

SAULO FABRÍCIO DA SILVA CHAVES

**MODELOS PARA ANÁLISE DE MEDIDAS REPETIDAS APLICADOS NO
MELHORAMENTO DO CUPUAÇUZEIRO**

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

Orientador: Luiz Antônio dos Santos Dias

Coorientador: Rodrigo Silva Alves

**VIÇOSA - MINAS GERAIS
2022**

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

T

C512m
2022

Chaves, Saulo Fabrício da Silva, 1997-
Modelos para análise de medidas repetidas aplicados no
melhoramento do cupuaçuzeiro / Saulo Fabrício da Silva
Chaves. – Viçosa, MG, 2022.
1 dissertação eletrônica (63 f.): il. (algumas color.).

Inclui apêndice.

Orientador: Luiz Antonio dos Santos Dias.

Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Agronomia, 2022.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2022.153>

Modo de acesso: World Wide Web.

1. *Theobroma grandiflorum* - Melhoramento genético.
2. Máxima verossimilhança. 3. Modelos multiníveis (Estatística).
4. Cupuaçuzeiro - Seleção . I. Dias, Luiz Antonio dos Santos,
1957-. II. Universidade Federal de Viçosa. Departamento de
Agronomia. Programa de Pós-Graduação em Genética e
Melhoramento. III. Título.

CDD 22. ed. 634.42

Bibliotecário(a) responsável: Alice Regina Pinto CRB6 2523

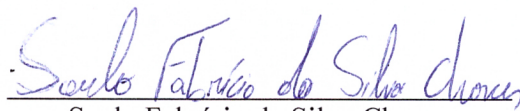
SAULO FABRÍCIO DA SILVA CHAVES

**MODELOS PARA ANÁLISE DE MEDIDAS REPETIDAS APLICADOS NO
MELHORAMENTO DO CUPUAÇUZEIRO**

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

APROVADA: 22 de fevereiro de 2022.

Assentimento:



Saulo Fabrício da Silva Chaves
Autor



Luiz Antônio dos Santos Dias
Orientador

Aos meus pais, Rubens e Ana Chaves

AGRADECIMENTOS

A Deus, por tudo.

Aos meus pais, Rubens e Ana Chaves, por todo amor, carinho, dedicação e confiança, base para minha jornada.

À minha família, pelo incentivo à busca ao conhecimento e formação pessoal. À minha companheira, Ariane Carneiro, pelo apoio e suporte emocional durante esta fase. Aos amigos de Belém e dos laboratórios de biometria e agroenergia, pelo companheirismo e parceria.

Aos meus orientadores, Luiz Antônio dos Santos Dias e Rodrigo Silva Alves, pelos ensinamentos, parceria e confiança, fundamentais para a conclusão desta etapa.

Ao pesquisador Rafael Moysés Alves, não só pela cessão dos dados, mas também por todo o conhecimento compartilhado, confiança depositada e amizade construída.

Aos estimados membros da banca avaliadora, professor dr. Kaio Dias e dr. Rafael Alves, pelas contribuições para a melhoria da qualidade desta dissertação.

À Universidade Federal de Viçosa e ao Programa de Pós-graduação em Genética e Melhoramento, pela oportunidade ímpar de fazer parte destas estimadas instituições.

À todos que contribuíram direta e indiretamente para a minha formação pessoal e profissional.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

RESUMO

CHAVES, Saulo Fabrício da Silva, M.Sc., Universidade Federal de Viçosa, fevereiro de 2022. **Modelos para análise de medidas repetidas aplicados no melhoramento do cupuaçuzeiro.** Orientador: Luiz Antônio dos Santos Dias. Coorientador: Rodrigo Silva Alves.

A modernização da cadeia produtiva do cupuaçuzeiro (*Theobroma grandiflorum*), espécie perene frutífera com ciclo de melhoramento longo, depende da adoção de eficientes estratégias de melhoramento e seleção. Os caracteres agronômicos da espécie geralmente envolvem medidas repetidas, necessitando determinar o número ótimo de safras para uma seleção genética acurada, bem como, estudar a adaptabilidade e estabilidade temporal dos híbridos, fundamentais para a seleção genética. Neste sentido, o emprego de métodos estatísticos adequados é imprescindível para o alcance de resultados críveis e úteis ao programa de melhoramento. Esta dissertação objetivou avaliar diferentes modelos e estratégias de seleção, visando a determinação do número ótimo de safras a serem avaliados (Capítulo 1) e da estrutura de covariância mais adequada (Capítulo 2), além da seleção genética de híbridos (ambos capítulos). Com este fim, uma população com 34 híbridos triplos de cupuaçuzeiro, delineada em blocos casualizados e com número variável de repetições, foi avaliada durante nove safras consecutivas quanto a produtividade e resistência à vassoura-de-bruxa. Os procedimentos estatísticos foram baseados em modelos lineares mistos. Em todos os cenários, a flexibilização das estruturas de (co)variância dos modelos proporcionaram resultados mais precisos. Três e duas safras foram suficientes para selecionar acuradamente considerando produtividade e resistência, respectivamente, quando o objetivo for recombinar. Para recomendação, nove safras fornecerão a acurácia necessária. Foram selecionados dez híbridos com bons índices de produtividade, adaptabilidade, estabilidade e resistência. Tais híbridos deverão ser incorporados ao programa de melhoramento e as matrizes destaques dentre as plantas referentes à tais híbridos serão selecionadas e empregadas em futuras hibridações e/ou ensaios de competição clonal.

Palavras-chave: *Theobroma grandiflorum*. REML. BLUP. Modelos mistos. Seleção genética.

ABSTRACT

CHAVES, Saulo Fabrício da Silva, M.Sc., Universidade Federal de Viçosa, February 2022. **Models for repeated measures analysis applied to the cupuassu tree breeding.** Advisor: Luiz Antônio dos Santos Dias. Co-advisor: Rodrigo Silva Alves.

The modernization of the productive chain of cupuassu tree (*Theobroma grandiflorum*), a fruitbearing perennial species with a long breeding cycle, depends on the adoption of efficient breeding and selection strategies. The agronomic characters of the species usually involve repeated measures, requiring the determination of the optimal number of harvests for an accurate genetic selection, as well as studying the adaptability and temporal stability of hybrids, which are fundamental for genetic selection. In this context, the use of adequate statistical methods is essential to achieve credible and useful results for the breeding program. This dissertation aimed to evaluate different models and selection strategies to determine the optimal number of harvests to be evaluated (Chapter 1) and the most adequate covariance structure (Chapter 2), in addition to the genetic selection of hybrids (both chapters). For this purpose, a population with 34 triple hybrids, delineated in randomized blocks and with a variable number of repetitions, was evaluated during nine consecutive seasons in terms of yield and resistance to witches' broom. Statistical procedures were based on mixed linear models. In all scenarios, the flexibility of the (co)variance structures of the models provided more accurate results. Three and two crops were enough to select accurately considering yield and resistance, respectively, when the objective is to recombine. For recommendation, nine harvests will provide the required accuracy. Ten hybrids with good levels of productivity, adaptability, stability and resistance were selected. Such hybrids should be incorporated into the breeding program and the outstanding matrices within those hybrids will be selected and used in future hybridizations and/or clonal competition assays.

Keywords: *Theobroma grandiflorum*. REML. BLUP. Mixed models. Genetic selection.

SUMÁRIO

1 INTRODUÇÃO GERAL	8
REFERÊNCIAS	12
AVALIAÇÃO GENÉTICA E SELEÇÃO DE GENÓTIPOS PRODUTIVOS E RESISTENTES À VASSOURA-DE-BRUXA EM <i>Theobroma grandiflorum</i>	15
RESUMO.....	16
1 INTRODUÇÃO.....	17
2 MATERIAIS E MÉTODOS.....	18
2.1 Condições experimentais	18
2.2 Genótipos avaliados	19
2.3 Análises estatísticas	19
3 RESULTADOS	23
4 DISCUSSÃO	29
5 CONCLUSÃO.....	33
REFERÊNCIAS	33
APÊNDICE A.....	38
MODELOS PARA ANÁLISE DE MEDIDAS REPETIDAS: APLICAÇÃO NA AVALIAÇÃO GENÉTICA DO CUPUAÇUZEIRO.....	40
1 INTRODUÇÃO.....	42
2 MATERIAIS E MÉTODOS.....	43
2.1 Condições experimentais	43
2.2 Genótipos avaliados	44
2.3 Análises estatísticas	44
2.3.1 Modelagem dos efeitos residuais	45
2.3.2 Modelagem dos efeitos de ambiente permanente de indivíduo	46
2.3.3 Modelagem dos efeitos de parcela	46
2.3.4 Modelagem dos efeitos genotípicos	46
2.3.3 Seleção do modelo	48
2.3.4 Estimativas dos parâmetros genéticos e não genéticos	48
3 RESULTADOS	49
4 DISCUSSÃO	55
5 CONSIDERAÇÕES FINAIS	59
REFERÊNCIAS	59

1 INTRODUÇÃO GERAL

O cupuaçuzeiro (*Theobroma grandiflorum* (Willd ex. Spreng) Schumm.) é uma espécie arbórea perene da família Malvaceae, subfamília Byttnerioideae, tribo Theobromeae (SILVA et al., 2017). Trata-se de uma espécie frutífera intimamente relacionada com a tradição e cultura amazônica que remonta a épocas pré-colombianas. Corrobora com essa afirmativa a dispersão da espécie ao longo das margens dos rios amazônicos, a partir de seu suposto centro de origem, no Sul e Sudeste Paraense e Noroeste Maranenhense, possivelmente uma ação antrópica dos primeiros habitantes da região, atraídos pelo aroma e sabor da polpa de seu fruto, o cupuaçu (ALVES et al., 2007).

Até o fim dos anos 1960, a exploração do cupuaçuzeiro era exclusivamente extrativista, com cultivos esporádicos em pequenos quintais agroflorestais. No Estado do Pará, o cultivo em maior escala teve início com o declínio da lavoura da pimenteira-do-reino (*Piper nigrum* L.) devido à fusariose (*Fusarium solani* f. sp. *Piperis* Alb.). Os pomares pioneiros foram formados com sementes de matrizes não selecionadas (SOUZA et al., 2012). Este fato proporcionou aos cultivos desuniformidade e suscetibilidade à doença vassoura-de-bruxa, causada pelo fungo *Moniliophthora perniciosa* (Stahel) Aime & Philips-Mora, principal gargalo do cultivo do cupuaçuzeiro, especialmente na região amazônica (ALVES et al., 2003).

O cupuaçu tem aproveitamento integral pela agroindústria. A casca, que detém quase metade do peso total do fruto (ALVES et al., 2020a), pode ser utilizada como composto orgânico, produção de energia, alimentação animal ou mesmo humana (SALGADO et al., 2011). A polpa, alicerce atual da cadeia produtiva da espécie, presta-se para a fabricação de sucos, doces, geleias e outros produtos alimentícios (PEREIRA; ABREU; RODRIGUES, 2018). As amêndoas, produto com maior potencial comercial, são utilizadas na indústria alimentícia para fabricação do chocolate de cupuaçu, o “cupulate” (COSTA et al., 2019); bem como, nas indústrias cosmética e farmacêutica, devido às propriedades antioxidantes do seu óleo (GONÇALVES; LAJOLO; GENOVESE, 2010).

A espécie possui as maiores flores do gênero *Theobroma*, fazendo jus ao seu epíteto *grandiflorum*. A época de floração corresponde ao período de menor umidade, que na região amazônica, estende-se de julho a setembro, enquanto a frutificação ocorre, majoritariamente, durante o período chuvoso, nos primeiros meses do ano (VENTURIERI, 2011). A espécie é alógama, com autoincompatibilidade tardia e, em alguns casos, inter-incompatibilidade (ALVES et al., 2007; RAMOS et al., 2005). Demonstra altas taxas de divergência interpopulacional o que, por um lado, influencia positivamente no ganho genético pela

exploração da heterose, e por outro, aliada à inter- e autoincompatibilidade, dificulta o processo de melhoramento em si (ALVES et al., 2003, 2007).

O programa de melhoramento genético do cupuaçuzeiro da Embrapa teve início na década de 1980, com as primeiras prospecções realizadas ao longo da calha do rio Amazonas nos estados do Pará, Amazonas e Amapá. Tais expedições foram responsáveis pelas coletas dos acessos que formaram os bancos ativos de germoplasma (BAGs) pioneiros na Embrapa Amazônia Ocidental (253 acessos), Embrapa Amazônia Oriental (198 acessos) e no INPA (79 acessos). Posteriormente, formaram-se mais três BAGs: no Amapá (68 acessos), em Rondônia (36 acessos) e no Acre (12 acessos), todos compostos por acessos coletados nos respectivos estados (SOUZA et al., 2002). Após a formação das populações-base, o programa de melhoramento tem sido conduzido em três fases: a caracterização e avaliação dos acessos nas coleções de trabalho, a seleção de acessos superiores e a utilização destes genótipos para formação de populações melhoradas e desenvolvimento de cultivares, seminais ou clonais, visto a versatilidade propagativa da espécie (ALVES et al., 2020a; SOUZA et al., 2012).

Nas primeiras etapas do programa de melhoramento, quando é empregado um número grande de indivíduos, é comum a utilização da seleção massal (SOUZA et al., 2002). Com o afinamento que ocorre ao longo dos estádios de melhoramento, são priorizados métodos mais robustos, utilizando hibridações e realizando seleções com testes entre e dentro de progênies (SOUZA et al., 2012). Os ciclos de seleção compõem a seleção recorrente intrapopulacional, baseada na recombinação dos indivíduos selecionados para manutenção da variabilidade genética (SOUZA et al., 2002). Priorizam-se matrizes com altos índices produtivos (alta produção de frutos pesados, maior proporção de polpa e sementes, etc.) e com resistência às pragas e doenças, especialmente à vassoura-de-bruxa.

Os entraves no melhoramento do cupuaçuzeiro são inerentes às próprias características da espécie, que apresenta longo ciclo reprodutivo, com elevado período juvenil (três primeiros anos de campo) e necessidade de avaliações repetidas no tempo, por apresentar expressão fenotípica diferencial ao longo dos anos. Assim, cada ciclo de melhoramento leva em torno de 12 anos (ALVES et al., 2021; SOUZA et al., 2002, 2012).

Portanto, é patente a necessidade da determinação do número ótimo de safras a serem avaliadas, para a seleção genética que vise aumentar os ganhos com a seleção por unidade de tempo. Neste contexto, utilizando medidas repetidas, é possível estimar parâmetros úteis, tais como os coeficientes de herdabilidade e repetibilidade, bem como prever os efeitos da interação genótipos x medidas (GM), sendo este último utilizado para análises de adaptabilidade e estabilidade temporal (FERREIRA et al., 2020, 2021). No entanto, é comum

encontrar na literatura estudos que assumem certas pressuposições que nem sempre são atendidas, como por exemplo a homogeneidade da variância residual entre as safras; ou desconsideram as covariâncias entre as safras em relação aos efeitos genéticos e não-genéticos (SANCHÉZ et al., 2017; SILVA et al., 2020). Métodos estatísticos inadequados podem levar a resultados errôneos, os quais podem prejudicar a eficiência dos programas de melhoramento, sobretudo aqueles com ciclos de longa duração, como o do cupuaçuzeiro.

