estratégia ótima para a seleção entre e dentro de populações. Tal procedimento considera simultaneamente informações das populações e indivíduos dentro das

populações. Entretanto, o autor cita que tal estratégia raramente é adotada pelos programas de melhoramento devido à alta demanda operacional para a avaliação de plantas individuais.

Diante disto, Resende e Barbosa (2006) propuseram o uso do BLUP individual simulado (BLUPIS). Esta estratégia consiste na seleção baseada apenas nos valores genotípicos das populações ou famílias testadas a campo, sem a necessidade da avaliação em nível de indivíduos. Além disto, a metodologia permite indicar o número de indivíduos selecionados dentro de cada população ou família, o número total de progênies a ser avaliadas na próxima geração e excluir populações ou famílias com valores dos efeitos genotípicos negativos para o caractere de interesse. Mais tarde, Castro et al. (2016) propuseram modificações na metodologia BLUPIS (BLUPISM - BLUP individual simulado modificado) a fim de contornar o fato de o BLUPIS tender a selecionar um número grande número de indivíduos em um número pequeno de poucas populações mais promissoras. Nesta nova proposta o número de indivíduos selecionados na população estudada seria computado considerando o número de indivíduos selecionados na população imediatamente superior a ela mesma e não mais à melhor população.

Neste sentido, os objetivos do presente trabalho foram: a) propor e aplicar o BLUPISM com modificações na metodologia original proposta por Castro et al. (2016) na seleção de populações F₂ de soja; b) comparar os ganhos genéticos previstos pela adoção do BLUPISM na seleção entre e dentro de populações segregantes frente a outras sete modalidades de condução das populações; c) comparar o número de indivíduos indicados para seleção dentro de cada população pelo BLUPI e BLUPISM; e d) comparar a coincidência entre as populações indicadas como promissoras na geração F₂ com aquelas que derivaram as melhores progênies para produtividade de grãos na geração F_{2:4}.

2. MATERIAL E MÉTODOS

2.1 Obtenção do material genético

Na safra 2015/16 foram desenvolvidas 16 populações via hibridação artificial em esquema de dialelo parcial em casa de vegetação. O dialelo foi constituído por dois grupos de genitores (cultivares comerciais), sendo o primeiro grupo composto por cinco cultivares com alto potencial produtivo para as regiões do cerrado brasileiro (G1) e o segundo por quatro cultivares com boa performance para teor de óleo (G2), conforme apresentado na tabela 1. As sementes de cada cruzamento foram colhidas separadamente, trilhadas e posteriormente semeadas em casa de vegetação na safra 2016/16 para multiplicação do número de sementes e

obtenção das populações segregantes F_2 . Para tanto, a casa de vegetação foi equipada com iluminação artificial com a finalidade de proporcionar um ambiente adequado ao bom desenvolvimento da cultura, considerando a importância do fotoperíodo para a soja, o qual atua como indutor de florescimento (CÂMARA, 1998).

Tabela 1 – Populações obtidas por meio de cruzamentos entre dois grupos de genitores em esquema de dialelo parcial.

Grupo 1	Grupo 2			
	M 9056 RR	M 8221 RR	P98N71	CONQUISTA
BMX Desafio RR	POP4	POP10	POP19	POP30
M7739 IPRO	-	-	POP20	POP31
P98Y30 RR	POP5	POP11	POP21	POP32
M8349 IPRO	-	-	POP22	POP33
TMG 132 RR	POP6	POP12	POP23	POP34

Simultaneamente à condução e avaliação do experimento das populações segregantes F_2 na safra 2016/17, foi realizada a seleção individual de plantas, dando origem a progênies $F_{2:3}$. A seleção foi realizada considerando características de sanidade, acamamento e, sobretudo, o ajuste do grupo de maturidade relativa (ALLIPRANDINI et. al, 2009) para as regiões sojícolas do cerrado brasileiro. As progênies $F_{2:3}$ foram semeadas na safra 2017/17, as quais foram derivadas de plantas selecionadas individualmente na geração anterior. Ao atingir a senescência, ou seja, o estágio R8 (FEHR e CAVINESS, 1977), cada parcela foi colhida individualmente em bulk, o que deu origem a um total de 218 progênies $F_{2:4}$. Durante tal processo, as populações POP22, POP 23 e POP 33 não seguiram no programa devido ao baixo desempenho para caracteres de interesse ou disponibilidade de sementes para as gerações seguintes, totalizando 13 populações para as próximas etapas.

2.2 Condução dos experimentos a campo e análises dos dados

2.2.1 Populações F_2

O experimento foi implantado no Município de Cajuri – MG (20°45'51.0" S, e 42°49'26.9" O e 647 m de altitude), em delineamento em blocos ao acaso com três repetições, 16 tratamentos (populações F_2 , Tabela 1) e nove testemunhas (genitores utilizados para obter

as populações). As parcelas foram constituídas por quatro linhas de 5 m lineares espaçadas por 0,5 metro com densidade de 260.000 plantas ha⁻¹. Em cada parcela, foram colhidas 24 plantas das linhas laterais da parcela e trilhadas separadamente para quantificação da massa de grãos por planta. As duas linhas centrais de cada parcela foram colhidas, trilhadas separadamente e a produtividade de grãos foi convertida para kg ha⁻¹ e corrigido para 13% de umidade. Os dados de massa de grãos por planta foram utilizados para obter a predição dos valores genéticos dos indivíduos avaliados e computar as estimativas de ganho para as modalidades de condução de população segregante estudadas, enquanto os dados de produtividade de grãos foram utilizados para ranqueamento das populações e aplicação das modalidades de seleção, conforme mostra a Figura 1.

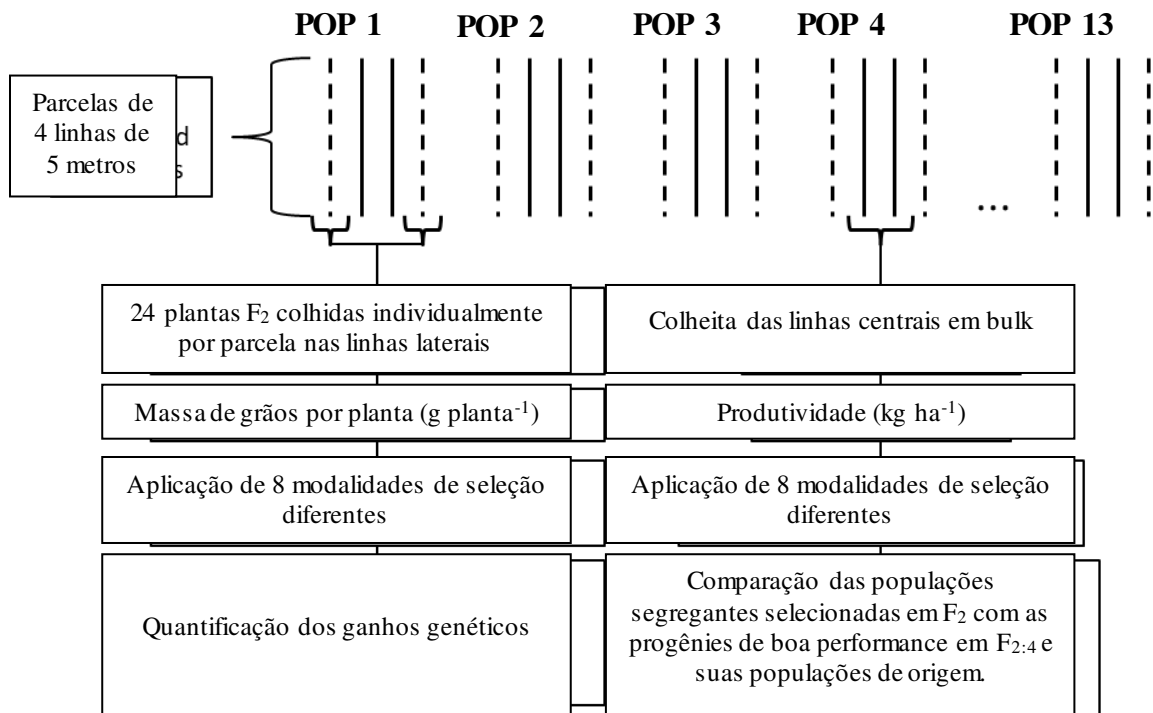


Figura 1 – Fluxograma de coleta e análise de dados do ensaio de campo das populações F₂.

O procedimento Restricted Maximum Likelihood/Best Linear Unbiased Prediction (REML/BLUP) foi adotado para as análises estatísticas, conforme proposto por Patterson e Thompson (1971) e Henderson (1975), respectivamente. O modelo estatístico associado à avaliação da massa de grãos por planta das populações segregantes F₂ foi o seguinte:

$$y_l = Xr + Zcgc_1 + Wcgc_2 + Tcec + Sp + e_l, \quad (\text{Eq. 1})$$

em que y_l é o vetor de dados fenotípicos individuais de massa de grãos por planta, r é o vetor dos efeitos fixos de repetição e testemunhas somados à média geral; $cgc1$ é o vetor dos efeitos de genitores pertencentes ao grupo 1 (aleatórios), em que $cgc1 \sim N(0, I\sigma_{cgc1}^2)$; $cgc2$ é o vetor dos efeitos de genitores pertencentes ao grupo 2 (aleatórios), em que $cgc2 \sim N(0, I\sigma_{cgc2}^2)$; cec é o vetor dos efeitos da capacidade específica de combinação dos genitores pertencentes ao grupo 1 com os pertencentes ao grupo 2 (aleatórios), em que $cec \sim N(0, I\sigma_{cec}^2)$; p é o vetor dos efeitos de parcela (aleatórios), em que $p \sim N(0, I\sigma_p^2)$; e_l é o vetor de erros (aleatórios), em que $e_l \sim N(0, I\sigma_{e_l}^2)$. As letras maiúsculas (**X**, **Z**, **W**, **T** e **S**) representam as matrizes de incidência para os referidos efeitos. Para os dados de produtividade de grãos em (kg ha^{-1}) foi ajustado um modelo semelhante ao apresentado (Eq. 1), porém o modelo não contempla o efeito de parcelas e y_l é o vetor de dados fenotípicos de produtividade de grãos em kg ha^{-1} .

Foram adotadas sete modalidades de seleção as quais variaram quanto aos valores dos efeitos genotípicos utilizados para a seleção e quanto ao número de indivíduos selecionados em cada população, conforme descrito na Tabela 2. Para tanto, foram consideradas três diferentes intensidades de seleção (IS), sendo 15%, 10% e 5%. Para o emprego das modalidades que incluem o BLUPISM foi adotado um número total de indivíduos ao invés do número de indivíduos selecionados na melhor população segregante, conforme proposto por Castro et al. (2016). Os autores sugerem o uso dos valores genotípicos (VG) para obtenção do número de indivíduos selecionados dentro de cada população, enquanto no presente trabalho, foram adotados diferentes valores dos efeitos genotípicos de genitores e de populações para obtenção do número de indivíduos a serem selecionados. O referido índice é apresentado abaixo, conforme proposto por Castro et al. (2016):

$$N_k = \left(\frac{vg_k}{vg_j} \right) n_{k-1} \quad (\text{Eq. 2})$$

$$N_k = u + g_k \quad (\text{Eq. 3})$$

No presente trabalho, foi proposta também uma modificação na fórmula original (Eq. 2) tal que o número de indivíduos a serem selecionados em uma população k é definido em função da razão entre o valor genotípico da população em questão (k) e a população imediatamente superior ($k-1$):

$$N_k = \left(\frac{vg_k}{vg_{k-1}} \right) n_{k-1} \quad (\text{Eq. 4})$$

$$N_k = u + g_k \quad (\text{Eq. 5})$$

em N_k é o número de indivíduos a serem selecionados na população k ; n_{k-1} se refere ao número de indivíduos selecionados na população imediatamente superiores, g_k é o valor do efeito genotípico da população k e Vg_k representa os valores genotípicos da população k , enquanto que g_j e Vg_j são o valor do efeito genotípico e o valor genotípico da melhor população, conforme descrito na Tabela 2 para cada modalidade.

Tal modificação foi realizada com a intenção preservar maior porção da variabilidade genética, por evitar a seleção de um número muito maior de plantas em uma população em detrimento daquelas com valor genotípico inferior.

Para definir qual a melhor modalidade entre as estudadas para a seleção em gerações precoces, foi calculado o ganho genético percentual dos indivíduos selecionados sobre a média geral das populações segregantes em três diferentes intensidades de seleção (15%, 10% e 5%). De acordo com Resende (2002) a metodologia ótima para seleção entre e dentro é o BLUPI (que consiste na modalidade E1) e, portanto, os ganhos obtidos pelas demais modalidades foram comparadas com esta.