Para contornar este risco, inserem-se métodos que, inicialmente, foram desenvolvidos para a avaliação genética animal, mas que podem – e são – aplicados ao melhoramento vegetal, como a Melhor Predição Linear Não-Viesada (BLUP) (HENDERSON, 1975); e métodos concebidos para análise de experimentos desbalanceados, como a Máxima Verossimilhança Restrita (REML) (PATTERSON; THOMPSON, 1971).

Considerando o seguinte modelo linear misto:

$$y = X\beta + Zu + e,$$

onde y é o vetor de dados fenotípicos, β é o vetor dos efeitos fixos, u é o vetor dos efeitos aleatórios, e é o vetor dos resíduos e X e Z são as matrizes de incidência para β e u , respectivamente; a obtenção dos BLUP e BLUE (Melhor Estimção Linear não Viesada, para efeitos fixos) perpassa pela Equação de Modelos Mistos (EMM) de Henderson (1975):

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + A^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix},$$

onde R é a matriz de (co)variâncias residuais e A é a matriz de parentesco. Considerando apenas uma característica avaliada em apenas um ambiente ou ano, $R = \sigma_e^2$, onde σ_e^2 é a variância residual. Desta forma, a equação de modelos mistos pode ser simplificada (ISIK; HOLLAND; MALTECCA, 2017; RESENDE, 2002):

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\lambda \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix},$$

Onde $\lambda = \sigma_e^2/\sigma_u^2$, sendo σ_u^2 a variância do efeito genético.

A utilização do BLUP, permite a minimização da variância do erro de predição (PEV) e, conseqüentemente, a maximização da acurácia seletiva; considera, quando disponível, a informação de parentesco entre os indivíduos avaliados, além de outras fontes de informação tais como várias safras, situação recorrente no melhoramento de fruteiras, maximizando o ganho genético (ISIK; HOLLAND; MALTECCA, 2017; RESENDE, 2004, 2016).

O BLUP genuíno assume que os componentes de variância são conhecidos (ALVES et al., 2018). Porém, na prática estes componentes são desconhecidos e devem ser estimados para a obtenção do BLUP empírico (RESENDE; SILVA; AZEVEDO, 2014). Para isto, vários

métodos foram desenvolvidos (Mínimos Quadrados; Métodos I, II e III de Henderson; Máxima Verossimilhança (ML); Máxima Verossimilhança Restrita (REML); entre outros).

O método REML é dito “restrito”, pois maximiza somente a porção da verossimilhança livre dos efeitos fixos, anulando, desta forma, o viés do método ML, porém mantendo suas vantagens diferenciais, isto é, suficiência, consistência, eficiência e invariância à parametrização (CORBEIL; SEARLE, 1976; MYUNG, 2003; RESENDE, 2002; RESENDE; SILVA; AZEVEDO, 2014). A partição dos dados em “contraste dos efeitos fixos” e “contraste dos erros”, isto é, aqueles efeitos com esperança igual à zero, é quem permite a particularização da maximização e, portanto, a estimação dos componentes de variância não-viesados (PATTERSON; THOMPSON, 1971; RESENDE; ALVES, 2020; RESENDE; SILVA; AZEVEDO, 2014). O método REML é composto de equações não lineares, com matrizes de grandes dimensões, o que demanda uma solução baseada em um processo iterativo até a convergência. As equações de modelos mistos podem auxiliar neste sentido. Em suma, no BLUP, assume-se o conhecimento dos componentes de variância, e estes podem ser estimados de maneira não-viesada via REML, que, por sua vez, vale-se das EMM para convergência dos componentes de variância e/ou parâmetros genéticos. Portanto, ambos procedimentos são realizados simultaneamente (ALVES et al., 2018; RESENDE, 2002).

Em dados balanceados e experimentos ortogonais, o método assemelha-se em eficiência ao método dos mínimos quadrados (LITTELL, 2002; RESENDE, 2002). Quando estas pressuposições não são atendidas, o método REML/BLUP é o mais adequado. A metodologia destaca-se, ainda, por gerar estimativas acuradas e não-viesadas quando os dados apresentam maior complexidade, como em ensaios em múltiplos ambientes. Ademais, permite a escolha de modelos com base em critérios estatísticos e a incorporação da informação de parentesco (ALVES et al., 2019; BERNARDO, 2020; RESENDE, 2016). Estas e outras vantagens, contribuem para a popularização do método e fazem com que ele seja cada vez mais utilizado nos programas de melhoramento (RESENDE et al., 2014).

No cupuaçuzeiro, Souza et al. (2002) indicaram a metodologia de modelos mistos (REML/BLUP) como a mais adequada para a avaliação genética. Desde então, esta metodologia vem auxiliando de maneira significativa na obtenção de resultados e maximização dos ganhos, como pode ser observado em Alves e Resende (2008), Maia et al. (2011), Alves et al. (2020b) e Alves et al. (2021).

Portanto, este estudo visa a avaliação de modelos lineares mistos com diferentes estruturas de (co)variância para avaliação de medidas repetidas e, também, a avaliação de

diferentes estratégias de seleção de modo a maximizar a eficiência do programa de melhoramento do cupuaçuzeiro.

REFERÊNCIAS

- ALVES, R. M.; ARTERO, A. S.; SEBBENN, A. M. Mating system in a natural population of *Theobroma grandiflorum* (Willd. ex Spreng.) Schum., by microsatellite markers. **Genetics and Molecular Biology**, v. 26, p. 373–379, 2003.
- ALVES, R. M.; SEBBENN, A. M.; ARTERO, A. S.; CLEMENT, C.; FIGUEIRA, A. High levels of genetic divergence and inbreeding in populations of cupuassu (*Theobroma grandiflorum*). **Tree Genetics & Genomes**, v. 3, n. 4, p. 289–298, 2007.
- ALVES, R. M.; CHAVES, S. F. S.; OLIVEIRA, R. P.; PEDROZA NETO, J. L.; SEBBENN, A. M. Canopy replacement used in the evaluation of cupuassu tree genotypes in the state of Pará. **Revista Brasileira de Fruticultura**, v. 42, n. 4, p. e-597, 2020a.
- ALVES, R. M.; CHAVES, S. F. S.; GAMA, M. A. P.; PEDROZA NETO, J. L.; SANTOS, T. G. Simultaneous selection of cupuassu tree and Brazilian mahogany genotypes in an agroforestry system in Pará state, Brazil. **Acta Amazonica**, v. 50, n. 3, p. 183–191, 2020b.
- ALVES, R. M.; CHAVES, S. F. S.; ALVES, R. S.; SANTOS, T. G.; ARAÚJO, D. G.; RESENDE, M. D. V. Cupuaçu tree genotype selection for an agroforestry system environment in the Amazon. **Pesquisa Agropecuária Brasileira**, v. 56, p. e02139, 2021.
- ALVES, R. M.; RESENDE, M. D. V. Avaliação genética de indivíduos e progênes de cupuaçuzeiro no estado do Pará e estimativas de parâmetros genéticos. **Revista Brasileira de Fruticultura**, v. 30, n. 3, p. 696–701, 2008.
- ALVES, R. S.; PEIXOTO, L. A.; TEODORO, P. E.; SILVA, L. A.; RODRIGUES, E. V.; RESENDE, M. D. V.; LAVIOLA, B. G.; BHERING, L. L. Selection of *Jatropha curcas* families based on temporal stability and adaptability of genetic values. **Industrial Crops and Products**, v. 119, p. 290–293, 2018.
- ALVES, R. S.; PEIXOTO, L. A.; TEODORO, P. E.; SILVA, L. A.; RODRIGUES, E. V.; RESENDE, M. D. V.; LAVIOLA, B. G.; BHERING, L. L. Multiple-trait BLUP in longitudinal data analysis on *Jatropha curcas* breeding for bioenergy. **Industrial Crops and Products**, v. 130, p. 558–561, 2019.
- BERNARDO, R. Reinventing quantitative genetics for plant breeding: something old, something new, something borrowed, something BLUE. **Heredity**, v. 125, n. 6, p. 375–385, 2020.
- CORBEIL, R. R.; SEARLE, S. R. Restricted Maximum Likelihood (REML) Estimation of Variance Components in the Mixed Model. **Technometrics**, v. 18, n. 1, p. 31, 1976.
- COSTA, R. S.; SANTOS, O. V.; LANNES, S. C. S.; CASAZZA, A. A.; ALIAKBARIAN,

- B.; PEREGO, P.; RIBEIRO-COSTA, R. M.; CONVERTI, A.; SILVA, J. O. C. Bioactive compounds and value-added applications of cupuassu (*Theobroma grandiflorum* Schum.) agroindustrial by-product. **Food Science and Technology**, v. 40, p. 401–407, 2019.
- FERREIRA, F. M.; ROCHA, J. R. A. S. C.; ALVES, R. S.; ELIZEU, A. M.; BENITES, F. R. G.; RESENDE, M. D. V.; SOUZA SOBRINHO, F.; BHERING, L. L. Estimates of repeatability coefficients and optimum number of measures for genetic selection of *Cynodon spp.* **Euphytica**, v. 216, n. 5, p. 70, 2020.
- FERREIRA, F. M.; ROCHA, J. R. A. S. C.; BHERING, L. L.; FERNANDES, F. D.; LÉDO, F. J. S.; RANGEL, J. H. A.; KOPP, M.; CÂMARA, T. M.; SILVA, V. Q. R.; MACHADO, J. C. Optimal harvest number and genotypic evaluation of total dry biomass, stability, and adaptability of elephant grass clones for bioenergy purposes. **Biomass and Bioenergy**, v. 149, p. 106104, 2021.
- GONÇALVES, A. E. S. S.; LAJOLO, F. M.; GENOVESE, M. I. Chemical Composition and Antioxidant/Antidiabetic Potential of Brazilian Native Fruits and Commercial Frozen Pulps. **Journal of Agricultural and Food Chemistry**, v. 58, n. 8, p. 4666–4674, 2010.
- HENDERSON, C. R. Best Linear Unbiased Estimation and Prediction under a Selection Model. **Biometrics**, v. 31, n. 2, p. 423, 1975.
- ISIK, F.; HOLLAND, J.; MALTECCA, C. **Genetic Data Analysis for Plant and Animal Breeding**. Cham: Springer International Publishing, 2017.
- LITTELL, R. C. Analysis of unbalanced mixed model data: A case study comparison of ANOVA versus REML/GLS. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 7, n. 4, p. 472, 2002.
- MAIA, M. C. C.; RESENDE, M. D. V.; OLIVEIRA, L. C.; ÁLVARES, V. S.; MACIEL, V. T.; LIMA, A. C. Seleção de clones experimentais de cupuaçu para características agroindustriais via modelos mistos. **Revista Agro@ambiente on-line**, v. 5, n. 1, p. 35, 2011.
- MYUNG, I. J. Tutorial on maximum likelihood estimation. **Journal of Mathematical Psychology**, v. 47, n. 1, p. 90–100, 2003.
- 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, A. L. F.; ABREU, V. K. G.; RODRIGUES, S. Cupuassu—*Theobroma grandiflorum*. In: RODRIGUES, S.; SILVA, E. DE O.; BRITO, E. S. DE (Eds.). **Exotic Fruits**. [s.l.] Academic Press, 2018. p. 159–162.
- RAMOS, A. R. VENTURIERI, G. A.; CUCO, S. M.; CASTRO, N. M. The site of self-incompatibility action in cupuassu (*Theobroma grandiflorum*). **Brazilian Journal of Botany**, v. 28, p. 569–578, 2005.
- RESENDE, M. D. V. **Genética biométrica e Estatística no melhoramento de plantas perenes**. 1. ed. Brasília: Embrapa Informação Tecnológica, 2002. 975p.
- RESENDE, M. D. V. **Métodos estatísticos ótimos para análise de experimentos de campo**,

2004. Disponível em: <<https://www.infoteca.cnptia.embrapa.br/handle/doc/305549>>
- RESENDE, M. D. V. Software Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology**, v. 16, n. 4, p. 330–339, 2016.
- RESENDE, M. D. V.; ALVES, R. S. Linear, generalized, hierarchical, bayesian and random regression mixed models in genetic/genomics in plant breeding. **Functional Plant Breeding Journal**, v. 2, n. 2, p. 1–31, 2020.
- RESENDE, M. D. V.; SILVA, F. F.; AZEVEDO, C. F. **Estatística matemática, biométrica e computacional: modelos mistos, multivariados, categóricos e generalizados (REML/BLUP), inferência bayesiana, regressão, aleatória, seleção genômica, QTL, GWAS, estatística espacial e temporal, competição, sobrevivência**. Viçosa, MG, Brazil: UFV, 2014. 881p.
- SALGADO, J. M.; RODRIGUES, B. S.; DONADO-PESTANA, C. M.; DIAS, C. T. S.; MORZELLE, M. C. Cupuassu (*Theobroma grandiflorum*) peel as potential source of dietary fiber and phytochemicals in whole-bread preparations. **Plant Foods for Human Nutrition**, v. 66, n. 4, p. 384–390, 2011.
- SANCHÉZ, C. F. B.; ALVES, R. S.; GARCIA, A. P.; TEODORO, P. E.; PEIXOTO, L. A.; SILVA, L. A.; BHERING, L. L.; RESENDE, M. D. V. Estimates of repeatability coefficients and the number of the optimum measure to select superior genotypes in *Annona muricata* L. **Genetics and Molecular Research**, v. 16, n. 3, 2017.
- SILVA, J. O. C.; BRUCKNER, C. H.; CARNEIRO, P. C. S.; RESENDE, M. D. V.; ALVES, R. S.; MATIAS, R. G. P. Estimates of genetic parameters and repeatability coefficients in S2 peach genotypes grown in a subtropical environment. **Functional Plant Breeding Journal**, v. 2, n. 1, p. 25–33, 2020.
- SILVA, R. A.; SOUZA, G.; LEMOS, L. S. L.; LOPES, U. V.; PATROCÍNIO, N. G. R.; ALVES, R. M.; MARCELLINO, L. H.; CLEMENT, D.; MICHELI, F.; GRAMACHO, K. P. Genome size, cytogenetic data and transferability of EST-SSRs markers in wild and cultivated species of the genus *Theobroma* L. (Byttnerioideae, Malvaceae). **PLOS ONE**, v. 12, n. 2, p. e0170799, 2017.
- SOUZA, A. G. C.; RESENDE, M. D. V.; SILVA, S. E. L.; SOUSA, N. R. The cupuaçu genetic improvement program at Embrapa Amazônia Ocidental. **Crop Breeding and Applied Biotechnology**, v. 2, n. 3, p. 471–478, 2002.
- SOUZA, A. G. C.; ALVES, R. M.; SOUSA, N. R.; SOUZA, M. G. Domestication and breeding of the cupuassu tree. In: BORÉM, A.; LOPES, M. T. G.; CLEMENT, C. R.; NODA, H. (Eds.). **Domestication and breeding: Amazonian species**. 1. ed. Viçosa, MG, Brazil: Suprema, 2012. p. 319–331.
- VENTURIERI, G. A. Flowering levels, harvest season and yields of cupuassu (*Theobroma grandiflorum*). **Acta Amazonica**, v. 41, p. 143–152, 2011.

CAPÍTULO 1

**AVALIAÇÃO GENÉTICA E SELEÇÃO DE GENÓTIPOS PRODUTIVOS E
RESISTENTES À VASSOURA-DE-BRUXA EM *Theobroma grandiflorum***

VIÇOSA - MINAS GERAIS

2022

RESUMO

O cupuaçuzeiro (*Theobroma grandiflorum*) é uma espécie perene frutífera, com ciclo de melhoramento longo, em média 12 anos. Daí a importância da determinação do número ótimo de safras para uma seleção genética acurada, de modo a reduzir esse ciclo. Na análise de dados, frequentemente assumem-se pressuposições nem sempre atendidas, como a homocedasticidade e/ou realizam-se ajustes para superar o desbalanceamento dos dados, especialmente em plantas perenes. O objetivo deste estudo foi testar modelos de repetibilidade com variâncias residuais homogêneas e heterogêneas, definir o número ótimo de safras para uma seleção genética acurada, avaliar a estabilidade e adaptabilidade temporal de híbridos de cupuaçuzeiro e selecionar, de forma simultânea, para produção de frutos e resistência à vassoura-de-bruxa por meio de modelos mistos. Uma população com 34 híbridos triplos de cupuaçuzeiro, delineada em blocos casualizados e com número variável de repetições, foi avaliada durante nove safras quanto a produtividade e resistência à vassoura-de-bruxa. As estimativas dos parâmetros genéticos e as previsões dos valores genotípicos foram obtidas via metodologia de modelo misto (REML/BLUP). Para resistência à vassoura-de-bruxa, característica binomial, ajustou-se um modelo linear generalizado misto. O modelo com resíduos heterogêneos apresentou o melhor ajuste segundo o critério de informação de Akaike, e os resultados gerados a partir dele atestaram variabilidade genética entre híbridos. Ademais, três e duas safras são suficientes para selecionar acuradamente para produtividade e resistência, respectivamente, quando o objetivo é recombinar, e de nove safras quando o objetivo for a recomendação. Foram selecionados dez híbridos com bons índices de produtividade, adaptabilidade, estabilidade e resistência. Tais híbridos deverão ser incorporados ao programa de melhoramento e as matrizes destaques dentre as plantas referentes à tais híbridos serão priorizados para futuras hibridações ou ensaios de competição clonal.

Palavras-chave: *Theobroma grandiflorum*, REML/BLUP, repetibilidade, seleção genética, heterocedasticidade.

1 INTRODUÇÃO

A Amazônia brasileira, rica em biodiversidade, detém ampla variedade de espécies frutíferas, cujo potencial extrapola as suas fronteiras, despertando interesse do mercado internacional. Estima-se que 44% da diversidade de fruteiras nativas do Brasil estejam na Amazônia (NEVES et al., 2012), indicando o enorme potencial para exploração científica e comercial sustentável dessas espécies.

O cupuaçuzeiro (*Theobroma grandiflorum* Willd. (ex. Spreng) Schum.) é um exemplo de fruteira nativa que vem ganhando destaque. Desde meados de 1980, a espécie é alvo do programa de melhoramento da Embrapa, que visa obter genótipos produtivos e resistentes ao fungo *Moniliophthora perniciosa* Stahel & Phillips-Mora, agente etiológico da vassoura-de-bruxa, de grande importância no gênero *Theobroma* (ALVES et al., 2020; PATROCÍNIO et al., 2017).

Por ser uma espécie perene com longo período juvenil (cerca de três anos), o ciclo de melhoramento do cupuaçuzeiro é extenso (em média, 12 anos), o que dificulta o lançamento de novas cultivares (SOUZA et al., 2002). Pela sua alogamia auto-incompatível (RAMOS et al., 2005), a seleção recorrente intrapopulacional é o método de melhoramento mais utilizado, o qual demanda vários ciclos de avaliação, seleção e recombinação para o acúmulo de alelos favoráveis. A clonagem dos melhores indivíduos a cada ciclo é realizada com base no tamanho efetivo populacional (N_e) (RESENDE, 2002), o qual, se muito pequeno ($N_e < 5$), implica diminuição da variabilidade genética no pomar, que, aliado à autoincompatibilidade da espécie, pode comprometer a produtividade.

Portanto, há necessidade da determinação do número ótimo de safras para a seleção genética, de modo a reduzir a duração do ciclo de melhoramento e, por conseguinte, aumentar os ganhos com a seleção por unidade de tempo (CHAVES et al., 2021). O uso de medidas repetidas possibilita estimar o coeficiente de repetibilidade, além dos efeitos da interação genótipos x medições (GM). A quantificação desta interação é fundamental para o entendimento da variação temporal no desempenho dos genótipos (ATLIN; CAIRNS; DAS, 2017). Com os efeitos da interação GM é possível empregar métodos para capitalizar a adaptabilidade e a estabilidade temporal dos genótipos em avaliação, conferindo maior credibilidade à seleção (DIAS et al., 1998).

O uso de medidas repetidas permite, ainda, identificar a estrutura de variância residual ao longo das safras, o que tem sido pouco estudado (PEIXOTO et al., 2021). Deve-se considerar a eventual heterocedasticidade residual, haja vista a precisão e acurácia demandada

na seleção (FAVERI et al., 2015; VERBYLA et al., 2021). É fundamental a avaliação de diferentes estruturas de variância residual, de modo a selecionar o melhor modelo para a avaliação genética (PEREIRA et al., 2018; QUINTAL et al., 2017).

Neste sentido, a metodologia de modelos mistos (REML/BLUP) tende a ser mais adequada, por sua maior flexibilidade e robustez. Essa metodologia é tão efetiva quanto o método da ANOVA (quadrados mínimos), em condições em que todas as pressuposições são atendidas. Porém, supera este último método quando não há independência, homogeneidade residual e/ou balanceamento genético e/ou experimental dos dados. Ademais, permite a escolha de modelos com base em critérios estatísticos e a incorporação da informação de parentesco (BERNARDO, 2020; RESENDE, 2002; RESENDE; ALVES, 2020).

Este estudo objetivou: i) avaliar modelos de repetibilidade com variâncias residuais homogêneas e heterogêneas, ii) definir o número ótimo de medidas (safras) para seleção genética acurada, iii) avaliar a estabilidade e adaptabilidade temporal de híbridos triplos de cupuaçuzeiro e iv) selecionar híbridos com base na produtividade de frutos e resistência à vassoura-de-bruxa, simultaneamente.

2 MATERIAIS E MÉTODOS

2.1 Condições experimentais

O experimento foi instalado em uma propriedade rural no município de Tomé-Açu (latitude 02° 24' 05,6" S; longitude 48° 00' 08,8" W; altitude 45 m acima do nível do mar), Pará, em fevereiro de 2007. O clima da região corresponde ao tipo Ami, de acordo com a classificação de Köppen, e ao tipo B4rA'a', seguindo a classificação de Thornthwaite (MORAES et al., 2020). O solo é do tipo Latossolo Amarelo, textura média, com baixa fertilidade, profundo e bem drenado (EMBRAPA, 2016). A temperatura média, no período avaliado, foi de aproximadamente 27 °C, com umidade relativa do ar de 85,5% e precipitação pluviométrica média anual de 2.729 mm, de acordo com os dados fornecidos pela estação meteorológica da Embrapa Amazônia Oriental, instalada no município.

Por necessitar de sombreamento parcial, principalmente nos primeiros anos de cultivo, o cupuaçuzeiro foi implantado em consórcio com um plantio comercial de pimenteira do reino (*Piper nigrum* L.) que estava no quinto ano de campo, em espaçamento de fileiras duplas [(2 x

2) x 4 m]. As plantas de cupuaçuzeiro foram estabelecidas nas entrelinhas, no espaçamento de 6 x 4 m. Os tratos culturais seguiram as recomendações técnicas dos sistemas de produção de ambas as espécies (LEMOS; TREMACOLDI; POLTRONIERI, 2014; SOUZA, 2007).

2.2 Genótipos avaliados

As progênies avaliadas são oriundas de cruzamentos triplos: inicialmente, foram obtidos híbridos simples após cruzamentos entre pares de clones avaliados e selecionados no BAG da Embrapa Amazônia Oriental, os quais apresentavam características complementares (ALVES et al., 2013). Estes híbridos foram avaliados em campo durante 12 anos. Após a seleção entre e dentro das famílias, plantas matrizes foram selecionadas quanto à produtividade e resistência à vassoura-de-bruxa, sendo, posteriormente, clonadas e cruzadas com quatro cultivares: BRS Coari (174), BRS Codajás (186), BRS Manacapuru (215) e BRS Belém (286). A variabilidade entre os genótipos resultantes dos cruzamentos triplos é fruto da heterogeneidade dos parentais não-endogâmicos, visto que, na hibridação tripla, cruzaram-se matrizes selecionadas com um clone (C) diferente de um dos seus parentais, obedecendo o esquema $(C_1 \times C_2) \times C_3$, com contribuições gênicas de três clones distintos. Foram avaliados, ao todo, 34 híbridos triplos nãoconvencionais de cupuaçuzeiro, dispostos em blocos completos casualizados com cinco a dez repetições e três plantas por parcela (Apêndice A). Os híbridos triplos são chamados “nãoconvencionais” por serem frutos de cruzamentos entre indivíduos heterozigotos. O número de repetições variou conforme à disponibilidade de mudas para plantio.

A partir do terceiro ano de campo foram avaliadas as seguintes características: produtividade média de frutos por planta (kg), produto do número de frutos pelo peso médio dos frutos por planta; e a presença de sintomas da doença vassoura-de-bruxa nos ramos. A infecção da doença ocorreu de forma natural, afim de que os genótipos pudessem manifestar os mecanismos de resistência horizontal. Estas avaliações foram realizadas por nove safras consecutivas.

2.3 Análises estatísticas

Em virtude do desbalanceamento, causado tanto pelo número diferente de repetições, quanto pela morte de algumas plantas, a metodologia de modelos mistos (REML/BLUP) foi

utilizada para a estimação dos componentes de variância e parâmetros genéticos, bem como para a predição dos valores genotípicos (HENDERSON, 1975; PATTERSON; THOMPSON, 1971).

O modelo linear misto associado à avaliação de híbridos, em delineamento de blocos completos casualizados, com várias plantas por parcela e medidas repetidas para produtividade, foi dado por:

$$y = X_m + Z_g + W_p + T_{gm} + Q_i + e,$$

onde y é o vetor de dados fenotípicos; m é o vetor dos efeitos de medições, repetições e suas combinações (assumidos como fixos), somados à média geral; g é o vetor dos efeitos genotípicos, assumidos como aleatórios ($g \sim N [0, \sigma_g^2]$, onde σ_g^2 é a variância genotípica); p é o vetor dos efeitos de aleatórios de parcelas ($p \sim N [0, \sigma_p^2]$, onde σ_p^2 é a variância dos efeitos de parcela); gm é o vetor dos efeitos aleatórios da interação GM ($gm \sim N [0, \sigma_{gm}^2]$, onde σ_{gm}^2 é a variância dos efeitos da interação GM); i é o vetor dos efeitos aleatórios de ambiente permanente ($i \sim N [0, \sigma_i^2]$, onde σ_i^2 é a variância dos efeitos de ambiente permanente); e é o vetor dos resíduos (aleatórios) ($e \sim N [0, \sigma_e^2]$, onde σ_e^2 é a variância residual, para o modelo homocedástico; $e \sim N [0, R]$, onde R é a matriz de variância residual, para o modelo heterocedástico). X , Z , W , T e Q são as matrizes de incidência para m , g , p , gm e i , respectivamente.

A atribuição da natureza aleatória aos efeitos genéticos tem como base os trabalhos de Piepho et al. (2008) e Resende e Duarte (2007), os quais defendem que quando o número de tratamentos (genótipos) é maior que 10, a utilização de um estimador de natureza aleatória, tal qual o BLUP, é preferível. De fato, os efeitos genéticos só devem ser considerados de natureza fixa quando não for possível estimar componentes de variâncias a partir dos dados, ou quando o principal objetivo for a comparação estatística entre um número reduzidos de cultivares (PIEPHO et al., 2008).

Para a característica “presença de sintomas de vassoura-de-bruxa” - avaliada por 0 (ausência) e 1 (presença) - foi utilizado um modelo linear generalizado misto (MLGM), com os mesmos efeitos aleatórios do modelo ajustado à produtividade e considerando que os resíduos seguem distribuição binomial. A linearização dos dados foi realizada utilizando a função de ligação logarítmica [$\text{logit} = \mu / (1 - \mu)$, onde μ é a média na escala dos dados] útil no contexto binomial devido à facilidade de implementação e interpretação (VILLEMEREUIL, 2018). O software utilizado para a realização dessa análise (ASReml, BUTLER et al., 2018) usa a metodologia da quase-verossimilhança penalizada (PQL,

WOLFINGER; CLAYTON, 1993), e maximiza a função de verossimilhança fixando a variância residual em um. Posteriormente, este valor foi ajustado para $\sigma_e^2 = \pi^2/3 = 3,29$, para comportar a variância da distribuição logística (VILLEMEREUIL, 2018). A atribuição de uma constante (1) para a variância residual impede o ajuste de um MLGM de repetibilidade com resíduos heterogêneos.

As diferentes estruturas de variância residual (homogênea e heterogênea) foram testadas pelo critério de informação de Akaike (AIC) (AKAIKE, 1974), dado por:

$$AIC = -2\text{Log}L + 2p,$$

onde $\text{Log}L$ é o logaritmo do máximo da função de verossimilhança restrita e p é o número de parâmetros estimados.

A significância dos efeitos aleatórios do modelo estatístico foi testada utilizando os intervalos de confiança (IC) (GILMOUR et al., 2015):

$$ICx = \sigma_x^2 \pm (1,96 EPx),$$

onde σ_x^2 é a estimativa da variância do efeito x , EPx é o erro padrão da estimativa da variância do efeito x e 1,96 é o valor tabelado do teste t de Student ao nível de 5% de probabilidade de erro Tipo I, para infinitos graus de liberdade.