Para as modalidades as quais há a possibilidade de combinações diferentes de indivíduos selecionados (exceto a E1), a seleção ou amostragem de indivíduos foi repetida cinco vezes para cada modalidade a fim de reproduzir a subjetividade da seleção de plantas a campo. Este processo é mencionado neste trabalho como “amostragem simulada”. A cada repetição os dados de massa de grãos por planta foram embaralhados, e os ganhos com a seleção foram computados sobre a média destas cinco repetições.

2.2.2 Progênes F_{2:4}

Para a avaliação da produtividade de grãos das progênes na geração F_{2:4} foram realizados dois experimentos nas localidades de Capinópolis – MG (18°40'48"S, 49°33'58"O e 530 m altitude) e Sorriso – MT (12°03'05"S, 56°03'14"W e 470 m de altitude). Os experimentos foram conduzidos em delineamento de blocos aumentados com oito blocos, sendo que parte dos genitores envolvidos nos cruzamentos realizados para originar as populações foi utilizada como testemunhas. Cada parcela foi constituída por duas linhas de 5 m lineares espaçadas por 0,5 m com população 260.000 planta ha⁻¹. Após a colheita, a massa de grãos da parcela foi corrigida para 13% de umidade.

Tabela 2 – Valores preditos utilizados e descrição das modalidades de seleção entre e dentro de populações segregantes F2 com base valores genotípicos (VG) ou amostragem simulada (AS) adotadas no trabalho.

Abreviação	Modalidades	Valor predito	Descrição	Seleção	
				Entre	Dentro
E1	BLUPI	$u + a$	Seleção dos indivíduos com melhor posição no ranking independente da população	VG	VG
E2	Amostragem	-	Amostragem de indivíduos em todas as populações em mesma proporção	-	AS
E3	Seleção por CEC	$u + g_{cec}$	Amostragem de indivíduos em igual proporção em populações com CEC acima da média	VG	AS
E4	Seleção por VG	$u + g_{cec} + g_{cgc1} + g_{cgc2}$	Amostragem de indivíduos em igual proporção em populações com VG acima da média	VG	AS
E5		$u + g_{cec}$		VG	AS
E6	BLUPISM	g_{cec}	Amostragem de indivíduos em número e populações indicado pelo índice	VG	AS
E7		$u + g_{cec} + g_{cgc1} + g_{cgc2}$		VG	AS
E8		$g_{cec} + g_{cgc1} + g_{cgc2}$		VG	AS

BLUPI = Blup individual; BLUPISM = Blup Individual Simulado Modificado; CEC = Capacidade específica de combinação; VG = Valor genótipo da população; u = média geral; a = valor aditivo do indivíduo; g_{cec} = valor do efeito genotípico da capacidade de combinação específica; g_i = valor do efeito genotípico da capacidade geral de combinação do genitor pertencente ao grupo i (i = 1 ou 2).

Com a finalidade de identificar as melhores populações F₂ que originaram as progênes F_{2:4} com mérito genético para ambos os ambientes, foi adotado o índice SIPP-BLUP (*Selection Index with Populations and Progenies – Best Linear Unbiased Predictor*) proposto por Resende et al. (2016) e adaptado por Volpato et al. (2018), o qual é equivalente ao seguinte modelo:

$$y_2 = \mathbf{H}b + \mathbf{C}g + \mathbf{K}i + \mathbf{J}f + e_2, \quad (\text{Eq 6})$$

em que y_2 é o vetor de dados, b é o vetor dos efeitos fixos de repetição e testemunhas somados à média geral; g é o vetor dos efeitos de progênes F_{2:4} (aleatórios), em que $g \sim N(0, I\sigma_g^2)$; i é o vetor dos efeitos da interação progênie x local (aleatórios) em que $i \sim N(0, I\sigma_i^2)$; f é o vetor dos efeitos de população (aleatórios), em que $f \sim N(0, I\sigma_f^2)$ e e_2 é o vetor de erros (aleatórios), em que $e_2 \sim N(0, I\sigma_{e_2}^2)$. As letras maiúsculas (**H**, **C**, **K** e **J**) representam as matrizes de incidência para os referidos efeitos.

Para verificar a relação entre populações promissoras entre as gerações F₂ e F_{2:4} foi computado o índice de coincidência das populações selecionados em cada uma das gerações, sendo que em F_{2:4}, considerou-se a predição dos valores dos efeitos genotípicos nas análises individuais e conjunta. Para as análises individuais o modelo utilizado foi similar ao apresentado na equação 4, não contemplando os efeitos de interação genótipo x ambiente. Todas as análises foram realizadas utilizando os Softwares SELEGEN REML/BLUP (RESENDE, 2016) e R (R CORE TEAM, 2020). O script em R para o BLUPISM baseado no número total de indivíduos que se deseja selecionar, conforme proposto neste trabalho é apresentado no Apêndice 1.

3. RESULTADOS E DISCUSSÃO

3.1 Modalidades de condução e seleção das populações segregantes

As estimativas dos componentes de variância da capacidade geral de combinação dos genitores dos grupos G₁ e G₂ (CGC₁ e CGC₂, respectivamente) se mostraram significativas (Tabela 3), o que indica que os genitores contribuíram com a transferência de diferentes alelos que controlam o caractere (DOVALE & FRITSCHÉ-NETO, 2013).

Tabela 3 – Estimativas de parâmetros genéticos e fenotípicos para o caractere massa de grãos por planta em plantas individuais de populações F₂ de soja.

Parâmetros	Estimativas
σ^2_{cgc1}	6,78 *
σ^2_{cgc2}	4,78 *
σ^2_a	23,13
σ^2_{cec}	4,94 *
σ^2_p	1,47
σ^2_e	85,37
σ^2_f	103,34
h^2_{g1}	0,26
h^2_{g2}	0,19
h^2_a	0,22
h^2_{dom}	0,19
c^2_{cec}	0,05
h^2_g	0,41
r^2_{ggpop}	0,62
r^2_{gggen}	0,75
CV_e	0,22
Média	41,48

* significativo a 5% de probabilidade pelo teste da razão de verossimilhança; σ^2_{cgc1} = variância entre genitores pertencentes ao grupo 1; σ^2_{cgc2} = variância entre genitores pertencentes ao grupo 2; σ^2_a = variância genética aditiva média; σ^2_{cec} = variância da capacidade específica de combinação entre genitores dos diferentes grupos; σ^2_p = variância entre parcelas; σ^2_e = variância residual; σ^2_f = variância fenotípica; h^2_{g1} = herdabilidade no sentido restrito no grupo 1 de genitores; h^2_{g2} = herdabilidade no sentido restrito no grupo 2 de genitores; h^2_a = herdabilidade interpopulacional individual média no sentido restrito; h^2_{dom} = herdabilidade individual dos efeitos interpopulacionais de dominância; c^2_{cec} = coeficiente de determinação dos efeitos da capacidade específica de combinação; h^2_g = herdabilidade interpopulacional individual no sentido amplo; r^2_{ggpop} = acurácia na seleção das populações; r^2_{gggen} = acurácia na seleção dos genitores; CV_e = coeficiente de variação residual.

A variância da capacidade específica de combinação (CEC) entre os genitores dos grupos G₁ e G₂ apresentaram significância, o que sugere a presença de efeitos não aditivos no controle do caractere e indica a possibilidade de seleção entre as populações segregantes, dado a presença de desvios devido à dominância (Pandini et al., 2001).

As estimativas de herdabilidade encontradas foram similares as encontradas por Rocha et al. (2018) na avaliação de populações F₂ de soja. De acordo com Pereira et al. (2017), em fases iniciais do programa de melhoramento é comum encontrar herdabilidade de magnitude reduzida, uma vez que as populações ou progênes são testadas em poucas ou sem repetições e em um ou poucos locais. O coeficiente de variação residual (CV_e) encontrado neste

experimento foi de 0,22 e está de acordo com a literatura (VOLPATO et al., 2018; RAMALHO et al., 2012) e é satisfatório para o caractere em questão. O valor de acurácia encontrado para a seleção de populações foi de 0,62, enquanto para a seleção entre genitores foi de 0,75. Diante do alto valor de acurácia encontrado para a seleção de genitores, e a importância dos efeitos aditivos para a seleção em gerações precoces é aconselhável incluir as informações dos genitores nas análises estatísticas para aumentar as chances de sucesso com a seleção.

A seleção entre e dentro de populações segregantes é uma das estratégias adotadas na condução de populações segregantes no melhoramento de autógamias (BERNARDO, 2010) e de acordo com Witcombe e Virk (2001), programas de melhoramento genético, por muitas vezes, são conduzidos com a participação de agricultores, os quais cedem área e recursos limitados para a condução do programa. Nestes casos os autores citam que nem sempre é possível implantar experimentos com grande número de tratamentos, sendo viável o uso da seleção para reduzir o número de tratamentos e utilizar os recursos disponíveis para condução das progênes mais promissoras.

As maiores estimativas de ganho com a seleção para todas as intensidades de seleção estudadas foram obtidas pela modalidade de seleção BLUPI (E1) (Tabela 4), a qual se baseia no ranking dos valores dos efeitos genotípicos dos indivíduos, conforme relatado por Resende (2002) na seleção de famílias de cana-de-açúcar. Ganhos genéticos de magnitudes altas, não superando apenas a adição da modalidade E1, foram obtidos pelo emprego do BLUPISM com a adoção de valores dos efeitos genotípicos $g_{cec} + g_{cgc1} + g_{cgc2}$ (E8), com estimativas percentuais variando entre 6,24% e 6,43%. Com exceção da modalidade E3 e E6, as quais consideram apenas os efeitos da CEC como valores genotípicos, foi observada a tendência de aumento dos ganhos com a seleção à medida que houve o aumento da intensidade de seleção.

A amostragem de indivíduos em populações em igual proporção (E2), apresentou ganhos negativos próximos a 1% (Tabela 4). Tal modalidade é comumente empregada em programas de melhoramento de autógamias, visto que não é corriqueira a avaliação de populações em geração F₂ e, portanto, o mérito genético de cada população não é levado em consideração. A seleção entre ou dentro de populações com baixo mérito genotípico, conforme ocorre na adoção da modalidade E2, faz com que o melhorista ocupe parte de sua capacidade de teste e direcione recursos à condução de populações com baixa probabilidade de derivar linhagens promissoras (HUEHN, 2006).

Tabela 4 – Ganhos genéticos percentuais obtidos para massa de grãos por planta em todas as modalidades adotadas para a seleção em populações segregantes F₂.

Modalidade de seleção	Intensidade de seleção		
	15%	10%	5%
E1	9,99	11,10	13,01
E2	-1,21	-1,07	-0,77
E3	0,46	0,56	0,44
E4	4,92	5,01	5,04
E5	-1,28	-1,17	-1,06
E6	3,93	3,03	2,79
E7	-1,31	-1,17	-1,06
E8	6,24	6,41	6,43

Em todos os três diferentes cenários de intensidade de seleção, a modalidade E8 ($BLUPISM - g_{cec} + g_{cgc1} + g_{cgc2}$) apresentou ganhos genéticos superiores aos da modalidade E6 ($BLUPISM - g_{cec}$) com diferenças entre 2,31% e 3,37% para 15% e 10% de intensidade, respectivamente. Para ambas as modalidades (E6 e E8) foi adotado o BLUPISM, porém, sobre valores diferentes, de modo que a modalidade E8 contemplou os valores dos efeitos das capacidades gerais de combinação. O mesmo é notado entre as modalidades E3 e E4, onde foram selecionados indivíduos em igual proporção dentro das populações com valores genotípicos positivos. Neste caso, a modalidade que contempla as informações de CGC e CEC (E4) superou os ganhos obtidos por aquela que contemplou apenas a CEC (E3). Estes resultados reforçam a vantagem de considerar as informações de genitores para a seleção em populações segregantes.

O BLUPISM foi empregado em quatro modalidades diferentes (E5, E6, E7 e E8). Em cada uma destas modalidades foram utilizados valores diferentes para representar o mérito genético das populações. Foi observado que nos casos em que os valores dos efeitos genotípicos referentes às populações foram somados a média geral ($u + g_{cec}$ ou $u + g_{cec} + g_{cgc1} + g_{cgc2}$), ou E5 e E7, os ganhos genéticos esperados com a seleção entre e dentro de populações foram próximos à zero, e neste caso, negativos. Por outro lado, quando valores dos efeitos genotípicos não foram somados a média (E6 e E8), os ganhos encontrados foram positivos 2,79% e 6,43%. Tal fato ocorreu devido à pequena magnitude dos valores de g_{cec} , g_{cgc1} e g_{cgc2} (variaram entre -2,71 e 2,96 - dados não apresentados) quando comparados à média geral (41,48). Sendo assim, quando os valores genotípicos foram somados a média geral, a razão entre os valores genotípicos das populações k e j ($u + g_{cec}$ ou $u + g_{cec} + g_{cgc1} + g_{cgc2}$) se tornou muito pequena em comparação com quando a média geral não foi somada a estes valores. Portanto, adotar as

modalidades E5 e E7 apresentou resultados muito similares aos encontrados para a modalidade E2 em termos de número de indivíduos selecionados, conforme apresentado na Tabela 5.