A partir das estimativas dos componentes de variância, os seguintes parâmetros foram estimados:

1- Variância fenotípica individual (σ_f^2):

$$\sigma_f^2 = \sigma_g^2 + \sigma_{gm}^2 + \sigma_p^2 + \sigma_i^2 + \sigma_{em}^2, \text{ onde } \sigma_{em}^2 \text{ é a variância residual da safra } m;$$

2- Herdabilidade média no sentido amplo (H_g^2):

$$H_g^2 = 1 - \frac{\overline{\Delta}_g}{2\sigma_g^2}, \text{ em que } \overline{\Delta}_g \text{ é a variância do erro de predição média entre genótipos par a par}$$

(CULLIS; SMITH; COOMBES, 2006);

3- Coeficiente de repetibilidade (ρ):

$$\rho = \frac{\sigma_g^2 + \sigma_i^2}{\sigma_f^2};$$

4- Coeficiente de determinação dos efeitos de parcela (c_p^2)

$$c_p^2 = \frac{\sigma_p^2}{\sigma_f^2};$$

5- Coeficiente de determinação dos efeitos da interação GM (c_{gm}^2)

$$c_{gm}^2 = \frac{\sigma_{gm}^2}{\sigma_f^2};$$

6- Coeficiente de determinação dos efeitos de ambiente permanente (c_i^2):

$$c_i^2 = \frac{\sigma_i^2}{\sigma_f^2};$$

7- Coeficiente de determinação dos efeitos residuais (c_e^2)

$$c_e^2 = \frac{\sigma_e^2}{\sigma_f^2};$$

8- Correlação genotípica através de medições (r_{gm}):

$$r_{gm} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{gm}^2};$$

9- Acurácia de seleção (r):

$$r = \sqrt{1 - \left(\frac{PEV}{\sigma_g^2}\right)}, \text{ em que PEV é a variância do erro de predição.}$$

A eficiência (E) e a acurácia (r_m) do uso de m medidas em comparação com o uso de uma única medida foram obtidas, respectivamente, por meio das seguintes equações (RESENDE; ALVES, 2020):

$$E = \left\{ \frac{m}{[1+(m-1)\rho]} \right\};$$

e

$$r_m = \sqrt{(m\rho)/(m\rho + 1 - \rho)}.$$

Visando a seleção de híbridos com maior resiliência às oscilações ambientais que ocorrem ano-a-ano foram avaliadas a adaptabilidade e a estabilidade temporais, associadas à produtividade, como proposto por Resende (2004). A estabilidade foi estimada pela Média Harmônica dos Valores Genotípicos (MHVG), a qual considera os valores genotípicos penalizados pela instabilidade, obtida por:

$$MHVG = \frac{m}{\sum_{j=1}^m \frac{1}{VG_{ij}}},$$

onde m é o número de medições e VG_{ij} é o valor genotípico da progênie i na safra j .

A adaptabilidade foi estimada pela Performance Relativa dos Valores Genotípicos (PRVG), que capitaliza a capacidade de resposta do genótipo às eventuais condições favoráveis, obtida por:

$$PRVG = \frac{1}{m} \frac{\sum_{j=1}^m VG_{ij}}{\mu_j},$$

onde μ_j é a média da safra j .

O método que compila ambos (MHVG e PRVG) e permite a seleção com base na adaptabilidade, estabilidade e produtividade é a Média Harmônica da Performance Relativa dos Valores Genotípicos (MHPRVG), obtida por:

$$\text{MHPRVG} = \frac{m}{\sum_{j=1}^m \frac{1}{\sqrt{G_{ij}/\mu_j}}}$$

O índice aditivo (IA) foi utilizado para a seleção simultânea quanto a produtividade média de frutos e resistência à vassoura-de-bruxa. Este índice foi obtido por (RESENDE, 2007):

$$\text{IA} = \sum_{c=1}^2 \text{CV}_{g_c} \frac{G}{\sigma_{g_c}},$$

onde G representa a MHPRVG da produtividade média de frutos e o valor genotípico para resistência à vassoura-de-bruxa, σ_{g_c} é o desvio padrão genotípico para a característica c e CV_{g_c} é o coeficiente de variação genotípico da característica c , utilizado como peso (FERREIRA et al., 2020), obtido por $\text{CV}_{g_c} = (100\sigma_{g_c})/\mu$, onde μ é a média geral. A direção da seleção foi positiva para MHPRVG da produtividade média de frutos e negativa para presença de sintomas de vassoura-de-bruxa. Todas as análises foram realizadas no software ASReml-R (BUTLER et al., 2018).

3 RESULTADOS

De acordo com o AIC, o modelo heterocedástico (2) apresentou o melhor ajuste (Tabela 1). Assim, todos os resultados foram obtidos a partir deste modelo.

Tabela 1. Modelos homocedástico (1) e heterocedástico (2), número de parâmetros, acurácia seletiva e critério de informação de Akaike (AIC) para a característica produtividade média de frutos (kg/planta) em híbridos triplos de cupuaçuzeiro.

Modelo	Número de parâmetros	Acurácia seletiva	AIC
1	5	0,791	39847,85
2	13	0,823	39074,07

Ao considerar o modelo heterocedástico para a característica produtividade média de frutos, as variâncias residuais e fenotípicas foram particularizadas, bem como os parâmetros genéticos derivados destes componentes, para cada safra (Tabela 2).

Tabela 2. Componentes de variância e parâmetros genéticos e não genéticos estimados a partir do modelo heterocedástico para produtividade média de frutos (kg/planta) e para presença de vassoura-de-bruxa, avaliadas durante nove safras, em híbridos triplos de cupuaçuzeiro.

Característica	Safras	Componentes de variância [†]						H_g^2	Coeficiente de determinação				ρ	r_{gm}
		σ_g^2	σ_{gm}^2	σ_i^2	σ_p^2	σ_e^2	σ_f^2		c_{gm}^2	c_i^2	c_p^2	c_e^2		
Produtividade média de frutos (kg/planta)	1 ^a					70,50	153,98	0,57	0,07	0,32	0,07	0,61	0,41	0,56
	2 ^a					64,42	147,90	0,57	0,07	0,33	0,07	0,46	0,42	
	3 ^a					80,11	163,59	0,53	0,06	0,30	0,07	0,44	0,38	
	4 ^a					129,21	212,69	0,45	0,05	0,23	0,05	0,49	0,29	
	5 ^a	13,39*	10,35*	10,83*	48,91*	142,69	226,17	0,43	0,05	0,22	0,05	0,61	0,28	
	6 ^a	± 1,53	± 0,68	± 1,11	± 1,64	263,89	347,37	0,30	0,03	0,14	0,03	0,63	0,18	
	7 ^a					384,56	467,99	0,23	0,02	0,11	0,02	0,76	0,13	
	8 ^a					203,60	287,08	0,35	0,04	0,17	0,04	0,82	0,22	
	9 ^a					206,63	290,11	0,35	0,04	0,17	0,04	0,71	0,22	
	Média	-	-	-	-	-	-	0,57	0,05	0,22	0,05	0,71	0,28	
Resistência à vassoura-de-bruxa	1 ^a a 9 ^a	0,68*	0,02 ^{ns}	1,13*	0,26 ^{ns}	3,29	5,39	0,59	0,004	0,21	0,05	0,61	0,34	0,97
		± 0,29	± 0,15	± 0,36	± ,30									

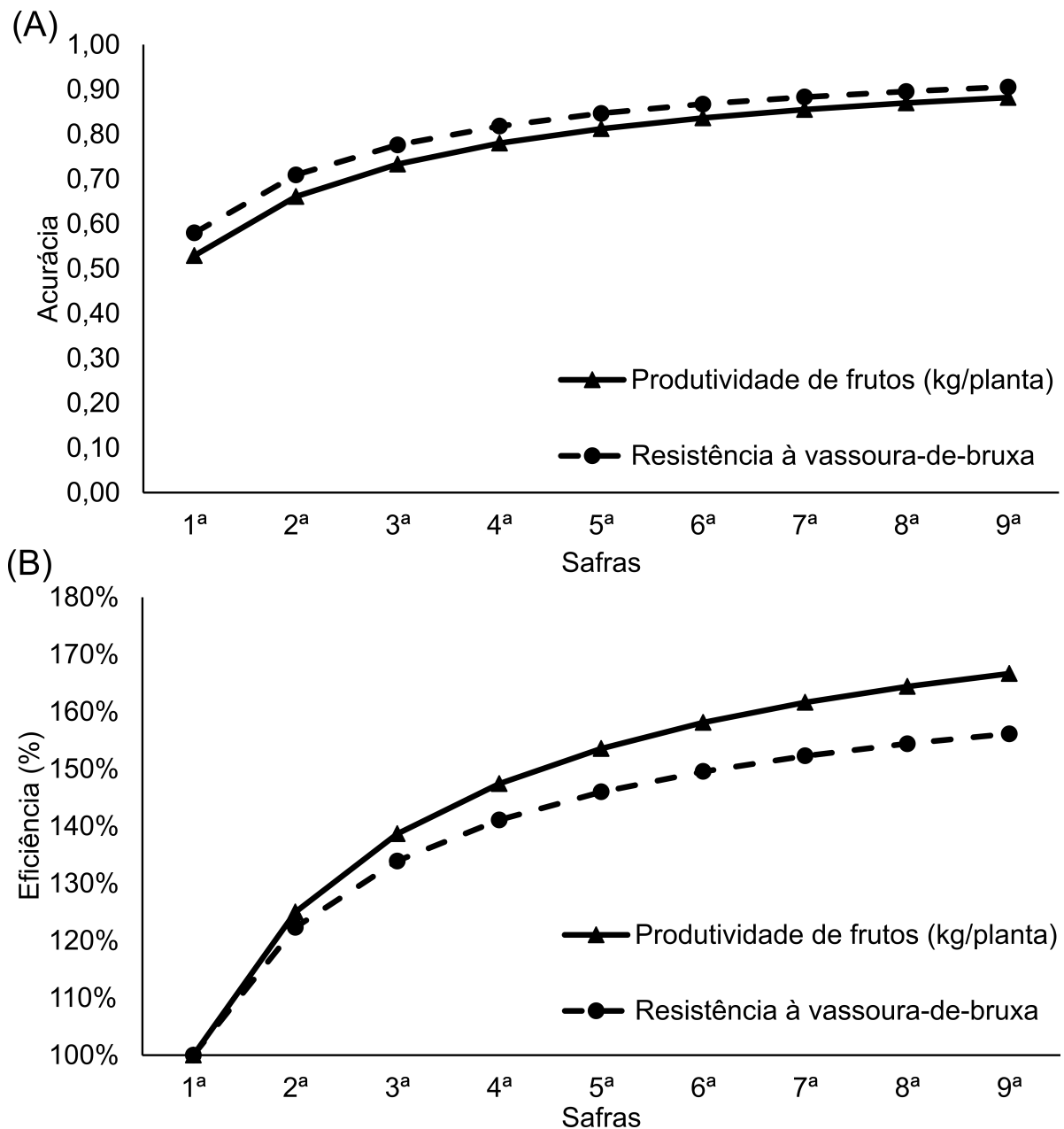
[†] σ_g^2 : variância genotípica; σ_{gm}^2 : variância da interação genótipos x medições (GM); σ_p^2 : variância dos efeitos de parcela; σ_i^2 : variância dos efeitos de ambiente; σ_e^2 : variância residual; σ_f^2 : variância fenotípica, H_g^2 : herdabilidade média no sentido amplo, c_{gm}^2 , c_p^2 , c_i^2 e c_e^2 : coeficientes de determinação dos efeitos da interação GM, de parcelas, de ambiente permanente e residual, respectivamente; ρ : coeficiente de repetibilidade; e r_{gm} : correlação genotípica através das medições; * e ^{ns}: significativo e não significativo à $P < 0,05$, respectivamente.

Todos os efeitos aleatórios contribuíram para explicar a variância fenotípica da característica produtividade de frutos (Tabela 2). O mesmo não pode ser dito para a resistência à vassoura-de-bruxa, onde os efeitos da interação GM e de parcelas foram não significativos, corroborado pelos baixos valores de seus coeficientes de determinação (Tabela 2). A baixa influência da interação GM refletiu na elevada correlação genética através das medições para resistência a vassoura-de-bruxa (0,97), o que não ocorreu para produtividade de frutos (0,56) (Tabela 2).

De forma geral, a variância residual explicou uma parte significativa da variância fenotípica total para ambas características, fato refletido no coeficiente de determinação dos efeitos residuais, 0,61 para resistência à vassoura-de-bruxa, e variando de 0,44 a 0,82 para produtividade de frutos. No modelo heterocedástico, empregado para a análise da produtividade por planta, é possível observar maior ou menor influência dos fatores não controláveis na manifestação do fenótipo, a depender da safra. Com isto, observa-se a flutuação da herdabilidade média, variando entre valores moderados (0,3 a 0,5) a altos ($> 0,5$) para esta característica (Tabela 2). A resistência à vassoura-de-bruxa também apresentou uma alta herdabilidade média (0,59) (Tabela 2). O coeficiente de repetibilidade teve grande variação para a produtividade, com amplitude de 0,13 a 0,42, acompanhando a flutuação da variância residual, com média de 0,28. Tratando-se da resistência, esse coeficiente apresentou valor ligeiramente maior (0,34) (Tabela 2).

As diferenças entre os coeficientes de repetibilidade dos caracteres em estudo refletem nas divergências observadas nas acurácias e eficiências de seleção ao longo das safras (Figura 1). Para a resistência à vassoura-de-bruxa, característica de maior repetibilidade, a avaliação em duas safras é suficiente para atingir acurácia de 0,70, satisfatória para seleção visando recombinação (RESENDE; DUARTE, 2007). Por outro lado, a produtividade precisa de três avaliações para alcançar este valor de acurácia (Figura 1A). Este fato é refletido na eficiência, a qual apresenta taxas de crescimentos cada vez menores para ambos caracteres, porém uma curva mais acentuada para a produtividade, ou seja, esta é mais beneficiada pelas repetidas avaliações (Figura 1B). Quando objetiva-se recomendar materiais genéticos, isto é, necessita-se acurácia de alta magnitude (0,90) (RESENDE; DUARTE, 2007), as características exigem o rigor de várias safras avaliadas, havendo necessidade de todas as nove medidas para resistência e produtividade (Figura 1A).

Figura 1. Acurácia (A) e eficiência (B) de seleção para produtividade de frutos e resistência à vassoura-de-bruxa avaliadas em híbridos triplos de cupuaçuzeiro, com o uso de m de safras.



Fonte: O autor (2022).

O ranqueamento pelos valores genotípicos e MHPRVG tiveram coincidência de 90% para a produtividade, considerando os 10 melhores híbridos (H117, H120, H125, H127, H129, H140, H162, H165, H167 e H169). A seleção destes híbridos proporcionará ganho de 12,54% em produtividade de frutos (Tabela 3). A consideração da adaptabilidade e estabilidade temporal, associada à produtividade, torna-se patente ao analisar conjuntamente a Figura 2 e o valor moderado da correlação genética através das medições, evidenciando a variação no comportamento dos genótipos através dos anos.

Tabela 3. Valor genotípico (BLUP), média harmônica da performance relativa do valor genotípico, multiplicada pela média geral ($MHPRVG_i * \mu$) e índice aditivo para os caracteres

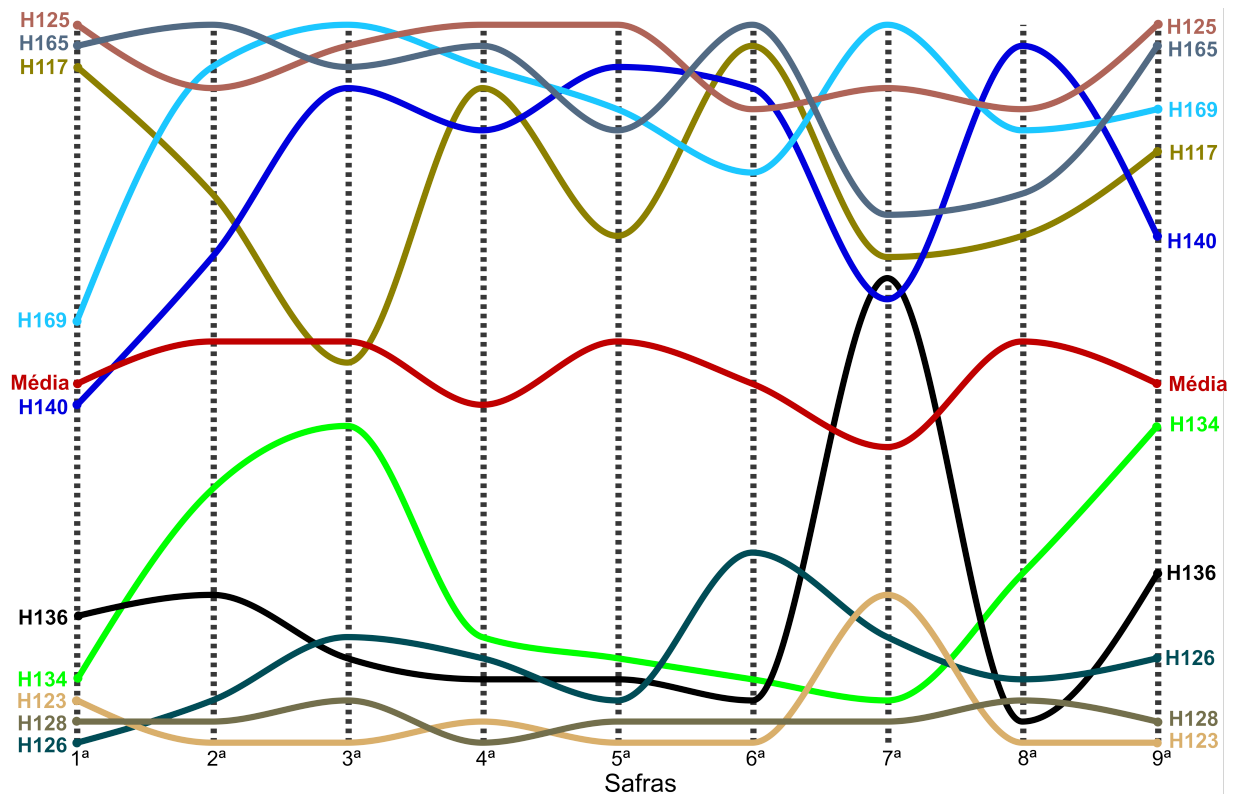
produtividade média de frutos (kg/planta) e resistência à vassoura-de-bruxa, em híbridos triplos de cupuaçuzeiro.

Genótipo	Produtividade média de frutos (kg/planta)		Resistência à vassoura-de-bruxa	Índice aditivo
	BLUP	MHPRVGi * u	BLUP	
H117	31,235	31,599	0,068	209,812
H118	26,510	26,038	-0,573	202,168
H120	30,781	30,799	0,391	203,597
H121	27,255	26,738	0,495	189,616
H123	21,040	19,493	0,042	174,022
H124	27,525	27,025	-0,805	209,08
H125	34,425	35,299	-0,316	226,578
H126	23,912	23,017	0,306	180,447
H127	30,142	30,444	-0,065	207,826
H128	22,467	21,428	-0,024	180,016
H129	31,113	30,788	-0,068	211,311
H130	28,751	28,633	-0,017	202,213
H131	28,228	28,119	-0,189	202,802
H132	25,902	25,514	-0,118	193,546
H133	25,330	25,052	-0,238	193,214
H134	25,190	24,616	-0,325	193,965
H135	26,902	26,865	-0,853	207,547
H136	24,655	23,785	-0,430	193,565
H137	27,627	27,866	-1,082	213,383
H138	28,158	27,809	0,395	194,238
H140	31,673	31,742	0,199	209,492
H143	30,129	29,827	-0,669	216,373
H144	29,631	29,014	0,110	203,52
H149	27,824	27,680	-0,445	205,015
H150	26,375	26,527	0,092	192,22
H152	26,358	26,242	0,130	191,632
H157	27,727	26,876	-0,520	205,727
H161	28,143	27,972	1,597	177,069
H162	31,899	31,145	1,555	190,993
H163	28,618	28,394	1,187	184,603
H165	33,473	34,438	-0,393	224,299
H166	27,118	27,370	-0,141	198,18
H167	29,901	30,019	-0,492	213,046
H169	32,689	33,063	1,197	198,887

Em negrito, os dez melhores híbridos para ambos os caracteres, segundo o índice aditivo

Figura 2. Valores genotípicos da produtividade média de frutos (kg/planta), através das safras avaliadas. Em destaque, o comportamento dos cinco melhores (H125, H165, H169, H117 e

H140) e cinco piores (H134, H136, H126, H128 e H123) genótipos para esta característica (vide Tabela 3) e a média.

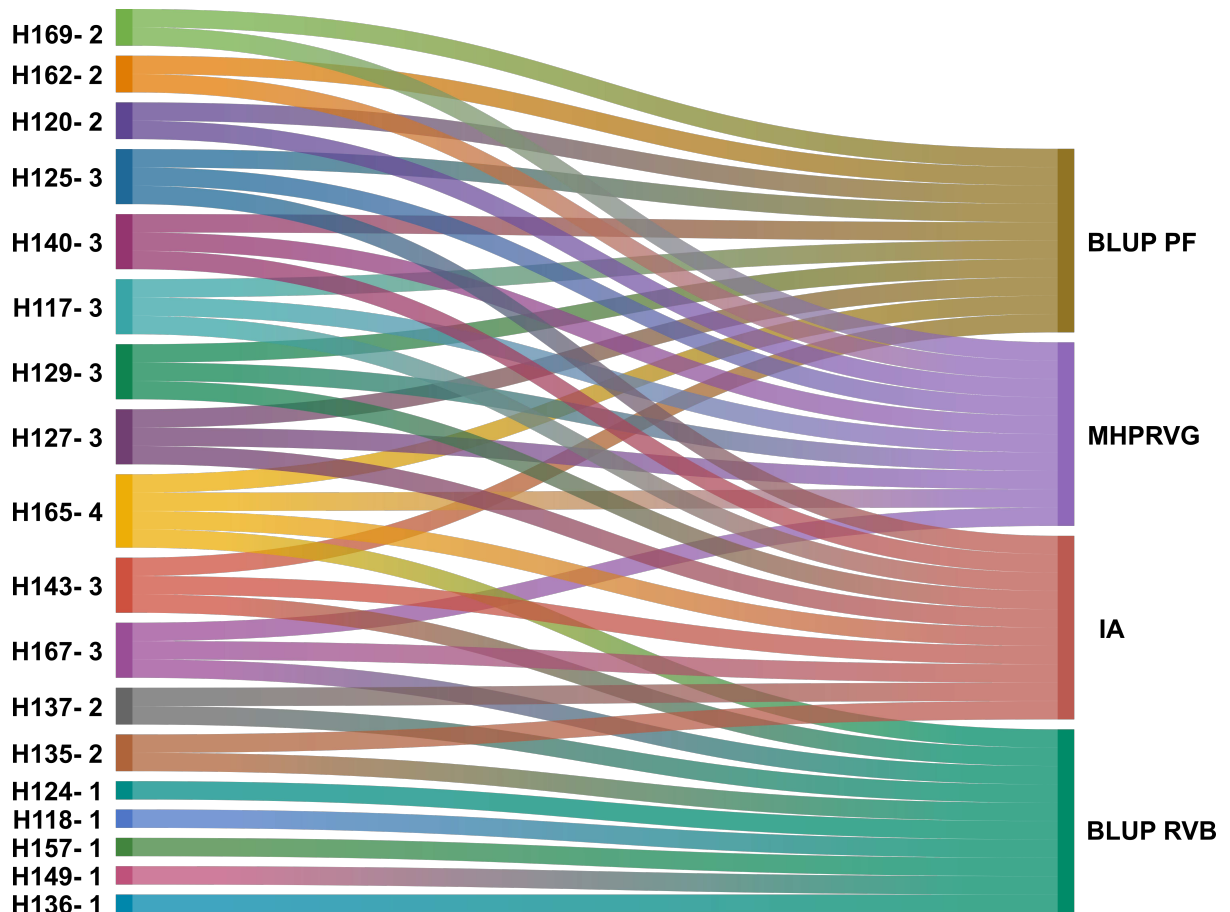


Fonte: O autor (2022)

Devido à não-significância dos efeitos da interação genótipos x medições para a resistência à vassoura-de-bruxa, o valor genotípico é suficiente para selecionar os melhores híbridos. Uma vez que atribuiu-se “1” para plantas sintomáticas e o interesse é identificar híbridos resistentes ao agente etiológico, quanto menor (ou mais negativo) o valor genotípico, mais adequado para a seleção é o híbrido. Neste sentido, destacam-se os 10 primeiros: H118, H124, H135, H136, H137, G143, H149, G157, H165 e G167, os quais proporcionarão um ganho estimado de 11,01% para resistência à vassoura-de-bruxa (Tabela 3).

Utilizando o índice aditivo para avaliar o desempenho simultâneo quanto a produtividade de frutos, considerando sua adaptabilidade e estabilidade, e resistência à vassoura-de-bruxa, dez híbridos destacam-se: H117, H125, H127, H129, H135, H137, H140, H143, H165 e H167. Dentre estes, sete encontram-se entre os 10 melhores para adaptabilidade, estabilidade e produtividade (H117, H125, H127, H129, H140, H165 e H167) e cinco configuram-se entre os dez melhores para resistência à vassoura-de-bruxa (H135, H137, H143, H165 e H167), com destaque para os genótipos H165 e H167, presentes em ambos critérios (Figura 3). A seleção dos materiais genéticos seguindo o índice aditivo proporcionará um ganho de 9,55% na produtividade de frutos e 5,23% na resistência à vassoura-de-bruxa.

Figura 3. Híbridos triplos de cupuaçuzeiro destaques (à esquerda) de acordo com quatro critérios (à direita): valor genotípico da produtividade média de frutos/planta (kg) (BLUP PF), média harmônica da performance relativa do valor genotípico da produtividade média de frutos/planta (kg) (MHPRVG), índice aditivo (IA) e valor genotípico da resistência à vassoura-de-bruxa (BLUP RVB). O número que acompanha o código do híbrido representa em quantos critérios ele se destaca, e as barras o relacionam para o referido critério. Foram considerados os 10 primeiros híbridos de cada critério



Fonte: O autor (2022)

4 DISCUSSÃO

Em experimentos com plantas perenes repetidos no tempo (safra) ou no espaço (locais), a probabilidade de haver homogeneidade entre variâncias residuais é pequena, devido as particularidades de cada safra/local (EDWARDS; JANNINK, 2006; JAIMEZ et al., 2020). Por este motivo, a melhor adequação do modelo heterocedástico era esperada para uma característica quantitativa como a produtividade de frutos, a qual é controlada por um grande número de genes de pequeno efeito e altamente influenciados pelos efeitos ambientais. Este resultado tem grande importância metodológica, por destacar a necessidade de considerar a

heterogeneidade das variâncias residuais nas análises de medidas repetidas (MARTÍNEZ-GARCÍA et al., 2017; SILVA et al., 2020). Tal fato é relevante em frutíferas perenes, como o cupuaçuzeiro, em que são realizadas repetidas mensurações no mesmo indivíduo, em várias safras (CHAVES et al., 2021).

A particularização das variâncias residuais para cada safra faz com que todos parâmetros derivados deste componente (variância fenotípica, herdabilidade, repetibilidade e coeficientes de determinação) também sejam particularizados. Com isto, os resultados refletem com maior verossimilitude o comportamento dos híbridos em estudo ao longo das safras avaliadas, concedendo maior robustez nas inferências geradas (FAVERI et al., 2015). Em paralelo, o modelo linear generalizado misto (MLGM) ajustado para a característica de resistência gera valores únicos.

A variabilidade genética para produtividade e resistência à vassoura-de-bruxa, atestada pela significância da variância genotípica, comprova a aptidão dessa população para seleção dos melhores genótipos, tanto para produtividade de frutos quanto para resistência à vassoura-de-bruxa. Por outro lado, há um comportamento diferencial de cada característica quanto a interação GM, significativa para produtividade, mas não para a resistência. Este padrão aponta que as variações ano-a-ano possuem maior efeito na produtividade de frutos, indicativo de que esta característica possui maior influência ambiental (ALVES et al., 2021).

A não significância da interação GM e a alta correlação genética através das medições sugerem que a resistência genética à vassoura-de-bruxa foi estável ao longo dos anos. Deste resultado, pode-se inferir que a presença de alelos favoráveis de diferentes fontes, dada pela composição tripla dos híbridos, contribuiu para a estabilidade da resistência, mesmo face a diferentes pressões do patógeno e condições ambientais (ALBUQUERQUE et al., 2010; BENJAMIN et al., 2016; PEREIRA et al., 2021). Estudos que abordem avaliação genética da resistência em campo são raros em cupuaçuzeiro (ALVES et al., 2009) sendo o presente estudo pioneiro no emprego de MLGM para análise da característica.

Segundo Resende e Alves (2020), a herdabilidade média no sentido amplo pode ser considerada alta para vassoura de bruxa ($> 0,50$) e mediana para produtividade ($0,15 < H_g^2 < 0,5$). Apesar de tais valores, a interferência ambiental na manifestação fenotípica evidencia-se ao considerar o decréscimo da herdabilidade a partir da quarta safra, para produtividade de frutos. Os diferentes padrões climáticos de cada safra, o estágio de maturação dos indivíduos e a intensificação da incidência de vassoura-de-bruxa são fatores que podem ter contribuído para este comportamento. O coeficiente de herdabilidade é um dos pilares nas avaliações genéticas,

prestando-se como base para tomadas de decisões acerca da melhor estratégia a ser adotada, refletindo na precisão e na eficácia da seleção (SCHMIDT et al., 2019).