Tabela 5 – Número de indivíduos selecionados (IS = 15%) dentro de cada população F₂ sob a adoção das modalidades E2, E5, E6, E7 e E8 e respectivos valores genotípicos das populações em modalidades que contemplaram a seleção entre.

População	Estratégia de seleção									
	E2		E5		E6		E7		E8	
	n	VG (u + g _{cec})	n	VG (g _{cec})	n	VG (u + g _{cec} + g _{cgcl} + g _{cgcl2})	n	VG (g _{cec} + g _{cgcl} + g _{cgcl2})	n	
POP33	9	43,64	9	2,10	40	46,89	9	5,35	46	
POP12	9	42,16	9	0,62	12	45,32	9	3,78	32	
POP30	9	43,46	9	1,92	37	43,82	9	2,28	19	
POP6	9	42,36	9	0,82	16	43,71	9	2,17	18	
POP22	9	40,45	9	-1,09	0	42,74	9	1,20	10	
POP34	9	41,37	9	-0,17	0	42,71	9	1,17	10	
POP11	9	42,82	9	1,28	25	42,55	9	1,01	9	
POP31	9	40,07	9	-1,47	0	41,38	9	-0,16	0	
POP23	9	40,72	9	-0,82	0	41,10	9	-0,44	0	
POP32	9	41,36	9	-0,18	0	39,27	9	-2,27	0	
POP10	9	42,12	9	0,58	11	39,04	9	-2,50	0	
POP21	9	41,46	9	-0,08	0	38,41	9	-3,13	0	
POP5	9	39,97	9	-1,57	0	37,89	9	-3,65	0	
POP4	9	41,96	9	0,42	8	37,07	9	-4,47	0	
POP20	9	40,93	9	-0,61	0	36,02	9	-5,52	0	
POP19	9	39,79	9	-1,75	0	33,93	9	-7,61	0	

n = número de indivíduos selecionados dentro de cada população; VG = valor genotípicos utilizados para a estratégia em questão.

Castro et al. (2016) realizaram a seleção de família de cana-de-açúcar com o uso do BLUPISM conforme proposto e descrito na equação 2. Mesmo utilizando o VG em detrimento aos valores dos efeitos genotípicos (g) conforme usado no presente trabalho, os autores encontraram número diferente de indivíduos selecionados ao longo das famílias, uma vez que os valores genotípicos das famílias representavam maior proporção do VG quando os efeitos dos valores genotípicos foram somados a média.

Conforme relatado por Resende e Barbosa (2006) e Castro et al., (2016), as estratégias que envolveram o uso do BLUPISM não selecionaram indivíduos dentro de populações com VG negativo, independente dos valores genéticos utilizados. A exclusão de populações pouco promissoras é viável, pois evita o desperdício de recursos e eficiente quando realizada em fases

iniciais do programa (ROCHA et al., 2018; BOREM et al., 2017; CARVALHO et al., 2009; ROSAL et al., 2000; FEHR, 1987). Outra vantagem do uso do BLUPISM apontada por Castro et al. (2016), é a de que a metodologia reproduziu quase exatamente o BLUPI em termos de número de indivíduos selecionados por família. (0,98 de correlação), bem como os resultados encontrados no presente trabalho, já que foram encontradas correlações entre 0,90 e 0,92 entre as metodologias (Tabela 6).

Tabela 6 – Número de indivíduos selecionados em cada população de acordo com as estratégias E1 (BLUPI) e E8 (BLUPISM com o emprego de $g_{cec} + g_{cgc1} + g_{cgc2}$) e correlação para as intensidades de seleção de 15%, 10% e 5%.

População	15%		10%		5%	
	BLUPI	BLUPISM	BLUPI	BLUPISM	BLUPI	BLUPISM
POP33	49	46	39	30	17	15
POP12	39	32	25	21	15	11
POP30	5	19	2	13	1	7
POP6	10	18	6	12	3	7
POP22	20	10	9	7	6	4
POP34	7	10	4	7	2	4
POP11	1	9	1	6	0	3
POP4	0	0	0	0	0	0
POP5	1	0	1	0	0	0
POP10	0	0	0	0	0	0
POP19	0	0	0	0	0	0
POP21	0	0	0	0	0	0
POP23	3	0	3	0	2	0
POP20	0	0	0	0	0	0
POP31	2	0	2	0	0	0
POP32	0	0	0	0	0	0
Correlação	0,92		0,91		0,89	

Foi observada a tendência do número de indivíduos selecionados por população pela estratégia E8 ser mais brando, ou seja, mais próximo da média quando comparado ao BLUPI. Por exemplo, a população POP12 teve menor número de indivíduos selecionados pelo BLUPISM, enquanto a POP11 apresentou a tendência inversa. Discussão similar foi apresentada por Castro et al. (2016). Isto sugere que a adoção do BLUPISM promove a melhor amostragem da variabilidade genética que por sua vez é fundamental para projetar ganhos ao longo do tempo em um programa de seleção recorrente (SILVA et al., 2017 e RAMALHO et al., 2012).

3.2 Comparação entre as populações indicadas em F2 e F2:4

De acordo com Lima et al. (2015) existem várias maneiras possíveis para se conduzir populações segregantes e todas estas seriam eficientes caso não houvesse interação genótipo x ambiente (GxA). Este componente é responsável pelo comportamento diferencial dos genótipos ao longo dos ambientes (PANTHEE et al., 2007) e perturba a seleção (FALCONER e MACKAY, 1996). A interação GxA pode reduzir a eficiência da seleção ao longo das gerações (MENDES et al., 2012). Além disto, o número de sementes nas primeiras gerações filiais é reduzido, culminando em um ou poucos locais de experimentação (PEREIRA et al., 2017). Portanto, se faz necessário encontrar metodologias que permitem identificar as populações com capacidade de derivar progênes com bom desempenho produtivo em gerações avançadas. Ademais, as primeiras gerações devem ser conduzidas em locais onde as populações indicadas como mais promissoras derivem progênes com bom desempenho para o caractere de interesse nos locais de teste nas gerações futuras.

Para as análises da geração F_{2:4} foi verificado o efeito significativo para a interação progênie x local, o que indica o comportamento diferencial dos genótipos ao longo dos ambientes (Tabela 7). Apesar disto, os valores estimados para o coeficiente de determinação da interação ($c^2_{int} = 0,10$) e a correlação entre locais ($r_{loc} = 0,89$) demonstram que a variação devido aos efeitos da interação não ocupa grande proporção da variação fenotípica e que o ranqueamento das populações não deve sofrer grandes variações entre os ambientes, não oferecendo grandes riscos a seleção indireta entre locais (Cruz et al., 2014).

O efeito de populações contemplado no modelo estatístico (Eq. 4) para as análises das progênes em F_{2:4} permitem obter a predição dos valores genotípicos das populações e, portanto, inferir sobre as populações com maior potencial em derivar progênes com bom desempenho para a produtividade grãos. Além disto, a literatura relata que a inclusão dos efeitos de populações conduz a melhores estimativas dos componentes de variância e predição dos valores genotípicos (VOLPATO et al., 2018; PEREIRA et al., 2017; DUARTE e VENCOVSKY, 2001).

As estimativas dos componentes de variância para as análises individuais e conjunta das progênes na geração F_{2:4} são apresentadas na tabela 7. Para a análise conjunta, as estimativas de variância associadas à população, progênes e interação progênie x local foram significativos, o que significa que as progênes responderam de modo diferente aos diferentes ambientes onde foram testadas (TENG et al., 2017; PANTHEE et al., 2007). Para análise

individual de Sorriso, a estimativa de variância de progênies foi significativa, a de Capinópolis, apenas as estimativas de variância associadas às populações foram significativas.

Tabela 7 – Estimativas de parâmetros genéticos e fenotípicos para o caractere produtividade de grãos (kg ha^{-1}) em progênies $F_{2:4}$ de soja.

Parâmetros	Análise conjunta	Análises individuais	
		Sorriso - MT	Capinópolis - MG
$\sigma^2_{\text{pop}}^1$	50332,71 *	5902,15	194983,45 **
σ^2_{prog}	305371,91 **	382727,63 **	304220,90
σ^2_{int}	90127,77 **	-	-
σ^2_e	422649,65	256961,15	639947,98
σ^2_f	868482,03	645590,94	1139152,32
h^2_{pop}	0,06	0,01	0,17
h^2_a	0,35	0,59	0,27
h^2_{aT}	0,41	0,60	0,44
c^2_{int}	0,10	-	-
r_{loc}	0,89	-	-
CV_e	0,20	0,17	0,24
Média	3219,02	3028,74	3400,23

* e ** significativo a 5% e 1% de probabilidade pelo teste da razão de verossimilhança; $^1\sigma^2_{\text{pop}}$ = variância entre populações; σ^2_{prog} = variância entre progênies ou variância entre progênies dentro de populações; σ^2_{int} = variância da interação progênie x local ou variância da interações entre progênies dentro de populações x local; σ_e = variância residual; σ_f = variância fenotípica; h^2_{pop} = herdabilidade entre populações; h^2_a = herdabilidade dentro de populações; h^2_{aT} = herdabilidade total = herdabilidade entre populações + herdabilidade dentro de populações; c^2_{int} = coeficiente de determinação dos efeitos da interação progênies x ambientes; r_{loc} = correlação entre ambientes; CV_e = coeficiente de variação residual.

Os valores de CV_e encontrados para as análises individuais e conjunta variaram entre 0,17 e 0,24 e estão de acordo com os encontrados na literatura para produtividade de grãos (VOLPATO et al., 2018; RAMALHO et al., 2012), especialmente para experimentação em programas de melhoramento genético de soja (OLIVEIRA et al., 2014). As estimativas de herdabilidade total, dos efeitos aditivos e de populações encontrados neste trabalho são compatíveis com aqueles encontrados por Volpato et al. (2018).

Os valores dos efeitos genotípicos para as avaliações nas gerações F_2 e $F_{2:4}$, tanto para as análises individuais quanto para a conjunta são apresentados na tabela 8. No experimento da geração $F_{2:4}$ conduzido em Capinópolis – MG não foi avaliada nenhuma progênie derivada da população POP32 devido ao número de sementes reduzido obtido em $F_{2:3}$.

Tabela 8 – Valores dos efeitos genotípicos das populações nas gerações F₂ e F_{2:4} para as análises individuais e conjunta.

Geração	F ₂		F _{2:4}	
	Local	(Cajuri - MG)	(Sorriso - MT)	(Capinópolis - MG)
População	$g_{cec} + g_{cgc1} + g_{cgc2}$	g	g	g
POP12	3,78	9,95	320,31	129,76
POP30	2,28	20,74	241,25	157,21
POP6	2,17	-5,13	-421,02	-123,76
POP34	1,17	-15,7	392,67	53,00
POP11	1,01	16,67	452,31	209,80
POP31	-0,16	-48,62	-409,65	-248,5
POP32	-2,27	-9,35	-	-51,67
POP10	-2,5	-22,86	-78,14	-21,5
POP4	-3,13	29,46	-139,57	-9,69
POP21	-3,65	-2,82	-167,52	-56,81
POP5	-4,47	0,48	255,69	62,81
POP20	-5,52	-14,36	-514,82	-234,62
POP19	-7,61	41,53	68,49	133,97

*Populações F₂ com valor do efeito genotípico positivo em Cajuri – MG.

Na geração F₂, conduzida em Cajuri, as populações POP12, POP30, POP6, POP34 e POP11 mostraram ser promissoras para o caractere massa de grãos por planta, pois apresentaram valores dos efeitos genotípicos positivos. Mesmo conduzidas em ambientes diferentes, as populações promissoras na geração F₂, também se mostraram promissoras em F_{2:4} ao derivar progênies com bom desempenho produtivo para pelo menos um dos locais, com exceção da POP6.

Por outro lado, foi notada a vantagem de descartar populações pouco promissoras na geração F₂ com o uso do BLUPISM ($g_{cec} + g_{cgc1} + g_{cgc2}$). Entre as oito populações com valores dos efeitos genotípicos negativos em F₂, apenas duas se mostraram promissoras em Capinópolis e na análise conjunta, enquanto três se mostraram promissoras em Sorriso. Outros autores que realizaram a seleção precoce em populações de autógamias também apontam como vantajoso o descarte de populações pouco promissoras (ROCHA et al., 2018; CARVALHO et al., 2009; ROSAL et al., 2000; FEHR, 1987).

Se for considerada a possibilidade da condução do programa de melhoramento para ambos os locais (Sorriso - MT e Capinópolis – MG) de teste das progênies na geração F_{2:4}, deveriam ser selecionadas as populações com alto valor dos efeitos genotípicos preditos na análise conjunta. Neste caso, apenas a população POP6, não promissora na análise conjunta, seria indicada para a seleção na geração F₂. Mesmo que a coincidência não seja total, os

resultados mostram que é vantajoso realizar a seleção entre populações F₂ em Cajuri - MG nas gerações iniciais, visando derivar linhagens para as localidades de Sorriso e Capinópolis, uma vez que, as populações selecionadas em F₂, originam progênes promissoras para estes locais em gerações futuras.