O coeficiente de repetibilidade é outro parâmetro intimamente ligado à influência dos fatores aleatórios não controláveis e à variabilidade genética na população para determinada característica, visto que representa o limite superior da herdabilidade individual no sentido amplo (SILVA et al., 2020). Prova disso é a sua grande amplitude para a produtividade de frutos, partindo de um valor considerado mediano (0,40) para um valor baixo (0,13), seguindo a classificação de Resende e Alves (2020), a qual também pode ser utilizada para classificar a repetibilidade da resistência a vassoura-de-bruxa como moderada (0,34). As várias definições de repetibilidade estão relacionadas, em suma, com a correlação entre medidas repetidas em um mesmo indivíduo, considerando tanto efeitos genéticos quanto ambientais no seu cômputo (FERREIRA et al., 2021; LAVIOLA et al., 2013; MALIKOUSKI et al., 2021).

Face à baixos coeficientes de repetibilidade, necessita-se do emprego de um maior número de avaliações para uma seleção eficiente. Tendo uma acurácia de 0,70 como meta, proposta por Resende e Duarte (2007) como o valor mínimo para selecionar com maior segurança visando a recombinação, a produtividade deve ser avaliada por, no mínimo, três safras consecutivas, em concordância com o determinado por Mustiga et al. (2018) e Carvalho et al. (2002) para cacauzeiro, e Chaves et al. (2021) para cupuaçuzeiro. A resistência à vassoura-de-bruxa, característica com repetibilidade moderada, necessita de duas safras para uma seleção segura. Considerando os três primeiros anos de período juvenil, o ciclo atual de seleção do programa de melhoramento da espécie poderá reduzir sua duração pela metade. Este resultado possui grande relevância prática, e o presente estudo é o primeiro relato para a resistência à vassoura-de-bruxa.

Somente nas fases finais, quando o objetivo é recomendar ao produtor, a avaliação por 12 anos, considerando o período juvenil, como é usualmente realizada, fornecerá a confiabilidade necessária. Vale ressaltar que este resultado é válido para estudos semelhantes ao referido nesta pesquisa, onde uma maior quantidade de híbridos é testada em detrimento ao número de repetições. No cupuaçuzeiro, experimentos que visam recomendar materiais genéticos possuem maior escala, com um menor número de tratamentos avaliados em mais repetições e vários locais. Portanto, é possível que, sob estas condições, o número de safras necessárias para recomendação seja menor que o encontrado no presente estudo.

A predição dos efeitos da interação GM permite o cálculo da MHPRVG, método que considera conjuntamente a produtividade, a estabilidade e a adaptabilidade do genótipo (SOUZA et al., 2022). A coincidência entre nove dos 10 melhores híbridos, considerando os

valores genotípicos e a MHPRVG, indica que os híbridos mais produtivos também possuem maior estabilidade e adaptabilidade. Apesar da correspondência entre os métodos, a ordem dos híbridos não é a mesma em cada um, sugerindo que, a despeito dos altos índices produtivos, os materiais genéticos possuem estabilidade e adaptabilidade distintas. Percebe-se isso ao visualizar o comportamento médio dos híbridos em cada safra (Figura 2).

A ausência de interação GM para a resistência dispensa a utilização de métodos que capitalizam a estabilidade e adaptabilidade. Portanto, a seleção de híbridos para esta característica foi realizada exclusivamente via BLUP. O índice aditivo destacou os melhores híbridos para ambas características simultaneamente. Prova disto é a coincidência entre os dez melhores classificados pelo índice e pelos estimadores univariados (MHPRVG para produtividade e BLUP para resistência à vassoura-de-bruxa, Figura 3). A utilização de índices de seleção é recorrente nos programas de melhoramento, haja vista a constante necessidade de desenvolver cultivares produtivas e resistentes, aliadas a outras características (JAIMEZ et al., 2020). A eficiência do índice aditivo já foi atestada em outras culturas perenes, como *Jatropha curcas* L. (ALVES et al., 2018) e *Citrus spp.* (MALIKOUSKI et al., 2021). Dentre os materiais genéticos selecionados, somente dois destacam-se em ambos caracteres, o G165 e G167, indicando que há baixa correlação entre essas características (DUVAL et al., 2017) e que, inobstante à este fato, a genealogia empregada nos cruzamentos que originaram ambos híbridos foi bem sucedida em agregar alelos favoráveis, tanto para produtividade quanto para resistência.

A identificação destes híbridos é de suma importância para respaldar o programa de melhoramento com genótipos portadores de alelos de diferentes fontes, capazes de diversificar a base genética e, com isto, manter a variabilidade com os ciclos de seleção e recombinação. Desta forma, consegue-se realizar a piramidação dos genes, elevando os ganhos com a seleção e, em última instância, fornecendo genótipos produtivos e seguramente resistentes aos produtores (PILET-NAYEL et al., 2017; WILLOCQUET; SAVARY; YUEN, 2017).

Os resultados discutidos neste trabalho evidenciam a importância da modelagem para um programa de melhoramento. Esta ação permite a maximização da acurácia e, conseqüentemente, a melhoria da avaliação genética da população em estudo. Aliar a utilização de modelos e métodos estatísticos adequados à uma fenotipagem precisa e um bom controle experimental permitirá a seleção de genótipos verdadeiramente superiores e, como resultado, a otimização do programa de melhoramento genético (STRINGER et al., 2017).

5 CONCLUSÃO

O modelo heterocedástico apresentou menor AIC e permitiu maximizar a acurácia seletiva.

A duração de cada ciclo do programa de melhoramento do cupuaçuzeiro pode ser reduzida à metade, quando o objetivo for a recombinação. Somente quando o interesse for em recomendação para cultivo, nas fases finais do programa, recomenda-se a avaliação por nove safras ou 12 anos, contando com o período juvenil, nas condições de cultivo do presente estudo.

Considerando, simultaneamente, a produtividade, a adaptabilidade e a estabilidade, bem como, a resistência à vassoura-de-bruxa, os híbridos H125, H165, H143, H137, H167, H129, H117, H124, H140 e H135 foram selecionados.

REFERÊNCIAS

AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v. 19, n. 6, p. 716–723, 1974.

ALBUQUERQUE, P. S. B.; SILVA, S. D. V. M.; LUZ, E. D. M. N.; PIRES, J. L.; VIEIRA, A. M. C.; DEMÉTRIO, C. G. B.; PASCHOLATTI, S. F.; FIGUEIRA, A.

Novel sources of witches' broom resistance (causal agent *Moniliophthora perniciosa*) from natural populations of *Theobroma cacao* from the Brazilian Amazon. **Euphytica**, v. 172, n. 1, p. 125–138, 2010.

ALVES, R. M.; RESENDE, M. D. V.; BANDEIRA, B. S.; PINHEIRO, T. M.; FARIAS, D. C. R. Evolução da vassoura-de-bruxa e avaliação da resistência em progênies de cupuaçuzeiro. **Revista Brasileira de Fruticultura**, v. 31, n. 4, p. 1022–1032, 2009.

ALVES, R. M.; SILVA, C. R. S.; SILVA, M. S. C.; SILVA, D. C. S.; SEBBENN, A. M. Diversidade genética em coleções amazônicas de germoplasma de cupuaçuzeiro [*Theobroma grandiflorum* (Willd. ex Spreng.) Schum.]. **Revista Brasileira de Fruticultura**, v. 35, n. 3, p. 818–828, 2013.

ALVES, R. M.; CHAVES, S. F. S.; OLIVEIRA, R. P.; PEDROZA NETO, J. L.; SEBBENN, A. M. Canopy replacement used in the evaluation of cupuassu tree genotypes in the state of Pará. **Revista Brasileira de Fruticultura**, v. 42, n. 4, p. e-597, 2020.

ALVES, R. M.; CHAVES, S. F. S.; ALVES, R. S.; SANTOS, T. G.; ARAÚJO, D. G.; RESENDE, M. D. V. Cupuaçu tree genotype selection for an agroforestry system environment in the Amazon. **Pesquisa Agropecuária Brasileira**, v. 56, p. e02139, 2021.

ALVES, R. S. PEIXOTO, L. A.; TEODORO, P. E.; SILVA, L. A.; RODRIGUES, E. V.; RESENDE, M. D. V.; LAVIOLA, B. G.; BHERING, L. L. Selection of *Jatropha curcas*

families based on temporal stability and adaptability of genetic values. **Industrial Crops and Products**, v. 119, p. 290–293, 2018.

ATLIN, G. N.; CAIRNS, J. E.; DAS, B. Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change. **Global Food Security**, v. 12, p. 31–37, 2017.

BENJAMIN, C. S.; LUZ, E. D. M. N.; SANTOS, W. O.; PIRES, J. L. Cacao families and parents selected as resistant to natural infection of *Moniliophthora perniciosa*. **Crop Breeding and Applied Biotechnology**, v. 16, p. 141–146, 2016.

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

BUTLER, D. G.; CULLIS, B. R.; GILMOUR, A. R.; GOGEL, B. J.; THOMPSON, R. **ASReml-R reference manual Version 4**. Hemel Hempstead, UK: VSN International, 2018.

CARVALHO, C. G. P.; CRUZ, C. D.; ALMEIDA, C. M. V. C.; MACHADO, P. F. R. Yield repeatability and evaluation period in hybrid cocoa assessment. **Crop Breeding and Applied Biotechnology**, v. 2, n. 1, p. 149–156, 2002.

CHAVES, S. F. S.; ALVES, R. M.; ALVES, R. S.; SEBBENN, A. M.; RESENDE, M. D. V.; DIAS, L. A. S. *Theobroma grandiflorum* breeding optimization based on repeatability, stability and adaptability information. **Euphytica**, v. 217, n. 12, p. 211, 2021.

CULLIS, B. R.; SMITH, A. B.; COOMBES, N. E. On the design of early generation variety trials with correlated data. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 11, n. 4, p. 381–393, 2006.

DIAS, L. A. S.; SOUZA, C. A. S.; AUGUSTO, S. G.; SIGUEIRA, P. R.; MÜLLER, M. W. Performance and temporal stability analyses of cacao cultivars in Linhares, Brazil. **Plantations, recherche, développement**, v. 5, n. 5, p. 343–355, 1998.

DUVAL, A.; GEZAN, S. A.; MUSTIGA, G.; STACK, C.; MARELLI, J-P.; CHAPARRO, J.; LIVINGSTONE, D.; ROYAERT, S.; MOTAMAYOR, J. Genetic Parameters and the Impact of Off-Types for *Theobroma cacao* L. in a Breeding Program in Brazil. **Frontiers in Plant Science**, v. 8, p. 2059, 2017.

EDWARDS, J. W.; JANNINK, J. Bayesian Modeling of Heterogeneous Error and Genotype \times Environment Interaction Variances. **Crop Science**, v. 46, n. 2, p. 820–833, 2006.

EMBRAPA. **Mapas de solos e de aptidão agrícola de áreas alteradas do Pará**. 1. ed. Belém, Pará: Embrapa Amazônia Oriental, 2016.

FAVERI, J.; VERBYLA, A. P.; PITCHFORD, W. S.; VENKATANAGAPPA, S.; CULLIS, B. R. Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. **Crop and Pasture Science**, v. 66, n. 9, p. 947, 2015.

FERREIRA, F. M.; ROCHA, J. R. A. S. C.; ALVES, R. S.; ELIZEU, A. M.; BENITES, F. R. G.; RESENDE, M. D. V.; SOUZA SOBRINHO, F.; BHERING, L. L. Estimates of

repeatability coefficients and optimum number of measures for genetic selection of *Cynodon spp.* **Euphytica**, v. 216, n. 5, p. 70, 2020.

FERREIRA, F. M.; ROCHA, J. R. A. S. C.; BHERING, L. L.; FERNANDES, F. D.; LÉDO, F. J. S.; RANGEL, J. H. A.; KOPP, M.; CÂMARA, T. M. M.; SILVA, V. Q. R.; MACHADO, J. C. Optimal harvest number and genotypic evaluation of total dry biomass, stability, and adaptability of elephant grass clones for bioenergy purposes. **Biomass and Bioenergy**, v. 149, p. 106104, 2021.

GILMOUR, A. R.; GOGEL, B. J.; CULLIS, B. R.; WELHAM, S. J.; THOMPSON, R. **ASReml User Guide Release 4.1 Structural Specification**. Hemel Hempstead, UK: VSN International, 2015.

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

JAIMEZ, R. E.; VERA, D. I.; MORA, A.; LOOR, R. G.; BAILEY, B. A. A disease and production index (DPI) for selection of cacao (*Theobroma cacao*) clones highly productive and tolerant to pod rot diseases. **Plant Pathology**, v. 69, n. 4, p. 698–712, 2020.

LAVIOLA, B. G.; OLIVEIRA, A. M. C.; BHERING, L. L.; ALVES, A. A.; ROCHA, R. B.; GOMES, B. E. L.; CRUZ, C. D. Estimates of repeatability coefficients and selection gains in *Jatropha* indicate that higher cumulative genetic gains can be obtained by relaxing the degree of certainty in predicting the best families. **Industrial Crops and Products**, v. 51, p. 70–76, 2013.

LEMONS, O. F.; TREMACOLDI, C. R.; POLTRONIERI, M. C. **Boas práticas agrícolas para aumento da produtividade e qualidade da pimenta-do-reino no Estado do Pará**. Brasília, DF: Embrapa, 2014.

MALIKOUSKI, R. G.; PEIXOTO, M. A.; MORAIS, A. L.; ELIZEU, A. M.; ROCHA, J. R. A. S. C.; ZUCOLOTO, M.; BHERING, L. L. Repeatability coefficient estimates and optimum number of harvests in graft/rootstock combinations for “tahiti” acid lime. **Acta Scientiarum. Agronomy**, v. 43, p. e51740–e51740, 2021.

MARTÍNEZ-GARCÍA, P. J.; FAMULA, R. A.; LESLIE, C.; MCGRANAHAN, G. H.; FAMULA, T. R.; NEALE, D. B. Predicting breeding values and genetic components using generalized linear mixed models for categorical and continuous traits in walnut (*Juglans regia*). **Tree Genetics & Genomes**, v. 13, n. 5, p. 109, 2017.

MORAES, J. R. S. C.; ROLIM, G. S.; MARTORANO, L. G.; APARECIDO, L. E. O.; OLIVEIRA, M. S. P.; FARIAS NETO, J. T. Agrometeorological models to forecast açai (*Euterpe oleracea* Mart.) yield in the Eastern Amazon. **Journal of the Science of Food and Agriculture**, v. 100, n. 4, p. 1558–1569, 2020.

NEVES, L. C.; CAMPOS, A. J.; BENEDETTE, R. M.; TOSIN, J. M.; CHAGAS, E. A. Characterization of the antioxidant capacity of natives fruits from the Brazilian Amazon Region. **Revista Brasileira de Fruticultura**, v. 34, p. 1165–1173, 2012.

PATROCÍNIO, N. G. R. B.; CERESINI, P. C.; GOMES, L. I. S.; RESENDE, M. L. V.; MIZUBUTI, E. S. G.; GRAMACHO, K. P. Population structure and migration of the

witches' broom pathogen *Moniliophthora perniciosa* from cacao and cultivated and wild solanaceous hosts in southeastern Brazil. **Plant Pathology**, v. 66, n. 6, p. 900–911, 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.

PEIXOTO, M. A.; EVANGELISTA, J. S. P. C.; ALVES, R. S.; FARIAS, F. J. C.; CARVALHO, L. P.; TEODORO, L. P. R.; TEODORO, P. E.; BHERING, L. L. Models for optimizing selection based on adaptability and stability of cotton genotypes. **Ciência Rural**, v. 51, n. 5, p. e20200530, 2021.

PEREIRA, A. S.; LOPES, U. V.; PIRES, J. L.; BARRETO, M. A.; SANTOS, L. P.; GONZAGA, R. S.; SILVA, L. S.; ANDRADE, F. J.; BONILLA, J. L. S.; GRAMACHO, K. P. Quantitative resistance to witches' broom disease in progenies of different sources of resistance. **Crop Protection**, v. 146, p. 105678, 2021.

PEREIRA, F. A. C.; CARVALHO, S. P.; REZENDE, T. T.; OLIVEIRA, L. L.; MAIA, D. R. B. Selection of *Coffea arabica* L. hybrids usgin mixed models with different structures of variance-covariance matrices. **Coffee Science**, v. 13, n. 3, p. 8, 2018.

PIEPHO, H.-P.; MÖRING, J.; MELCHINGER, A. E.; BÜCHSE, A. BLUP for phenotypic selection in plant breeding and variety testing. **Euphytica**, v. 161, n. 1–2, p. 209–228, 2008.

PILET-NAYEL, M.-L.; MOURY, B.; CAFFIER, V.; MONTARRY, J.; KERLAN, M.-C.; FOURNET, S.; DUREL, C.-E.; DELOURME, R. Quantitative Resistance to Plant Pathogens in Pyramiding Strategies for Durable Crop Protection. **Frontiers in Plant Science**, v. 8, p. 1838, 2017.

QUINTAL, S. S. R.; VIANA, A. P.; CAMPOS, B. M.; VIVAS, M.; AMARAL JÚNIOR, A. T. Analysis of structures of covariance and repeatability in guava segregating population. **Revista Caatinga**, v. 30, n. 4, p. 885–891, 2017.

RAMOS, A. R.; VENTURIERI, G. A.; CUCO, S. M.; CASTRO, N. M. The site of selfincompatibility action in cupuassu (*Theobroma grandiflorum*). **Brazilian Journal of Botany**, v. 28, p. 569–578, 2005.

RESENDE, M. D. V. **Genética biométrica e Estatística no melhoramento de plantas perenes**. 1. ed. Brasília: Embrapa Informação Tecnológica, 2002. 975p.

RESENDE, M. D. V. **Métodos estatísticos ótimos para análise de experimentos de campo**, 2004. Disponível em:

<<https://www.infoteca.cnptia.embrapa.br/handle/doc/305549>>

RESENDE, M. D. V. **Matemática e estatística na análise de experimentos e no melhoramento genético**. 1. ed. Colombo: Embrapa Florestas, 2007. 561p.

RESENDE, M. D. V.; ALVES, R. S. Linear, generalized, hierarchical, bayesian and random regression mixed models in genetic/genomics in plant breeding. **Functional Plant Breeding Journal**, v. 2, n. 2, p. 1–31, 2020.

- RESENDE, M. D. V.; 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.
- SCHMIDT, P.; HARTUNG, J.; BENNEWITZ, J.; PIEPHO, H.-P. Heritability in Plant Breeding on a Genotype-Difference Basis. **Genetics**, v. 212, n. 4, p. 991–1008, 2019.
- SILVA, J. O. C.; BRUCKNER, C. H.; CARNEIRO, P. C. S.; RESENDE, M. D. V.; ALVES, R. S.; SILVA, D. F. P. Estimates of genetic parameters and repeatability coefficients in S2 peach genotypes grown in a subtropical environment. **Functional Plant Breeding Journal**, v. 2, n. 1, p. 25–33, 2020.
- SOUZA, A. G. C.; RESENDE, M. D. V.; SILVA, S. E. L.; SOUSA, N. R. The cupuaçu genetic improvement program at Embrapa Amazônia Ocidental. **Crop Breeding and Applied Biotechnology**, v. 2, n. 3, p. 471–478, 2002.
- SOUZA, A. G. C. **Boas práticas agrícolas da cultura do cupuaçuzeiro**. Manaus: Embrapa Amazônia Ocidental, 2007.
- SOUZA, B. M.; AGUIAR, A. V.; DAMBRAT, H. M.; GALUCHA, S. C.; TAMBARUSSI, E. V.; SESTREM, M. S. C.; TOMIGIAN, D. S.; FREITAS, M. L. M.; VENSON, I.; TORRES-DINI, D.; LONGUI, E. L. Effects of previous land use on genotype-by-environment interactions in two loblolly pine progeny tests. **Forest Ecology and Management**, v. 503, p. 119762, 2022.
- STRINGER, J.; ATKIN, F.; GEZAN, S. Statistical approaches in plant breeding: maximising the use of genetic information. In: CAMPOS, H.; CALIGARI, P. D. S (Eds.). **Genetic improvement of tropical crops**. 1. ed. Cham, Switzerland: Springer International Publishing, 2017. p. 3-17.
- TAHI, M.; TREBISSOU, C.; RIBEYRE, F.; GUIRAUD, B. S.; POKOU, D.; CILAS, C. Variation in yield over time in a cacao factorial mating design: changes in heritability and longitudinal data analyses over 13 consecutive years. **Euphytica**, v. 215, n. 6, p. 106, 2019.
- VERBYLA, A. P.; FAVERI, J.; DEERY, D. M.; REBETZKE, G. J. Modelling temporal genetic and spatio-temporal residual effects for high-throughput phenotyping data. **Australian & New Zealand Journal of Statistics**, v. 63, n. 2, p. 284–308, 2021.
- VILLEMEREUIL, P. Quantitative genetic methods depending on the nature of the phenotypic trait. **Annals of the New York Academy of Sciences**, v. 1422, n.1, p. 2947, 2018
- WILLOCQUET, L.; SAVARY, S.; YUEN, J. Multiscale Phenotyping and Decision Strategies in Breeding for Resistance. **Trends in Plant Science**, v. 22, n. 5, p. 420–432, 2017.
- WOLFINGER, R.; O'CONNEL, M. Generalized linear mixed models: a pseudolikelihood approach. **Journal of Statistical Computation and Simulation**, v. 48, n.3-4, p. 233-243, 2018

APÊNDICE A

Discriminação dos híbridos triplos de cupuaçuzeiro, parentais e número de repetições no estande inicial, do experimento instalado no município de Tomé Açu, Pará, Brasil.

Progênie	Parentais	Número de repetições
H117	174 x (186 x 434)	10
H118	174 x (186 x 434)	10
H120	174 x (186 x 434)	5
H121	174 x (186 x 434)	10
H123	174 x (186 x 554)	10
H124	174 x (186 x 554)	10
H125	174 x (186 x 554)	5
H126	174 x (186 x 554)	5
H127	174 x (186 x 1074)	10
H128	174 x (186 x 1074)	5
H129	174 x (186 x 1074)	5
H130	174 x (186 x 513)	10
H131	174 x (186 x 513)	10
H132	174 x (286 x 513)	10
H133	186 x (286 x 215)	5
H134	186 x (174 x 286)	5
H135	186 x (174 x 286)	10
H136	186 x (174 x 286)	10
H137	186 x (286 x 513)	10
H138	215 x (174 x 186)	10
H140	215 x (186 x 434)	5
H143	215 x (186 x 1074)	5
H144	215 x (186 x 1074)	5
H149	SEKÓ x (186 x 434)	5
H150	SEKÓ x (186 x 434)	5
H152	SEKÓ x (186 x 554)	10
H157	554 x (174 x 186)	5
H161	1074 x (186 x 434)	5
H162	1074 x (286 x 215)	5
H163	1074 x (286 x 215)	5
H165	1074 x (174 x 286)	10

H166	1074 x (186 x 554)	10
H167	1074 x (186 x 554)	5
H169	1074 x (186 x 513)	5

CAPÍTULO 2

MODELOS PARA ANÁLISE DE MEDIDAS REPETIDAS: APLICAÇÃO NA AVALIAÇÃO GENÉTICA DO CUPUAÇUZEIRO

VIÇOSA - MINAS GERAIS

2022

RESUMO

A dinamização da cadeia produtiva do cupuaçuzeiro (*Theobroma grandiflorum*) depende, substancialmente, do melhoramento genético. Essa abordagem deve ser respaldada por ensaios bem conduzidos, fenotipagem acurada e modelos estatísticos adequados para a avaliação e seleção genética. Assim, este estudo objetivou avaliar dez modelos com diferentes estruturas de covariância para os efeitos genéticos e não-genéticos, visando identificar aquele de melhor ajuste para a avaliação e seleção de híbridos de cupuaçuzeiro. Para isso, 34 híbridos triplos de cupuaçuzeiro, implantados em blocos completos casualizados com número variável de repetições, foram avaliados durante nove safras quanto a produtividade média de frutos (kg/planta). O ajuste de cada modelo foram avaliados pelo critério de informação de Akaike (AIC). O modelo mais adequado para a estimação dos componentes de variância e para a predição dos valores genotípicos foi aquele em que as matrizes dos efeitos residuais, de ambiente permanente, de parcelas e genéticos foram modeladas pelas estruturas autoregressiva heterogênea de primeira ordem, bloco diagonal, simetria composta heterogênea e fator analítico, respectivamente. A partir desse modelo, dez híbridos foram selecionados, os quais deverão ser incorporados ao programa de melhoramento, sendo as matrizes destaques dentre as plantas referentes à tais híbridos priorizados para futuras hibridações ou ensaios de competição clonal.

Palavras-chave: *Theobroma grandiflorum*, estruturas de covariâncias, fator analítico, autoregressivo heterogêneo, simetria composta heterogênea.

1 INTRODUÇÃO

O cupuaçuzeiro (*Theobroma grandiflorum* Willd. (Ex. Spreng.) Schum.), Malvaceae, é uma espécie frutífera nativa da região amazônica que tem a polpa e as amêndoas do fruto, o cupuaçu, como principais alicerces da cadeia produtiva (PEREIRA; ABREU; RODRIGUES, 2018). O valor agroindustrial do fruto atribuí à espécie bom potencial mercadológico, capaz de transcender as fronteiras regionais. Possui ampla aceitação, sobretudo na região Norte brasileira, onde existem mais de 16 mil hectares dedicados ao seu cultivo, gerando emprego e renda para a população local (IBGE, 2017).

A modernização da cadeia produtiva do cupuaçuzeiro depende do melhoramento genético da cultura. A utilização de sementes provenientes de matrizes não selecionadas nos plantios pioneiros, isto é, que não passaram por nenhum processo de melhoramento, causou expressiva desuniformidade na maioria dos pomares amazônicos (ALVES; CHAVES, 2020). A produtividade está aquém do potencial da cultura mesmo nos estados de maior produção, 3,2 t de frutos/ha no Pará e 2,0 t de frutos/ha no Amazonas (AMAZONAS, 2019; PARÁ, 2018). A grande divergência intra e inter-populacional, condicionada pela alogamia ampliada pela autoincompatibilidade do cupuaçuzeiro também contribuiu para este fato (ALVES et al., 2007).

O programa de melhoramento da cultura está ativo há quase quatro décadas para superar tais problemas. No entanto, a natureza perene da espécie demanda o emprego de métodos experimentais e estatísticos diferenciados (SOUZA et al., 2002). A análise de medidas repetidas, realizada em dados coletados no mesmo indivíduo em diferentes safras, possui peculiaridades, a exemplo das covariâncias (genéticas e/ou ambientais) entre as várias medidas (ANDRADE et al., 2016; PEIXOTO et al., 2021; PIEPHO; ECKL, 2014). Tais peculiaridades não podem ser ignoradas, visto que, para o alcance de resultados concretos, confiáveis e acurados, deve-se estabelecer a base que irá fundamentar as estratégias de melhoramento: i) delineamento experimental apropriado; ii) coleta de dados acurada; iii) análises estatísticas adequadas (FAVERI et al., 2015; SMITH; SPANGENBERG, 2014; STRINGER; ATKIN; GEZAN, 2017).

Dentre as várias alternativas existentes para a análise de medidas repetidas não é possível estabelecer, a priori, a mais apropriada a qualquer situação, mas sim a que mais se adequa ao conjunto de dados. Em termos de modelos lineares mistos, as alternativas vão desde o modelo de repetibilidade, o qual considera homogeneidade de variâncias dos efeitos genéticos e não genéticos, bem como, alta correlação genética ($>0,80$) através das medições (RESENDE; SILVA; AZEVEDO, 2014); até o modelo multivariado com emprego da matriz

de variância não-estruturada, o qual considera que cada medida se comporta como um caráter distinto (ALVES et al., 2019). O modelo de repetibilidade é o mais simples e parcimonioso, porém pode não traduzir o real comportamento das plantas em campo, haja vista as expressivas diferenças observadas ano-a-ano, as quais podem tornar as variâncias heterogêneas, proporcionando correlações genéticas menores que 0,8 (EDWARDS; JANNINK, 2006; MRODE, 2014). Por outro lado, o modelo multivariado, por particularizar as medidas, dispõe de estimativas mais acuradas e fiéis ao comportamento do genótipo em cada safra, porém traz consigo uma complexidade estatística e computacional potencialmente impeditiva, principalmente quando mais de três medidas são consideradas, devido a super-parametrização (DIGGLE, 1988; WARD et al., 2019).

Em vista do supracitado, supõem-se que por meio da modelagem das matrizes de covariância dos efeitos genéticos e não-genéticos será possível determinar o modelo mais adequado para a análise de medidas repetidas. Neste contexto, este trabalho objetivou avaliar modelos com diferentes estruturas de covariância, para os efeitos genéticos e não-genéticos, visando identificar aquele com melhor ajuste para a avaliação e seleção de genótipos de cupuaçuzeiro.

2 MATERIAIS E MÉTODOS

2.1 Condições experimentais

O experimento foi instalado em uma propriedade rural (latitude 02° 24' 05,6" S; longitude 48° 00' 08,8" W; altitude 45 m acima do nível do mar) no município de ToméAçu, Pará, em fevereiro de 2007. O clima da região corresponde ao tipo Ami, de acordo com a classificação de Köppen, e ao tipo B4rA'a', de acordo com a classificação de Thornthwaite (MORAES et al., 2020). O solo é do tipo Latossolo Amarelo, textura média, com baixa fertilidade, profundo e bem drenado (EMBRAPA, 2016). A temperatura média, no período avaliado, foi de aproximadamente 27 °C, com umidade relativa do ar de 85,5% e precipitação pluviométrica média anual de 2.729 mm, de acordo com os dados fornecidos pela estação meteorológica da Embrapa Amazônia Oriental, instalada no município.