4. CONCLUSÕES

O BLUPISM com o uso dos valores $g_{cec} + g_{cgc1} + g_{cgc2}$ (E8) apresentou ganhos mais altos entre as estratégias propostas e foi a que proporcionou ganho com a seleção mais próximos daqueles obtidos pela estratégia de referência BLUPI (E1), cuja prática é inviável.

As estratégias BLUPISM e BLUPI apresentaram alta concordância em números de indivíduos selecionados dentro de todas as populações.

Houve alta coincidência entre as populações selecionadas na geração F₂ pelo BLUPISM com uso de $(g_{cec} + g_{cgc1} + g_{cgc2})$ e aquelas indicadas como promissoras na geração F_{2:4}.

O BLUPISM é uma ferramenta útil e com potencial para ser aplicado de forma prática em programa de melhoramento de autógamias.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ALLIPRANDINI, L. F. et al. Understanding Soybean Maturity Groups in Brazil: Environment, Cultivar Classification, and Stability. **Crop Science**, v. 49, n. 3, p. 801–808, maio 2009.
- ROCHA, G. A. DE F.; PEREIRA, F. A. DE C.; VELLO, N. A. Potential of soybean crosses in early inbreeding generations for grain yield. **Crop Breeding and Applied Biotechnology**, v. 18, n. 3, p. 267–275, set. 2018.
- BERNARDO, R. N. **Breeding for quantitative traits in plants**. Woodbury, Minn.: Stemma Press, 2010.
- BORÉM, A.; MIRANDA, G. V.; FRITSCHÉ-NETO, R. **Melhoramento de plantas**. 7. ed. Viçosa, MG: Editora UFV, 2017.
- CÂMARA, G. M. DE S. **Soja: Tecnologia de produção**. Soja: Tecnologia de produção. **Anais...** In: SOJA: TECNOLOGIA DE PRODUÇÃO. Piracicaba: ESALQ, 1998.
- CARVALHO, A. D. F.; GERALDI, I. O.; SANTOS, V. DA S. Avaliação de progênies f2:4 e f4:6 de soja e perspectivas do uso de teste precoce para a produção de grãos. **Bragantia**, v. 68, n. 4, p. 857–861, 2009.
- CASTRO, R. D. et al. Selection between and within full-sib sugarcane families using the modified BLUPIS method (BLUPISM). **Genetics and Molecular Research**, v. 15, n. 1, 2016.
- CONAB. **Série Histórica**. Disponível em: <<http://www.conab.gov.br/info-agro/safras/serie-historica-das-safras/item/7666-soja/>>. Acesso em: 4 jul. 2021.
- COOPER, R. L. Modified Early Generation Testing Procedure for Yield Selection in Soybean. **Crop Science**, v. 30, n. 2, p. 417, 1990.
- CRUZ, C. D.; CARNEIRO, P. C. S.; REGAZZI, A. J. **Modelos Biométricos Aplicados ao Melhoramento Genético** - Vol 2 - 3a Edição. 3. ed. Viçosa, MG: Editora UFV, 2014. v. 2
- DOVALE, J. C.; FRITSCHÉ-NETO, R. Genetic control of traits associated with phosphorus use efficiency in maize by REML/BLUP. **Revista Ciência Agronômica**, v. 44, n. 3, p. 554–563, set. 2013.
- DUARTE, J. B.; VENCOSKY, R. Estimação e predição por modelo linear misto com ênfase na ordenação de médias de tratamentos genéticos. **Scientia Agricola**, v. 58, n. 1, p. 109–117, mar. 2001.

- FALCONER, D. S.; MACKAY, T. **Introduction to quantitative genetics**. 4. ed., [16. print.] ed. Harlow: Pearson, Prentice Hall, 1996.
- FEHR, W. R.; FEHR, E. L.; JESSEN, H. J. (EDS.). **Principles of cultivar development**. New York : London: Macmillan ; Collier Macmillan, 1987.
- GATES, C. E.; WEBER, C. R.; HORNER, T. W. A Linkage Study of Quantitative Characters in a Soybean Cross ¹. **Agronomy Journal**, v. 52, n. 1, p. 45–49, jan. 1960.
- HENDERSON, C. R. Best Linear Unbiased Estimation and Prediction under a Selection Model. **Biometrics**, v. 31, n. 2, p. 423–447, 1975.
- HUEHN, M. Optimum number of crosses and progeny per cross in breeding self-fertilizing crops. III. Economic constraints. **Cereal Research Communications**, v. 34, n. 2–3, p. 903–910, jun. 2006.
- LIMA, D. C. et al. Breeding common bean populations for traits using selection index. **Scientia Agricola**, v. 72, n. 2, p. 132–137, fev. 2015.
- MAUAD, M. et al. Influência da densidade de semeadura sobre características agronômicas na cultura da soja. **Agrarian**, v. 3, n. 9, p. 175–181, 2010.
- MENDES, M. P.; RAMALHO, M. A. P.; ABREU, A. F. B. Strategies in identifying individuals in a segregant population of common bean and implications of genotype x environment interaction in the success of selection. **Genetics and Molecular Research**, v. 11, n. 2, p. 872–880, 2012.
- OLIVEIRA, I. J. DE et al. Diallel among soybean genotypes with high oil content and resistance to sudden death syndrome. **Brazilian Archives of Biology and Technology**, v. 57, n. 2, p. 178–186, abr. 2014.
- PANDINI, F.; VELLO, N. A.; LOPES, Â. C. A. Performance of Agronomic Traits in a Soybean F1 Diallel System. **Cropp Breeding and Applied Biotechnology**, v. 1, n. 3, p. 229–244, 30 set. 2001.
- PANTHEE, D. R. et al. Quantitative trait loci for agronomic traits in soybean. **Plant Breeding**, v. 126, n. 1, p. 51–57, 2007.
- PATTERSON, H. D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v. 58, n. 3, p. 545–554, 1971.
- PEREIRA, F. DE C. et al. Implications of the population effect in the selection of soybean progeny. **Plant Breeding**, v. 136, n. 5, p. 679–687, 2017.
- R CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, 2020.

- RAMALHO, M. A. P. et al. **Aplicações da Genética Quantitativa no Melhoramento de Plantas Autógamas**. Lavras, MG: UFLA, 2012.
- RESENDE, M. D. V. DE; BARBOSA, M. H. P. Selection via simulated individual BLUP based on family genotypic effects in sugarcane. **Pesquisa Agropecuária Brasileira**, v. 41, n. 3, p. 421–429, mar. 2006.
- DE RESENDE, M. D. V. et al. Selection Index with Parents, Populations, Progenies, and Generations Effects in Autogamous Plant Breeding. **Crop Science**, v. 56, n. 2, p. 530–546, mar. 2016.
- RESENDE, M. D. V. DE. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília, DF: Embrapa Informação Tecnológica, Colombo: Embrapa Florestas, 2002.
- RESENDE, M. D. V. DE. Software Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology**, v. 16, n. 4, p. 330–339, dez. 2016.
- ROSAL, C. J. D. S. et al. Seleção precoce para a produtividade de grãos no feijoeiro. **Bragantia**, v. 59, n. 2, p. 189–195, 2000.
- SEDIYAMA, T.; SILVA, F. L.; BORÉM, A. **Livro Soja do Plantio à Colheita - UFV**. Viçosa, MG: Editora UFV, 2015.
- SILVA, F. L. DA et al. **Soybean breeding**. New York, NY: Springer Science+Business Media, 2017.
- TENG, W. et al. Dissection of the genetic architecture for soybean seed weight across multiple environments. **Crop and Pasture Science**, v. 68, n. 4, p. 358, 2017.
- VOLPATO, L. et al. Selection of inbred soybean progeny (*Glycine max*): an approach with population effect. **Plant Breeding**, v. 137, n. 6, p. 865–872, dez. 2018.
- WITCOMBE, J. R.; VIRK, D. S. Number of crosses and population size for participatory and classical plant breeding. **Euphytica**, v. 122, n. 3, p. 451–462, 1 dez. 2001.
- WYSMIERSKI, P. T.; VELLO, N. A. The genetic base of Brazilian soybean cultivars: evolution over time and breeding implications. **Genetics and Molecular Biology**, v. 36, n. 4, p. 547–555, 25 out. 2013.

**EMPREGO DO ÍNDICE DE SELEÇÃO COM GENITORES E POPULAÇÕES PARA
A SELEÇÃO DE PROGÊNIES DE SOJA**

CAPÍTULO 2

RESUMO

O ciclo de melhoramento pode ser compreendido como o tempo entre os primeiros cruzamentos e a adoção de um novo cultivar ou o tempo de “cruzamento a cruzamento” e a redução destes tempos é uma das maneiras de aumentar a eficiência dos programas de melhoramento. Reduzir o ciclo de melhoramento é uma das maneiras de elevar a eficiência dos programas. Para isto, se faz necessário realizar a seleção em estádios precoces do programa, o que requer a adoção de metodologias que permitam a seleção com alta acurácia devido aos fatores perturbadores à seleção neste cenário. Neste contexto, os objetivos deste trabalho foram: a) avaliar os efeitos do emprego da informação de genitores na avaliação de progênies $F_{2:4}$ de soja sobre a qualidade experimental; b) avaliar os efeitos do emprego do índice de seleção genitores + progênies sobre a acurácia de seleção de progênies frente à seleção baseada nos melhores preditores lineares não viesados (BLUP – *Best Linear Unbiased Prediction*). Para isso, foram desenvolvidas 16 populações de soja em esquema de dialelo parcial com dois grupos de genitores. Na geração F_2 houve a seleção de plantas individuais para a derivação de progênies as quais foram conduzidas em autofecundação até a obtenção das sementes $F_{2:4}$. Na geração $F_{2:4}$ foram conduzidos dois experimentos a campo nas localidades de Sorriso – MT e Capinópolis – MG em delineamento de blocos aumentados com oito repetições. As parcelas foram constituídas de duas linhas de cinco metros, sobre as quais foi avaliada a produtividade de grãos. Os resultados mostraram que a inclusão das informações de genitores proporciona uma seleção mais acurada das progênies, além de ser de fácil inclusão no modelo. A inclusão destas informações é mais efetiva quando há efeito significativo entre os genitores. O índice de seleção genitores + progênies apresentou melhores estimativas de acurácia em relação ao modelo completo (inclusos dados de genitores) e deve conduzir o melhorista a uma seleção mais confiável e maiores ganhos genéticos.

1 INTRODUÇÃO

A soja [*Glycine Max* (L.) Merr.] é uma das culturas mais cultivadas no mundo e uma das principais fontes de óleo e proteína vegetal. Nos últimos 45 anos a produtividade média da cultura no Brasil evoluiu de 1748 kg ha⁻¹ para 3517 kg ha⁻¹. (CONAB, 2021). Dentro dessa evolução, o melhoramento genético que visa a busca por maiores produtividades, tem sido um dos grandes responsáveis (TODESCHINI et al., 2019).

Programas de melhoramento de soja clássicos podem levar de 8 a 12 anos entre o cruzamento e o lançamento de uma nova variedade (BORÉM e MIRANDA, 2017). Entretanto, diversas técnicas modernas possibilitam reduzir esse tempo para 4 a 7 anos. Uma das maneiras conhecidas para melhorar a eficiências de programas de melhoramento é reduzir o ciclo de melhoramento, que é o tempo compreendido entre os primeiros cruzamentos e o emprego dos cultivares ou ainda o tempo de “cruzamento a cruzamento” (CECCARELLI, 2015).

A redução do ciclo de melhoramento para a cultura da soja pode ser alcançada pela doação de berçários em regiões de baixas latitudes ou no hemisfério oposto, conduzindo dois ciclos da cultura em situações de chuvas em distribuição bimodal; uso de *single-seeds descent* (SSD) em autógamas (GOULDEN, 1939); colheita antecipada de sementes (HICKEY, 2019); uso de duplo haploides ou emprego de ferramentas biotecnológicas tais como a introgressão de QTLs (RIBAUT e HOISINGTON, 1998); e condução de gerações de cruzamento e avanços de geração em ambientes com umidade, temperatura, fotoperíodo e suplementação de CO₂ controlados e o emprego da edição gênica (NAGATOSHI e FUJITA, 2019; HICKEY, 2019). De modo geral, estas estratégias para acelerar os ciclos de melhoramento são conhecidas como “*speed breeding*” e tem sido assuntos de muitos estudos na atualidade. O uso das técnicas citadas para diminuir o tempo de obtenção de uma cultivar, muitas vezes são de alta complexidade ou custo e têm sido comumente aplicadas com o intuito de reduzir o tempo para a geração de novas variedades e não para reduzir o tempo de “cruzamento a cruzamento” (WANGA et al., 2021).

Os programas de melhoramento de espécies cultivadas, sobretudo de soja, em sua maioria são desenhados sobre o esquema de seleção recorrente (SILVA et al., 2017). Nesta estratégia de melhoramento, as melhores progênies ou variedades obtidas ao fim de um ciclo de seleção, são utilizadas como genitores na formação de uma nova população (HALLAUER e DARRAH, 1985). Sendo assim, é válida a busca por alternativas que visam reduzir o ciclo de melhoramento (tempo de cruzamento a cruzamento) a fim de gerar novas variedades em menos tempo (HICKEY, 2019).

É importante salientar que a cada ciclo de melhoramento em um programa de seleção recorrente existem dois propósitos: a derivação de linhagens para testes de produtividade e seleção de progênies para recombinação. Neste contexto, quanto menor for o ciclo de melhoramento, maior será o ganho com a seleção por reduzir o valor de “ t ” na equação do ganho genético expressa por $\Delta G = \frac{i r \sigma_g}{t}$, onde i representa a intensidade de seleção; r representa a acurácia de seleção; σ_g é a raiz quadrada da variância genética; e t é o tempo do ciclo referente ao ciclo de melhoramento para o qual se estima o ganho.

Para alcançar a redução do tempo em que se conclui cada ciclo de melhoramento, o melhorista pode usar a seleção precoce. O emprego da seleção precoce pode ser uma ferramenta utilizada para aumentar a eficiência do programa, além de apresentar melhor custo-benefício quando comparado a situações de programas que executam a seleção em estágios mais tardios (CECCARELLI, 2015), pois permite além de reduzir o ciclo de melhoramento, avançar o programa com um menor número de progênies e concentrar recursos e esforços na avaliação de progênies promissoras (SILVA, et al., 2017, BERNARDO, 2010).

Para contornar os possíveis efeitos perturbadores à seleção que podem ser ocasionados pelos loci em heterozigose em fases precoces do programa de melhoramento, se faz necessário lançar mão de metodologias que permitam alcançar boa acurácia na seleção da avaliação das progênies, além de conduzir os ensaios com boa qualidade experimental. De acordo com Bauer e Léon (2008) a seleção efetuada contemplando as informações de genitores ou de pedigree tende a ser mais eficiente do que aquela efetuada considerando apenas a informação das progênies. Para Duarte e Vencovsky (2001) e Volpato et al. (2020) desconsiderar informações de parentais ou parentesco pode comprometer as estimativas dos componentes de variância, as predições dos valores genéticos e por consequência, o ranqueamento das progênies.

Sendo assim, Resende (2015) propôs o uso de um índice de seleção que considera as informações de genitores e progênies. O objetivo deste índice é proporcionar o ranqueamento das progênies de acordo com o seu mérito genético ponderado pelas suas acurácias e pelos méritos genéticos e acurácias de seus genitores. Ainda não foram encontrados relatos da aplicação deste índice em programas de melhoramento genético de soja.

Neste contexto, os objetivos deste trabalho foram: a) avaliar os efeitos do emprego da informação de genitores na avaliação de progênies $F_{2:4}$ de soja sobre a qualidade experimental; b) avaliar os efeitos do emprego do índice de seleção genitores + progênies proposto por Resende (2015) sobre a acurácia de seleção de progênies frente à seleção baseada nos melhores preditores lineares não viesados (BLUP – *Best Linear Unbiased Prediction*).

2 MATERIAL E MÉTODOS

2.1 Obtenção do material genético

Na safra 2015/16 foram desenvolvidas 16 populações via hibridação artificial em esquema de dialelo parcial em casa de vegetação. O dialelo foi constituído por dois grupos de genitores (cultivares comerciais), sendo o primeiro grupo composto por cinco cultivares (G1) e o segundo por quatro cultivares (G2), conforme apresentado na tabela 1. As sementes de cada cruzamento foram colhidas separadamente, trilhadas e posteriormente semeadas em casa de vegetação durante o inverno na safra 2016/16 para multiplicação do número de sementes e obtenção das populações segregantes F₂. Para tanto, a casa de vegetação foi equipada com iluminação artificial com a finalidade de proporcionar um ambiente adequado ao bom desenvolvimento da cultura, considerando a importância do fotoperíodo para a soja, o qual atua como indutor de florescimento (Câmara, 1998).

Tabela 1 – Populações obtidas por meio de cruzamentos entre dois grupos de genitores em esquema de dialelo parcial.

Grupo 1	Grupo 2			
	M 9056 RR	M 8221 RR	P98N71	CONQUISTA
BMX Desafio RR	POP4	POP10	POP19	POP30
M7739 IPRO	-	-	POP20	POP31
P98Y30 RR	POP5	POP11	POP21	POP32
M8349 IPRO	-	-	POP22	POP33
TMG 132 RR	POP6	POP12	POP23	POP34

Simultaneamente à condução e avaliação do experimento das populações segregantes F₂ na safra 2016/17, foi realizada a seleção individual de plantas, dando origem a progênies F_{2:3}. A seleção foi realizada considerando características de sanidade, acamamento e, sobretudo, o ajuste do grupo de maturidade relativa (ALLIPRANDINI et. al, 2009) para as regiões de cultivo da cultura da soja. As progênies F_{2:3} foram semeadas no inverno do ano de 2017, as quais foram derivadas de plantas selecionadas individualmente na geração anterior. Ao atingir

a senescência, ou seja, o estágio R8 (FEHR e CAVINESS, 1977), cada parcela foi colhida individualmente em bulk, o que deu origem a um total de 218 progênies F_{2:4}. Durante tal processo, as populações POP22, POP23 e POP33 não seguiram no programa devido ao baixo desempenho para caracteres de interesse ou disponibilidade de sementes para as gerações seguintes.

2.2 Condução dos experimentos a campo e avaliação das progênies F_{2:4} a campo

Para a avaliação da produtividade de grãos das progênies na geração F_{2:4} foram realizados dois experimentos nas localidades de Sorriso – MT (12°03'05"S, 56°03'14"W e 470 m de altitude) e Capinópolis – MG (18°40'48"S, 49°33'58"O e 530 m altitude). Os ensaios de Sorriso e Capinópolis foram constituídos de 192 e 214 progênies F_{2:4}, respectivamente. Os experimentos foram conduzidos em delineamento de blocos aumentados com oito blocos, sendo que parte dos genitores envolvidos nos cruzamentos realizados para originar as populações foi utilizada como testemunhas. Para cada ensaio foram utilizadas quatro testemunhas diferentes. Cada parcela foi constituída por duas linhas de 5 m espaçadas entre si por 0,5 m. A densidade de semeadura utilizada foi de 13 plantas por metro linear. A condução do ensaio foi realizada conforme recomendações descritas por Sedyama et al. (2015). Após a colheita, a massa de grãos da parcela foi corrigida para 13% de umidade.

2.3 Análises estatísticas

Nas análises estatísticas para cada local, foi adotada a metodologia da máxima verossimilhança restrita/melhor predição linear não viesada (REML/BLUP) conforme proposto por Henderson (1975) e Patterson e Thompson (1971). Os dados foram analisados sob o ajuste de dois modelos diferentes a fim de investigar as consequências da adição dos efeitos de genitores na qualidade experimental e na acurácia das predições. O modelo ajustado apenas para o efeito de progênies como entidade genética é apresentado pela equação 1 (Eq. 1) e será referido neste trabalho como “modelo reduzido”, enquanto o modelo ajustado para o os efeitos de genitores e progênies contempladas como entidades genéticas é apresentado na equação 2 (Eq. 2) e será referido neste trabalho como “modelo completo”:

$$y_1 = X_c + T_p + S_r + e_1 \quad (\text{Eq. 1})$$

onde y_1 é o vetor dos dados fenotípicos de indivíduo, c é o vetor de efeitos fixos de testemunhas somados à média geral; p é o vetor de efeito das progênes oriundas dos cruzamentos entre os genitores dos grupos 1 e 2 (aleatório), no qual $p \sim N(0, I\sigma_p^2)$; e_1 é o vetor de erros (aleatório), no qual $e_1 \sim N(0, I\sigma_{e1}^2)$. As letras maiúsculas (\mathbf{X} , \mathbf{T} e \mathbf{S}) representam as matrizes de incidência dos referidos efeitos.

$$y_2 = \mathbf{X}_c + \mathbf{Z}_{g1} + \mathbf{W}_{g2} + \mathbf{T}_p + \mathbf{S}_r + e_2 \quad (\text{Eq. 2})$$

onde y_2 é o vetor dos dados fenotípicos de indivíduo, c é o vetor de efeitos fixos de testemunhas somados à média geral; g_1 é o vetor de efeitos devido a contribuição dos genitores do grupo 1 às progênes (aleatório), no qual $g_1 \sim N(0, I\sigma_{gca1}^2)$; g_2 é o vetor de efeitos devido a contribuição dos genitores do grupo 2 às progênes (aleatório), no qual, $g_2 \sim N(0, I\sigma_{g2}^2)$; p é o vetor de efeito das progênes oriundas dos cruzamentos entre os genitores dos grupos 1 e 2 (aleatório), no qual $p \sim N(0, I\sigma_p^2)$; e_1 é o vetor de erros (aleatório), no qual $e_2 \sim N(0, I\sigma_{e2}^2)$. As letras maiúsculas (\mathbf{X} , \mathbf{Z} , \mathbf{W} , \mathbf{T} e \mathbf{S}) representam as matrizes de incidência dos referidos efeitos.

Após serem obtidas as predições para os efeitos BLUPs das progênes e seus respectivos genitores, foi realizado o procedimento proposto por Resende (2015) para a obtenção dos valores preditos do “Índice de seleção com Genitores e Populações F1 ou F2 ou F3 em plantas autógamas”, neste trabalho referido como “Índice de seleção genitores + progênes”. O índice genitores + progênes considera as informações de genitores e progênes com o intuito de elevar a acurácia na seleção das progênes. O modelo utilizado para a obtenção do índice é descrito abaixo:

$$I_a = \hat{b}_1 \left(\frac{g_1 + g_2}{2} \right) + \hat{b}_2 g_{Fn} \quad (\text{Eq. 3})$$

Em que:

$$\hat{b}_1 = \frac{(1 - r_{aaFn}^2)}{1 - (r_{aaPais}^2 r_{aaFn}^2)} \quad (\text{Eq. 4})$$

$$\hat{b}_2 = \frac{(1 - r_{aaPais}^2)}{1 - (r_{aaPais}^2 r_{aaFn}^2)} \quad (\text{Eq. 5})$$

A acurácia do índice é dada por:

$$r_{\hat{a}aI} = \sqrt{\frac{(1-r_{\hat{a}aPais}^2)(1-r_{\hat{a}aFn}^2)}{1-(r_{\hat{a}aPais}^2 r_{\hat{a}aFn}^2)}} \quad (\text{Eq. 6})$$

onde, visando a geração F_{∞} : I_a é o valor do índice de seleção para a progênie a ; g_{Fn} é o valor genotípico da progênie em questão na geração filial n ; g_1 é o valor genotípico para o genitor do grupo 1 para a progênie em questão; g_2 é o valor genotípico para o genitor do grupo 2 para a progênie em questão; $r_{\hat{a}aPais}^2$ é o quadrado da média aritmética entre o quadrado das acurácias dos genitores $r_{\hat{a}ag1}^2$ e $r_{\hat{a}ag2}^2$; $r_{\hat{a}aFn}^2$ é o quadrado da acurácia da progênie em questão na geração filial n ; $r_{\hat{a}aI}$ é a acurácia do índice de seleção para a progênie em questão.

Para verificar as diferenças na seleção das progênies mais promissoras e descarte das progênies menos promissoras pelas predições obtidas pelo modelo completo e pelo índice de seleção, foi calculado o índice de coincidência para 10%, 15% e 20% de intensidades de seleção das melhores e piores progênies em cada ensaio. O índice de coincidência consistiu no resultado da razão entre o número de progênies coincidentes na seleção (ou descartadas) pelos dois modelos e o número total de progênies selecionadas (ou descartadas).

Todas as análises foram realizadas utilizando os Softwares SELEGEN REML/BLUP (RESENDE, 2016) e R (R CORE TEAM, 2020), sendo que o script em R para a obtenção do índice de seleção é apresentada no Apêndice 2.

3. RESULTADOS E DISCUSSÃO

3.1 Análises estatísticas para o modelo reduzido

As estimativas dos parâmetros genéticos e fenotípicos para as progênes $F_{2:4}$ avaliadas nos ensaios de Sorriso e Capinópolis sob ajuste do modelo reduzido são apresentados na Tabela 2. Para ambas as localidades foi encontrado o efeito significativo da variância entre progênes, o que releva a oportunidade da obter ganhos com a seleção entre progênes.

Tabela 2 – Estimativas de parâmetros genéticos e fenotípicos para produtividade de grãos (kg ha^{-1}) das progênes $F_{2:4}$ para os ensaios conduzidos em Sorriso – MT e Capinópolis – MG sob ajuste do modelo reduzido (Eq. 1).

Parâmetros	Estimativas	
	Sorriso - MT	Capinópolis - MG
σ^2_{prog}	449607,83 **	867282,44 **
σ^2_{rep}	65288,21	23318,47
σ^2_e	424839,06	642359,76
h^2_{prog}	0,48	0,57
r_{aa}^\dagger	0,69	0,75
CVg (%)	22,78	27,39
CVe (%)	22,13	23,57
CVr (%)	1,02	1,16
Média geral	2944,73	3400,23

** significante a 1% de probabilidade pelo teste da razão da máxima verossimilhança; σ^2_{prog} = variância entre as progênes; σ^2_{rep} = variância entre as repetições; σ^2_e = variância do erro; h^2_{prog} = herdabilidade entre progênes; r_{aa} = acurácia; CVg (%) = coeficiente de variação genético; CVe (%) = coeficiente de variação residual; CVr (%) = coeficiente de variação relativo que é calculado pela equação: $CVr (\%) = \frac{CVg (\%)}{CVe (\%)}$.

† Estimativas de acurácia obtidas pela equação $r_{\text{aa}} = \sqrt{h^2_{\text{prog}}}$.

A herdabilidade encontrada para a seleção de progênes de soja foi de 0,48 e 0,57, para os ensaios de Sorriso e Capinópolis, respectivamente e estão de acordo com o relata a literatura (VIOTTO DEL CONTE et al., 2020; VOLPATO et al., 2019; VOLPATO et al., 2018). Isto indica que a produtividade medida das progênes sofreu menor influência dos efeitos ambientais quando comparado ao ensaio de Sorriso. Tal inferência é corroborada pelas estimativas de acurácia, a qual também apresentou maior magnitude para a localidade de Capinópolis (0,75) frente ao ensaio conduzido em Sorriso (0,69), as quais são consideradas alta e média,

respectivamente segundo Resende e Duarte (2007). Os autores classificam como altas as estimativas de acurácias acima de 70%, como médias, acurácias de valores entre 30 e 70% e como baixas, abaixo de 30%.

Os coeficientes de variação residual estimados para ambos os ensaios apresentaram magnitudes semelhantes (22,13 e 23,57) e estão de acordo com o que é encontra na literatura (VIANNA et al., 2019, VOLPATO et al., 2018). Estes dados revelam uma boa qualidade experimental, uma vez que a variável estudada foi a produtividade de grãos e esta é governada por uma grande quantidade de genes (caractere quantitativo), portanto, bastante passível de ser afetado pelos efeitos ambientais e interações gênicas (PEREIRA, 2017). De acordo com Resende e Duarte, 2007, o coeficiente de variação relativo (CVr) também é uma medida de qualidade experimental para avaliações genéticas e deve ser calculada pela razão entre o coeficiente de variação genético e o coeficiente de variação residual. Para ambos os ensaios, estas estimativas apresentaram magnitude maior que a unidade, o que segundo os autores revela boa qualidade experimental.

3.2 Análises estatísticas para o modelo completo

As estimativas dos parâmetros genéticos e fenotípicos para as progênes $F_{2:4}$ sob ajuste do modelo completo para ambos os ensaios são apresentadas na Tabela 3. As estimativas de variância entre progênes apresentaram significância em ambos os ensaios, indicando a viabilidade da seleção entre progênes nos dois locais. Para o ensaio conduzido em Sorriso, não houve efeito significativo da variância entre genitores de um mesmo grupo, mas foi encontrada variância significativa para os efeitos das repetições o que indica a existência de heterogeneidade entre os blocos (TORRES et al., 2015).

Para a localidade de Capinópolis foi revelada variância significativa entre os genitores do grupo 1, mas não do grupo 2. Estes dados indicam que a informação de genitores em Capinópolis, neste caso, os do grupo 1, podem ser considerados na escolha das progênes mais promissoras. Isso porque os genitores com valores genéticos preditos positivos e de maior magnitude dentro deste grupo são os que deram origem às progênes mais produtivas, logo possuem uma concentração de alelos favoráveis expressos neste local maior que os demais genitores do mesmo grupo (CRUZ, 2005)

Tabela 3 - Estimativas de parâmetros genéticos e fenotípicos para produtividade de grãos (kg ha⁻¹) das progênes F_{2:4} para os ensaios conduzidos em Sorriso – MT e Capinópolis – MG sob ajuste do modelo completo (Eq. 2).

Parâmetros	Estimativas	
	Sorriso - MT	Capinópolis - MG
σ^2_{prog}	446770,08 **	747258,68 **
σ^2_{g1}	3787,97	321829,05 *
σ^2_{g2}	2347,44	16117,49
σ^2_{rep}	66581,42 **	22059,25
σ^2_e	424129,79	643883,39
h^2_{prog}	0,47	0,43
h^2_{g1}	0,00	0,18
h^2_{g2}	0,00	0,01
h^2_{geral}	0,48	0,62
$r_{\hat{a}prog}$	0,70	0,65
$r_{\hat{a}geral}$	0,70	0,79
CVg (%)	30,22	36,63
CVe (%)	22,12	23,60
CVr (%)	1,37	1,55
Média geral	2944,73	3400,23

* e ** significante a 5% e 1% de probabilidade pelo teste da razão da máxima verossimilhança; σ^2_{prog} = variância entre as progênes; σ^2_{g1} = variância entre os genitores do grupo 1; σ^2_{g2} = variância entre os genitores do grupo 2; σ^2_{rep} = variância entre as repetições; σ^2_e = variância do erro; h^2_{prog} = herdabilidade entre progênes; h^2_{g1} = herdabilidade entre os genitores do grupo 1; h^2_{g2} = herdabilidade entre os genitores do grupo 2; $h^2_{geral} = h^2_{prog} + h^2_{g1} + h^2_{g2}$; $r_{\hat{a}prog}$ = acurácia na seleção de progênes; $r_{\hat{a}geral}$ = acurácia geral (soma das acurácias de progênes e grupos de genitores); CVg (%) = coeficiente de variação genético; CVe (%) = coeficiente de variação residual; CVr (%) = coeficiente de variação relativo que é calculado pela equação: $CVr (\%) = \frac{CVg (\%)}{CVe (\%)}$.

† Estimativas de acurácia obtidas pela equação $r_{\hat{a}a} = \sqrt{h^2_{prog}}$.

A herdabilidade encontrada para as progênes no ensaio de Sorriso foi de 0,47, enquanto as estimativas de herdabilidade para os genitores foram muito próximas de zero. Para o ensaio de Capinópolis as herdabilidades estimadas para progênes, grupo 1 e grupo 2 de genitores foram de 0,43, 0,18 e 0,01, respectivamente. Os valores de herdabilidade encontrados para a avaliação da produtividade de grãos em soja estão de acordo com os encontrados na literatura (VIOTTO DEL CONTE et al., 2020; Volpato et al., 2019; Volpato et al., 2018).

A acurácia para seleção de progênes e a acurácia geral, a qual resulta da soma da acurácia de progênes e dos grupos de genitores, foram estimadas em 0,70 para Sorriso, não

diferindo entre si, devido às informações dos genitores não terem contribuído para apreciação dos efeitos genotípicos das progênes. Para os dados tomados em Capinópolis a acurácia para a seleção de progênes foi de 0,65 e a acurácia considerando todas as entidades genéticas contempladas no modelo, foi de 0,79, o que evidencia uma vantagem em considerar as informações de genitores para o processo de seleção de progênes, sobretudo em fases iniciais do programa. As acurácias gerais para ambos os ensaios podem ser consideradas altas de acordo com Resende e Duarte (2007) e apontam boa qualidade experimental.

Os coeficientes de variação residual estimados para os ensaios conduzidos nas diferentes localidades apresentaram magnitudes semelhantes (22,12 e 23,60) e estão de acordo com o que se encontra na literatura (VIANNA et al., 2019; VOLPATO et al., 2018). Os coeficientes de variação relativos apontam a boa qualidade experimental para ambos os ensaios, pois são superiores a um. Tal magnitude para o CVr indica que a variação devido aos efeitos genéticos é superior à variação devido aos efeitos residuais, portanto existe a possibilidade de sucesso com a seleção (VENCOSKY, 1987).

3.3 Modelo reduzido vs modelo completo

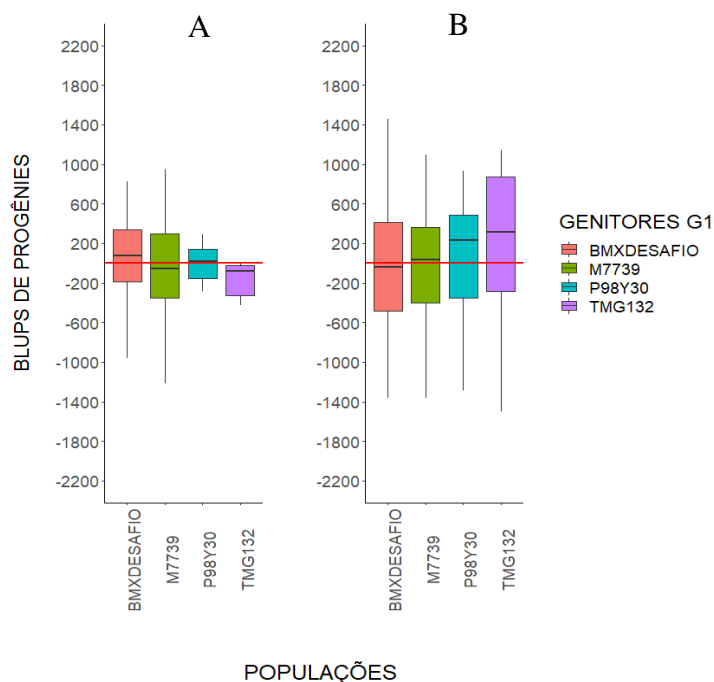
As estimativas para o coeficiente de variação residual entre os diferentes modelos ajustados para cada local foram de magnitudes muito próximas, indicando que a adesão de informações de genitores para a melhora desta medida de qualidade experimental não foi efetiva. No entanto, o aumento dos coeficientes de variação relativos calculados para Sorriso (1,02 no modelo reduzido e 1,37 no modelo completo) e Capinópolis (1,16 no modelo reduzido e 1,55 no modelo completo) mostram a melhora da qualidade experimental para avaliações genéticas com o uso do modelo completo. Resende e Duarte (2007) argumentam que quanto maior o CVr, melhores deverão ser as acurácias obtidas com o experimento, e, portanto, mais efetiva deve ser a seleção.

Não foram notadas melhoras nas medidas de herdabilidade de progênes (h^2_{prog}) e acurácia para a seleção de progênes (r_{aaprog}) do modelo reduzido para o completo para ambas as localidades. O fato é que para o ensaio conduzido em Capinópolis, a herdabilidade estimada para a seleção de progênes reduziu de 0,57 para (modelo reduzido) 0,43 do (modelo completo) enquanto a acurácia reduziu de 0,75 (modelo reduzido) para 0,65 (modelo completo). No entanto, quando se adiciona informações de genitores no modelo, é adequado que sejam considerados também os parâmetros genéticos e fenotípicos destas entidades para se avaliar a

qualidade experimental e para realizar o processo de seleção. Sendo assim, foram estimadas a herdabilidade geral (h^2_{geral}) e a acurácia geral (r_{aageral}), somando-se as herdabilidades e acurácias de todas as entidades genética contempladas no modelo completo. Para o ensaio de Capinópolis, quando são comparadas a herdabilidade do modelo reduzido (0,57) frente a herdabilidade geral do modelo completo (0,62) e a acurácia do modelo reduzido (0,75) com a acurácia geral do modelo completo (0,79), verificamos que houve aumento na magnitude para estas estimativas. Tal diferença nestas estimativas, proporciona ao melhorista ferramentas para acessar com melhor qualidade a variabilidade genética contida entre as progênes, pois desassocia do erro a porção da variância devido às variações genéticas e obtém previsões mais fidedignas, portanto, permite realizar a seleção com maior confiabilidade (RESENDE, 2015).

Estas discussões se aplicam apenas para o ensaio conduzido em Capinópolis, e não ao ensaio de Sorriso, uma vez que não houve diferença das estimativas de herdabilidade ou acurácia entre o ajuste do modelo reduzido ou completo. Isto se deu, pois, as estimativas de variância para ambos os grupos de genitores em Sorriso foram de baixa magnitude e não significativa, logo, não houve uma porção expressiva da variância fenotípica explicada pela variância entre os genitores de um mesmo grupo, conforme elucidado na Figura 1. Por consequência, as estimativas de herdabilidade e acurácia para estas entidades genéticas, se aproximaram de zero e não contribuem para a seleção.

Figura 1 – Distribuição de dos valores genotípicos (BLUP) das progênes agrupados por seus respectivos genitores do grupo 1 para as localidades de Sorriso – MT (A) e Capinópolis – MG



3.1 Modelo completo vs Índice de seleção genitores + progênes

Ao comparar as estimativas de acurácia de seleção obtidas pelo ajuste do modelo completo e aquelas obtidas pelo índice de seleção, é possível notar que o índice de seleção supera as estimativas de acurácia do modelo completo nos quesitos acurácia mínima, média e máxima individual (acurácia referente à seleção de determinada progênie) para ambos os locais (Tabela 4). As maiores acurácias de seleção encontradas pela aplicação do índice oferecem maior confiança ao melhorista no processo de seleção (Resende, 2015) e proporcionam maiores ganhos genéticos (COBB et al., 2019) em relação aos modelos de menor acurácia.

Tabela 4 – Estimativas de acurácia mínima, média e máxima para o modelo completo e o índice de seleção genitores + progênes.

	Estimativas de acurácia			
	Sorriso - MT		Capinópolis - MG	
	Modelo completo	Índice de seleção	Modelo completo	Índice de seleção
Média	0,70 [†]	0,83	0,68	0,86
Máximo	0,71	0,84	0,83	0,91
Mínimo	0,00	0,75	0,00	0,80

[†] Estimativas de acurácia obtidas pela equação $r_{\hat{a}a} = \sqrt{1 - \frac{PEV}{\sigma_{prog}^2}}$.

A Tabela 5 contém os índices de coincidência entre as progênes selecionadas e descartadas para as intensidades de seleção de 10%, 15% e 20%, as quais são intensidades de seleção comumente aplicadas nesta fase dos programas de melhoramento de autógamias, uma vez que após a seleção de progênes, pretende-se concentrar recursos e esforços em um menor número de progênes, aumentando o volume de sementes e testando-as em um maior número de locais (COBB et al., 2019; SILVA, et al., 2017, BERNARDO, 2010).

Tabela 5 – Índice de coincidência entre as progênes selecionadas entre as predições obtidas sob ajuste do modelo completo e do índice de seleção genitores + progênes para os ensaios conduzidos em Sorriso – MT e Capinópolis – MG, considerando as intensidades de seleção de 10, 15 e 20%.

Intensidades de seleção	Sorriso - MT		Capinópolis - MG	
	Seleção	Descarte	Seleção	Descarte
10%	89%	100%	74%	84%
15%	97%	97%	82%	85%
20%	95%	100%	86%	89%

Para o ensaio conduzido em Sorriso, as coincidências computadas para seleção e descarte variaram entre 89% e 100%, o que indica que mesmo com a melhora da estimativa da acurácia para seleção, não houve alteração significativa no ranqueamento das progênes, logo, conduzir a seleção baseado nos BLUPs das progênes ou nos valores do índice de seleção, não deveria alterar significativamente a composição do germoplasma para a condução do programa. Para esta localidade, a coincidência mínima encontrada entre as duas estratégias de seleção foi de 89% para a seleção com intensidade de 10%, o que representa a seleção divergente de duas progênes em um total de 19 selecionados.

Para a localidade de Capinópolis a coincidência de seleção e descarte foi tão menor quanto maior a intensidade de seleção imposta sobre as progênes. Para a seleção de 10% das progênes nesta localidade, foi observada a coincidência de 74% entre modelo completo e índice de seleção, o que representa de seis progênes selecionadas erroneamente em 21 possíveis.

De modo geral, os índices de coincidência calculados para o ensaio conduzido em Sorriso, mostram-se superiores àqueles calculados para o ensaio conduzido em Capinópolis. Portanto é evidente que para o ensaio conduzido em Capinópolis, além do índice proporcionar informações para uma seleção mais acurada, a alteração de ranqueamento das progênes promove alterações mais significativas na condução do programa, quando comparado ao ensaio de Sorriso.

De acordo com Bernardo (2020), no melhoramento de plantas os futuros genitores que serão utilizados nos próximos blocos de cruzamentos são selecionados primeiro por seu valor genotípico (mérito genético), os quais são mensurados em ensaios de avaliação de rendimento para o caractere de interesse. Nestes mesmos ensaios, as progênes de melhor mérito genético também são eleitas para seguirem no programa com o objetivo de gerar novas linhagens e entregá-las ao produtor (HICKEY, 2019). Sendo assim, a seleção precoce e acurada de progênes de bom mérito genético - conforme proposto pelo índice genitores + progênes utilizado neste trabalho - reduz o ciclo de melhoramento do programa que é o parâmetro mais expressivo e mais fácil de ser manejado para aumentar o ganho genético (COBB, 2019), seja para reduzir o tempo entre cruzamento a cruzamento ou o tempo entre o cruzamento e a adoção da cultivar.

Estes resultados nos levam a recomendar a adoção das fontes de variação de genitores na avaliação de progênes como rotina do programa de melhoramento, as quais podem ser obtidas a partir dos dados das progênes (HALLAUER, 2011; FALCONER e MACKAY, 1996) e são de fácil gerenciamento, uma vez que o programa de melhoramento já se preocupa em preservar a informação de pedigree por todo o programa. Com o emprego destas informações

no modelo, há a possibilidade de melhora na qualidade da predição dos valores genéticos e da qualidade experimental, principalmente quando há efeito significativo para as variâncias entre os genitores. Conforme relatado em diversos trabalhos encontrados na literatura, a inclusão de informações diferentes entidades genéticas no modelo, tais como os de genitores ou populações, podem causar alterações no ranqueamento das progênes avaliadas e nas estimativas dos componentes de variância (DUARTE e VENCOVSKY, 2001; PEREIRA et al., 2017; VOLPATO et al., 2020). Neste trabalho as informações de genitores podem ser interpretadas como informações de populações, pois ambas representariam as mesmas entidades, uma vez que cada população é formada por uma combinação única entre dois genitores. Ao agregar aos modelos as informações de genitores ao invés de populações, possibilitou o uso do índice de seleção proposto por Resende (2015).

4. CONCLUSÕES

A inclusão da informação referentes aos genitores (modelo completo) no modelo reduzido é de fácil execução, conduz a melhores inferências sobre às progênes e oferece a oportunidade de uma seleção mais acurada;

As informações de genitores contempladas no modelo são mais efetivas quando há efeitos significativos para os grupos de genitores, mas não prejudicam a qualidade das análises quando isto não acontece;

O índice de seleção genitores + progênes apresentou melhores estimativas de acurácia em relação ao modelo completo, o que deve conduzir o melhorista à maiores ganhos genéticos.

5 REFERÊNCIAS BIBLIOGRÁFICAS

BAUER, A. M.; LÉON, J. Multiple-trait breeding values for parental selection in self-pollinating crops. **Theoretical and Applied Genetics**, v. 116, n. 2, p. 235–242, jan. 2008.

BERNARDO, R. Reinventing quantitative genetics for plant breeding: something old, something new, something borrowed, something BLUE. **Heredity**, v. 125, n. 6, p. 375–385, dez. 2020.

BERNARDO, R. N. **Breeding for quantitative traits in plants**. Woodbury, Minn.: Stemma Press, 2010.

BORÉM, A.; MIRANDA, G. V.; FRITSCHÉ-NETO, R. **Melhoramento de plantas**. 7. ed. Viçosa, MG: Editora UFV, 2017.

CECCARELLI, S. Efficiency of Plant Breeding. **Crop Science**, v. 55, n. 1, p. 87–97, jan. 2015.

COBB, J. N. et al. Enhancing the rate of genetic gain in public-sector plant breeding programs: lessons from the breeder's equation. **Theoretical and Applied Genetics**, v. 132, n. 3, p. 627–645, mar. 2019.

CONAB. **Série Histórica**. Disponível em: <<http://www.conab.gov.br/info-agro/safras/serie-historica-das-safras/item/7666-soja/>>. Acesso em: 4 jul. 2021.

DUARTE, J. B.; VENCOVSKY, R. Estimaco e predico por modelo linear misto com ênfase na ordenaco de mdias de tratamentos genticos. **Scientia Agricola**, v. 58, n. 1, p. 109–117, mar. 2001.

FALCONER, D. S.; MACKAY, T. **Introduction to quantitative genetics**. 4. ed., [16. print.] ed. Harlow: Pearson, Prentice Hall, 1996.

GOULDEN, C. H. **Problems in plant selection**. Proceedings of the Seventh International Genetics Congress. **Anais...** In: SEVENTH INTERNATIONAL GENETICS CONGRESS. Cambridge: Cambridge Univ. Press, 1939.

HALLAUER, A. R. Evolution of plant breeding. **Crop Breeding and Applied Biotechnology**, v. 11, n. 3, p. 197–206, set. 2011.

HALLAUER, A. R.; DARRAH, L. L. Compendium of recurrent selection methods and their application. **Critical Reviews in Plant Sciences**, v. 3, n. 1, p. 1–33, jan. 1985.

HENDERSON, C. R. Best Linear Unbiased Estimation and Prediction under a Selection Model. **Biometrics**, v. 31, n. 2, p. 423–447, 1975.

HICKEY, L. T. et al. Breeding crops to feed 10 billion. **Nature Biotechnology**, v. 37, n. 7, p. 744–754, jul. 2019.

LIRA, E. G. et al. Genetic parameters, phenotypic, genotypic and environmental correlations and genetic variability on sunflower in the Brazilian Savannah. **Cincia Rural**, v. 47, n. 8, 2017.

PATTERSON, H. D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v. 58, n. 3, p. 545–554, 1971.

PEREIRA, F. DE C. et al. Implications of the population effect in the selection of soybean progeny. **Plant Breeding**, v. 136, n. 5, p. 679–687, 2017.

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, 2020.

RAMALHO, M. A. P. et al. **Aplicações da Genética Quantitativa no Melhoramento de Plantas Autógamas**. Lavras, MG: UFLA, 2012.

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

RESENDE, M. D. V. DE. Software Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology**, v. 16, n. 4, p. 330–339, dez. 2016.

RESENDE, M. D. V. DE. **Genética Quantitativa e de Populações**. 1. ed. Viçosa, MG: Produção Independente, 2015.

RIBAUT, J.-M.; HOISINGTON, D. Marker-assisted selection: new tools and strategies. **Trends in Plant Science**, v. 3, n. 6, p. 236–239, jun. 1998.

SEDIYAMA, T.; SILVA, F. L.; BORÉM, A. **Livro Soja do Plantio à Colheita - UFV**. Viçosa, MG: Editora UFV, 2015.

SILVA, F. L. DA et al. Soybean breeding. New York, NY: Springer Science+Business Media, 2017.

SILVA, F. L. DA et al. **Soybean breeding**. New York, NY: Springer Science+Business Media, 2017.

TODESCHINI, M. H. et al. Soybean genetic progress in South Brazil: physiological, phenological and agronomic traits. **Euphytica**, v. 215, n. 7, p. 124, jul. 2019.

TORRES, F. E. et al. Número de repetições para avaliação de caracteres em genótipos de feijão-caupi. **Bragantia**, v. 74, n. 2, p. 161–168, maio 2015.

VENCOCKSKY, R. Herança quantitativa. In: **Melhoramento e produção de milho**. 2. ed. Campinas, SP: Fundação Cargil, 1987. p. 795.

VIANNA, M. S. et al. Selection of lineages, genetic parameters, and correlations between soybean characters. **Bioscience Journal**, v. 35, n. 5, 27 jun. 2019.

VOTTO DEL CONTE, M. et al. Overcoming collinearity in path analysis of soybean [*Glycine max* (L.) Merr.] grain oil content. **PLOS ONE**, v. 15, n. 5, p. e0233290, 22 maio 2020.

VOLPATO, L. et al. Inference of population effect and progeny selection via a multi-trait index in soybean breeding. **Acta Scientiarum. Agronomy**, v. 43, p. e44623, 14 ago. 2020.

VOLPATO, L. et al. Multi-trait multi-environment models in the genetic selection of segregating soybean progeny. **PLOS ONE**, v. 14, n. 4, p. e0215315, 18 abr. 2019.

VOLPATO, L. et al. Selection of inbred soybean progeny (*Glycine max*): an approach with population effect. **Plant Breeding**, v. 137, n. 6, p. 865–872, dez. 2018.

WANGA, M. A. et al. Opportunities and challenges of speed breeding: A review. **Plant Breeding**, v. 140, n. 2, p. 185–194, abr. 2021.

NAGATOSHI, Y.; FUJITA, Y. Accelerating Soybean Breeding in a CO₂-Supplemented Growth Chamber. **Plant and Cell Physiology**, v. 60, n. 1, p. 77–84, 1 jan. 2019.

CONCLUSÕES GERAIS

O emprego do BLUPISM com valores dos efeitos genéticos das capacidades geral e específica de combinação se mostrou uma estratégia útil e praticável para a seleção de populações F₂ de soja;

O BLUPISM conduz a seleção de um número de indivíduos em cada população muito semelhante ao BLUPI que é considerada uma metodologia ótima para seleção de indivíduos em populações;

A inclusão da informação referentes aos genitores é de fácil execução, conduz a melhores inferências sobre as progênies e oferece a oportunidade de uma seleção mais acurada;

O índice de seleção genitores + progênies apresentou melhores estimativas de acurácia em relação à seleção de progênies baseada no BLUP obtido pelo ajuste do modelo que contempla informações de progênies e genitores.

APÊNDICES

Apêndice 1

```
#####
####                               ####
#### Best Linear Unbiased Predictor Individual   ####
#### Simulated Modified - BLUPISM              ####
#### para o número total                       ####
#### de indivíduos selecionados                 ####
####                                           ####
#####

##### Preparo do arquivo de dados
### O arquivo de dados deve conter duas colunas:
## GEN_ID: Nomenclatura das populações ou
# famílias;
## g: Valor genotípico ou valor dos efeitos
#genotípico da respectiva população/família.

#Encontrar o arquivo "datapar1"
{VGpop <- read.table(file.choose(), header = T)}

ntot<-144

n_BLUPISM <- function(VGpop, ntot)
{
  nini<-1
  i=1
  {{VGpop<-VGpop[with(VGpop, order(-VGpop$g)), ]
  popm1<-diag(VGpop[,2])
  popm2<-diag((popm1[1,1]),(nrow(VGpop)),(nrow(VGpop)))
  popm1.1<-as.matrix(diag(popm1))
  popm1.2<-rbind(popm1.1[1,],popm1.1)
  popm3<-diag(popm1.2[-(nrow(popm1.2)),])
  bm1<-as.matrix(diag(popm1%*(solve(popm2))))
  bm11<-as.matrix(diag(popm1%*(solve(popm3))))
  bm2<-as.data.frame(bm1)
  bm22<-as.data.frame(bm11)
  bm3<-cbind(VGpop[, -1],bm2)
  bm33<-cbind(VGpop[, -1],bm22)
  colnames(bm3)<-c("BLUP_pop", "BLUPISM_rate")
  colnames(bm33)<-c("BLUP_pop", "BLUPISM_rate")
  rownames(bm3)<-rownames(VGpop)
  rownames(bm33)<-rownames(VGpop)
  bm3$BLUP_pop[bm3$BLUP_pop <= 0 ] <- 0
  bm33$BLUP_pop[bm33$BLUP_pop <= 0 ] <- 0
  bm3$BLUPISM_rate[bm3$BLUPISM_rate <= 0 ] <- 0
  bm33$BLUPISM_rate[bm3$BLUPISM_rate <= 0 ] <- 0}
  {bm62<-as.data.frame(bm33[,2])
  k=1
  numind<-as.matrix(nini)
  while (k < ntot) {
    i=1
    numind<-as.matrix(nini)
    for (i in (1:(nrow(bm62)-1))) {
      numi<-((bm62[(1+i),])* (numind[i,1]))
```

```

        numir<-as.data.frame(round(numi))
        numind<-(rbind(numind,numir[1,]))
    }
    k<-colSums(numind)
    diff<-ntot-k
    nini<-as.matrix(nini+(ceiling(0.001*diff)))
    print(k)
}
row.names(numind) <- VGpop$GEN_ID
colnames(numind) <- "Nº Ind"
BLUPISMt <- numind
BLUPISMt}
cat("O número total de indivíduos selecionados: ", k,
    "\nO número total de indivíduos solicitado:  ", ntot,
    "\nDiferença :                               ", -1*diff,
    "\n\nNúmero de indivíduos selecionados dentro de cada\n
        população ou família\n\n")
BLUPISMt}
}

n_BLUPISM(VGpop, ntot)

```

Apêndice 2

```

# Para a confecção do índice de Seleção Populações + Genitores serão necessários 3
arquivos:
# - "datapop" - dados referentes às populações/progênes;
# - "datapar1" - dados referentes aos parentais do grupo 1 (dialelo parcial ou
esquemas fatoriais) ou parentais polinizados;
# - "datapar2" - dados referentes aos parentais do grupo 2 (dialelo parcial ou
esquemas fatoriais) ou parentais polinizadores;
# Cada arquivo deve conter 3 colunas, conforme descrito abaixo:
# - "Name" - nome dado às populações/progênes ou parentais dos grupos 1
(polinizados) ou 2 (polinizadores);
# - "BLUP" - BLUP (valor genético) das populações/progênes ou parentais dos grupos
1 (polinizados) ou 2 (polinizadores);
# - "PEV" - PEV populações/progênes ou parentais dos grupos 1 (polinizados) ou 2
(polinizadores).
#Encontrar o arquivo "datapop"
{datapop <- file.choose()
  source(file.choose())
  datapop$Name <- as.factor(datapop$Name)}
#Encontrar o arquivo "datapar1"
{datapar1 <- file.choose()
  source(file.choose())
  datapar1$Name <- as.factor(datapar1$Name)}
#Encontrar o arquivo "datapar2"
{datapar2 <- file.choose()
  source(file.choose())
  datapar2$Name <- as.factor(datapar2$Name)}
#Digite abaixo a variância genética estimada das populações/progênes
varpop <- ??
#Digite abaixo a variância genética estimada para o grupo 1 ou parentais polinizados
vargp1 <- ??
#Digite abaixo a variância genética estimada para o grupo 2 ou parentais
polinizadores
vargp2 <- ??
#Rode os comandos a partir desta linhas
PEV<-m1$PevU
PEVPOP<-as.matrix<-PEV$PROG$KG.HA
PEVPOP<-as.vector<-diag(PEVPOP)
VAPOP<-c(rep(varpop,length(PEVPOP)))
acPOP<-as.data.frame(cbind((levels(data$PROG)),(c(rep(1,length(PEVPOP)))-
(PEVPOP/VAPOP))))
PEVMOM<-as.matrix<-PEV$IDMOM$KG.HA
PEVMOM<-as.matrix<-diag(PEVMOM)
PEVDAD<-as.matrix<-PEV$IDDDAD$KG.HA
PEVDAD<-as.matrix<-diag(PEVDAD)
PEVGENITOR<-as.matrix<-c(PEVMOM,PEVDAD)
VAGENITOR<-c(rep((vargp1+vargp2)/2,length(PEVGENITOR)))
acGENITOR<-
as.data.frame(cbind(c(levels(data$IDMOM),levels(data$IDDDAD)),(c(rep(1,length(PEVGE
NITOR)))-(PEVGENITOR/VAGENITOR))))
ac<-as.data.frame(rbind(acGENITOR,acPOP))
ac$V1<-as.character(ac$V1)
ac<-ac[with(ac, order(V1)), ]
acn<-as.matrix(ac$V2)
acn<-as.numeric(acn)
ac<-cbind(ac,acn)
ac$acn<-round(ac$acn,8)

```

```

ac<-ac[,-2]
ac[ac=="0"]<-NA
ac<-ac[complete.cases(ac),]
newc<-
cbind((as.data.frame(data$PROG)),(as.data.frame(data$IDMOM)),(as.data.frame(data$IDDDAD)))
newc<-as.matrix(newc)
newc<-newc[!duplicated(newc),]
newc[newc==0]<-NA
newc<-data.frame(newc[complete.cases(newc),])
newc1<-newc
rownames(newc1)<-newc1$data.PROG
newc1<-merge(newc1,ac,by.x="data.IDMOM",by.y="V1",all.x=T)
newc1<-merge(newc1,ac,by.x="data.IDDDAD",by.y="V1",all.x=T)
newc1<-merge(newc1,ac,by.x="data.PROG",by.y="V1",all.x=T)
rownames(newc1)<-newc1$data.PROG
newc1<-newc1[,c(-1,-2,-3)]
colnames(newc1)<-c("data.IDMOM","data.IDDDAD","data.PROG")
newc1[newc1==0]<-NA
newc1<-data.frame(newc1[complete.cases(newc1),])
Fst<-
diag(c(rep(((m1$sigma$PROG)/(m1$sigma$IDMOM+m1$sigma$IDDDAD)),(length(newc1$data.PROG))))))
ac2pop<-((newc1$data.PROG))
ac2mom<-((newc1$data.IDMOM))
ac2dad<-((newc1$data.IDDDAD))
ac2par<-((ac2mom*0.5)+(ac2dad*0.5))
vg<-
(rbind(as.matrix(m1$U$PROG$KG.HA),as.matrix(m1$U$IDMOM$KG.HA),as.matrix(m1$U$IDDDAD$KG.HA)))
vg<-as.data.frame(cbind(rownames(vg),vg))
vg[] <- lapply(vg, gsub, pattern='PROG', replacement='')
vg[] <- lapply(vg, gsub, pattern='IDMOM', replacement='')
vg[] <- lapply(vg, gsub, pattern='IDDDAD', replacement='')
vg<-vg[with(vg, order(V1)), ]
vgn<-as.matrix(vg[,2])
vgn<-as.numeric(vgn)
vg<-cbind(vg,vgn)
vg$vgn<-round(vg$vgn,8)
vg<-vg[,-2]
vg[vg=="0"]<-NA
vg<-vg[complete.cases(vg),]
rownames(vg)<-vg$V1
newc2<-newc
rownames(newc2)<-newc2$data.PROG
newc2<-merge(newc2,vg,by.x="data.IDMOM",by.y="V1",all.x=T)
newc2<-merge(newc2,vg,by.x="data.IDDDAD",by.y="V1",all.x=T)
newc2<-merge(newc2,vg,by.x="data.PROG",by.y="V1",all.x=T)
rownames(newc2)<-newc2$data.PROG
newc2<-newc2[,c(-1,-2,-3)]
colnames(newc2)<-c("data.IDMOM","data.IDDDAD","data.PROG")
newc2<-newc2[complete.cases(newc2),]
#newc2 tem as dimensões corretas depois de ficar apenas POP e eliminar os pesos "0"
vgpop<-((as.matrix(newc2$data.PROG)))
vgpar<-
as.matrix(((as.numeric((as.vector(newc2$data.IDDDAD)))+(as.numeric((as.vector(newc2$data.IDMOM))))))/2)
w<-acPOP
wn<-as.matrix(w[,2])

```

```

wn<-as.numeric(wn)
w<-cbind(w,wn)
w$wn<-round(w$wn,8)
w<-w[,-2]
w[w=="0"]<-NA
w<-w[complete.cases(w),]
w<-w[with(w, order(V1)), ]
w<-w[,1]
ac2pop[is.na(ac2pop)] <- 0
ac2pop<-diag(ac2pop)
ac2par<-diag(ac2par)
dim(ac2pop)
dim(ac2par)
b1<-(((diag(1,(length(newc1$data.PROG)),(length(newc1$data.PROG))))-
      (ac2pop))%*(solve(((diag(1,(length(newc1$data.PROG)),
                          (length(newc1$data.PROG))))-(ac2pop*ac2par))))
b2<-(((diag(1,(length(newc1$data.PROG)),(length(newc1$data.PROG))))-(ac2par))%*
      (solve(((diag(1,(length(newc1$data.PROG)),(length(newc1$data.PROG))))-
              (ac2par*ac2pop))))
acIS<-diag(sqrt(((diag(1,(length(newc1$data.PROG)),(length(newc1$data.PROG))))
                -(((diag(1,(length(newc1$data.PROG)),(length(newc1$data.PROG))))-
                    ac2par))%*%
                ((diag(1,(length(newc1$data.PROG)),(length(newc1$data.PROG))))-(ac2pop))%*%
                (solve(((diag(1,(length(newc1$data.PROG)),(length(newc1$data.PROG))))-
                        (ac2pop%*%ac2par))))))))
dim(b1)
dim(vgpar)
dim(b2)
dim(vgpop)
head(newc1)
head(newc2)
SI<-((b1%vgpar)+(b2%vgpop))
SI<-data.frame(SI,row.names=row.names(newc1))
Selec.Index<-cbind(SI,acIS)
colnames(Selec.Index)<-c("Selec. Index","Accuracy")
BLUPpop<-data.frame((vgpop),row.names=row.names(newc1))
colnames(BLUPpop)<- "BLUPpop"
blupdad<-cbind(as.vector(newc2$data.IDDAD),as.vector(newc1$data.IDDAD))
blupdad<-blupdad[!duplicated(blupdad),]
BLUPdad<-as.data.frame(blupdad[,1])
colnames(BLUPdad)<- "BLUPdad"
rownames(BLUPdad)<-blupdad[,2]
blupmom<-cbind(as.vector(newc2$data.IDMOM),as.vector(newc1$data.IDMOM))
blupmom<-blupmom[!duplicated(blupmom),]
BLUPmom<-as.data.frame(blupmom[,1])
colnames(BLUPmom)<- "BLUPmom"
rownames(BLUPmom)<-blupmom[,2]
SI<-list(BLUPpop=BLUPpop,
          BLUPdad=blupdad,BLUPmom=blupmom,
          Selec.Index=Selec.Index[with(Selec.Index,
                                         order(-Selec.Index`Selec.
Index`)), ],
          Fst=as.vector(diag(Fst)))
SI$Selec.Index

```