O cupuaçuzeiro foi implantado em consórcio com um plantio comercial de pimenteira do reino (*Piper nigrum* L.), que estava no quinto ano de campo, em espaçamento de linhas duplas [(2 x 2) x 4 m]. As plantas de cupuaçuzeiro foram estabelecidas nas entrelinhas, no espaçamento de 6 x 4 m. Os tratos culturais seguiram as recomendações técnicas dos sistemas

de produção de ambas as espécies (LEMOS; TREMACOLDI; POLTRONIERI, 2014; SOUZA, 2007).

2.2 Genótipos avaliados

As progênes utilizadas no experimento foram oriundas de cruzamentos triplos, obtidas no esquema $(C_1 \times C_2) \times C_3$, valendo-se de contribuições gênicas de três clones (C) distintos. Foram avaliados 34 híbridos triplos não-convencionais de cupuaçuzeiro, dispostos no delineamento blocos completos casualizados com até dez repetições e três plantas por parcela. Tais híbridos são ditos “não-convencionais” pois não são provenientes de cruzamentos entre linhagens endogâmicas, como é usualmente realizado. O número distinto de repetições foi devido à disponibilidade de mudas para o plantio.

A partir do terceiro ano após a implantação, foi iniciada a avaliação da produtividade média de frutos por planta (kg), produto da multiplicação do número de frutos pelo peso médio dos frutos daquela planta. O valor de uma safra corresponde à soma das quatro avaliações. Os dados foram coletados em nível de planta durante nove safras.

2.3 Análises estatísticas

Todas as análises estatísticas realizadas no presente estudo tiveram como base o método da máxima verossimilhança restrita (REML, PATTERSON; THOMPSON, 1971), para estimação dos componentes de variância e parâmetros genéticos; e a melhor predição linear não-viesada (BLUP, HENDERSON, 1975), para predição dos valores genotípicos.

As modelagens dos efeitos genéticos e não genéticos foram realizadas a partir de um modelo-base, de modo sequencial (Ver Faveri et al., 2015 e Melo et al., 2020), associado à avaliação de híbridos em blocos completos casualizados, com mais de uma planta por parcela e com medidas repetidas, dado por:

$$y = X_m + Z_g + W_p + Q_i + e,$$

onde y é o vetor de dados fenotípicos; m é o vetor dos efeitos de medições, repetições e suas combinações (assumidos como fixos), somados à média geral; g é o vetor dos efeitos genotípicos, assumidos como aleatórios ($g \sim N [0, G_g]$); p é o vetor dos efeitos aleatórios de parcela ($p \sim N [0, G_p]$); i é o vetor dos efeitos aleatórios de ambiente permanente de indivíduo ($i \sim N [0, G_i]$); e e é o vetor dos resíduos (aleatórios) ($e \sim N [0, R]$). As letras X, Z, W e Q

representam as matrizes de incidência para m , g , p e i , respectivamente. G_g , G_p , G_i e R referem-se às matrizes de covariância para os efeitos genotípicos, de parcelas, de ambiente permanente e resíduos, respectivamente. Estas matrizes foram modeladas com diferentes estruturas de covariância.

A atribuição da natureza aleatória aos efeitos genéticos tem como base os trabalhos de Piepho et al. (2008) e Resende e Duarte (2007), os quais concluíram que quando o número de tratamentos (híbridos) é grande (> 10), tratar esse efeito como aleatório é vantajoso. De fato, os efeitos genéticos só devem ser considerados de natureza fixa quando não for possível estimar componentes de variâncias a partir dos dados e/ou quando o principal objetivo for a comparação entre um número reduzido de cultivares (PIEPHO et al., 2008).

2.3.1 Modelagem dos efeitos residuais

A matriz de covariância dos efeitos residuais foi modelada com três estruturas distintas. A primeira, chamada identidade de variâncias (IDV), considera variâncias homogêneas e ausência de covariâncias entre pares de medidas (ISIK; HOLLAND; MALTECCA, 2017):

$$R = \begin{bmatrix} \sigma_e^2 & 0 & \cdots & 0 \\ & \sigma_e^2 & \cdots & 0 \\ & & \ddots & 0 \\ \text{Sim.} & & & \sigma_e^2 \end{bmatrix}.$$

A consideração da estrutura IDV neste primeiro modelo o caracteriza como o modelo simples de repetibilidade (o mais simples e parcimonioso).

A segunda estrutura considera que as variâncias residuais entre medidas são heterogêneas, assumindo, ainda, a nulidade de covariâncias entre pares de medidas. Esta estrutura é denominada bloco diagonal (IDH) (ISIK; HOLLAND; MALTECCA, 2017):

$$R = \begin{bmatrix} \sigma_{e_1}^2 & 0 & \cdots & 0 \\ & \sigma_{e_2}^2 & \cdots & 0 \\ & & \ddots & 0 \\ \text{Sim.} & & & \sigma_{e_9}^2 \end{bmatrix}.$$

A terceira e última estrutura modelada para os efeitos residuais é a auto-regressiva de primeira ordem com variâncias heterogêneas (AR1H), a qual assume heterogeneidade de covariâncias entre os efeitos residuais das diferentes medidas, bem como, a presença de uma autocorrelação ou correlação serial, isto é, um coeficiente de correlação que diminui à medida que o intervalo entre safras aumenta (ISIK; HOLLAND; MALTECCA, 2017; RESENDE; SILVA; AZEVEDO, 2014):

$$R = \begin{bmatrix} \sigma_{e_1}^2 & \sigma_{e_1}\sigma_{e_2}\rho^{2-1} & \sigma_{e_1}\sigma_{e_3}\rho^{3-1} & \cdots & \sigma_{e_1}\sigma_{e_9}\rho^{9-1} \\ & \sigma_{e_2}^2 & \sigma_{e_2}\sigma_{e_3}\rho^{3-2} & \cdots & \sigma_{e_2}\sigma_{e_9}\rho^{9-2} \\ & & \sigma_{e_3}^2 & \cdots & \sigma_{e_3}\sigma_{e_9}\rho^{9-3} \\ & & & \ddots & \vdots \\ \text{Sim.} & & & & \sigma_{e_9}^2 \end{bmatrix}.$$

Nas notações acima, $\sigma_{e_m}^2$ é a variância residual na medição m , σ_{e_m} é o desvio padrão residual na medição m e ρ é o coeficiente de autocorrelação ou correlação serial. O termo “Sim.” na porção inferior das matrizes indica simetria.

2.3.2 Modelagem dos efeitos de ambiente permanente de indivíduo

Duas estruturas foram testadas para os efeitos de ambiente permanente de indivíduo: a IDV e a IDH, seguindo as notações demonstradas no tópico 2.3.1.

2.3.3 Modelagem dos efeitos de parcela

Com relação aos efeitos de parcela, foi ajustada, primeiramente, a estrutura bloco diagonal (IDH). Posteriormente, testou-se a modelagem via modelo auto-regressivo heterogêneo de primeira ordem (AR1H), seguindo a notação apresentada no tópico 2.3.1. Finalmente, o modelo de simetria composta heterogênea (CORH) foi ajustado, o qual considera heterogeneidade de variâncias e covariâncias, bem como a presença de um único coeficiente de correlação ponderando as covariâncias (GILMOUR et al., 2015):

$$G_p = \begin{bmatrix} \sigma_{p_1}^2 & \sigma_{p_1}\sigma_{p_2}\rho & \sigma_{p_1}\sigma_{p_3}\rho & \cdots & \sigma_{p_1}\sigma_{p_9}\rho \\ & \sigma_{p_2}^2 & \sigma_{p_2}\sigma_{p_3}\rho & \cdots & \sigma_{p_2}\sigma_{p_9}\rho \\ & & \sigma_{p_3}^2 & \cdots & \sigma_{p_3}\sigma_{p_9}\rho \\ & & & \ddots & \vdots \\ \text{Sim.} & & & & \sigma_{p_9}^2 \end{bmatrix},$$

onde $\sigma_{p_m}^2$ é a variância do efeito de parcela na medição m , σ_{p_m} é o desvio padrão do efeito de parcela na medição m e ρ é o coeficiente de correlação associado a cada covariância entre par de medições.

2.3.4 Modelagem dos efeitos genotípicos

Os efeitos genotípicos foram modelados seguindo três estruturas: bloco diagonal (IDH), simetria composta heterogênea (CORH), ambas já descritas nos tópicos 2.3.1 e 2.3.3, respectivamente; e a estrutura fator analítico (FA1), geralmente empregada na análise de ensaios em múltiplos ambientes, mas que pode ser ajustada à medidas repetidas, ao considerar cada safra como um caráter individual e a presença de correlação genética entre pares de safras (PIEPHO, 1997). No presente estudo, considerou-se um modelo de primeira ordem (SMITH et al., 2015):

$$G_p = \begin{bmatrix} \lambda_{11}^2 + \psi_1 & \lambda_{11}\lambda_{12} & \lambda_{11}\lambda_{13} & \cdots & \lambda_{11}\lambda_{19} \\ & \lambda_{12}^2 + \psi_2 & \lambda_{12}\lambda_{13} & \cdots & \lambda_{12}\lambda_{19} \\ & & \lambda_{13}^2 + \psi_3 & \cdots & \lambda_{13}\lambda_{19} \\ & & & \ddots & \vdots \\ \text{Sim.} & & & & \lambda_{19}^2 + \psi_9 \end{bmatrix} \otimes I_g,$$

onde λ_{1m}^2 é a primeira carga fatorial na medição m , ψ_m é a variância específica de cada medição m , \otimes é o produto de Kronecker e I_g é uma matriz identidade de ordem $g \times g$. Os elementos fora da diagonal representam as covariâncias. No modelo fator analítico, os EBLUPs (\hat{g}_i) são calculados pelo estimador:

$$\hat{g}_i = \lambda_{1m} + f_{1i} + \psi_m, \text{ em que } f_{1i} \text{ é o escore fatorial do genótipo } i \text{ no fator } 1.$$

Tal modelo permite, ainda, a estimação da correlação genética entre medições:

$$r'_{mm} = \frac{\lambda_{1m}\lambda_{1m'}}{\sqrt{\sigma_{g_m}^2 \sigma_{g_{m'}}^2}},$$

onde $\sigma_{g_m}^2$ é a variância genotípica da medição m .

As estruturas das matrizes R , G_i , G_p e G_g nos dez modelos ajustados estão resumidas na Tabela 1. Os efeitos residuais foram os primeiros a serem modelados, seguidos dos efeitos de ambiente permanente de indivíduo e parcela, finalizando com os efeitos de genótipos.

Tabela 1. Estruturas das matrizes de covariâncias dos efeitos de resíduo, ambiente permanente de indivíduo, parcela e genótipo, nos dez modelos ajustados para análise de medidas repetidas em cupuaçuzeiro.

Modelo	Efeito			
	Resíduo (R)	Ambiente permanente (G_i)	Parcela (G_p)	Genótipo (G_g)
1	IDV [†]	IDV	IDV	IDV
2	IDH [†]	IDV	IDV	IDV
3	AR1H [†]	IDV	IDV	IDV
4	AR1H	IDH	IDV	IDV
5	AR1H	IDH	IDH	IDV
6	AR1H	IDH	AR1H	IDV
7	AR1H	IDH	CORH [†]	IDV

8	AR1H	IDH	CORH	IDH
9	AR1H	IDH	CORH	CORH
10	AR1H	IDH	CORH	FA1 [†]

[†] IDV: Identidade de variâncias; IDH: Bloco diagonal (variâncias heterogêneas); AR1H: auto-regressiva de primeira ordem com variâncias heterogêneas; CORH: Simetria composta com variâncias heterogêneas; FA1: Fator analítico de primeira ordem.

2.3.3 Seleção do modelo

As diferentes estruturas de covariância foram avaliadas pelo critério de informação de Akaike (AIC) (AKAIKE, 1974), dado por:

$$AIC = -2\text{Log}L + 2p,$$

onde $\text{Log}L$ é o logaritmo do máximo da função de verossimilhança restrita e p é o número de parâmetros estimados. Quanto menor o AIC, mais adequado é o modelo para o conjunto de dados.

2.3.4 Estimativas dos parâmetros genéticos e não genéticos

A significância dos componentes de variância obtidos diretamente dos modelos foi testada utilizando os intervalos de confiança (IC) (GILMOUR et al., 2015):

$$ICx = \sigma_x^2 \pm (1,96 EPx),$$

onde σ_x^2 é a estimativa do componente de variância x , EPx é o erro padrão da estimativa do componente de variância x e 1,96 é o valor tabelado do teste t de Student em nível de 5% de probabilidade de erro Tipo I, com infinitos graus de liberdade.

A título de comparação, as inferências foram realizadas a partir do modelo simples de repetibilidade (modelo 1 - comumente empregado na avaliação genética de fruteiras perenes) e do modelo de melhor ajuste. Foram estimadas a variância fenotípica ($\sigma_f^2 = \sigma_g^2 + \sigma_{gm}^2 + \sigma_p^2 + \sigma_i^2 + \sigma_e^2$), a herdabilidade média no sentido amplo [$H_g^2 = 1 - \frac{\overline{\Delta}_g}{2\sigma_g^2}$, em que $\overline{\Delta}_g$ é a variância do erro de predição média entre genótipos par a par (CULLIS; SMITH; COOMBES, 2006)], os coeficientes de determinação dos efeitos de parcela ($c_p^2 = \frac{\sigma_p^2}{\sigma_f^2}$) e de ambiente permanente de indivíduo ($c_i^2 = \frac{\sigma_i^2}{\sigma_f^2}$), além dos efeitos residuais ($c_e^2 = \frac{\sigma_e^2}{\sigma_f^2}$); o coeficiente de repetibilidade ($\rho = \frac{\sigma_g^2 + \sigma_i^2}{\sigma_f^2}$) e a acurácia seletiva [$r = \sqrt{1 - \left(\frac{PEV}{\sigma_g^2}\right)}$], onde PEV é a variância do erro de predição,

obtida da diagonal da inversa generalizada da matriz de coeficientes das equações de modelos mistos (RESENDE; SILVA; AZEVEDO, 2014).

Todas as análises foram realizadas no software R (R CORE TEAM, 2022), utilizando o pacote ASReml-R (BUTLER et al., 2018).

3 RESULTADOS

A estrutura AR1H foi a que apresentou menor valor de AIC e, conseqüentemente, o melhor ajuste para os efeitos residuais (Tabela 2). O número de parâmetros (10) referentes à matriz R (Tabela 2) associados a este modelo refere-se às variâncias residuais das nove safras e o coeficiente de autocorrelação associado às covariâncias entre pares de safras.

Tabela 2. Estruturas da matriz de covariância dos efeitos residuais (R), logaritmo da verossimilhança (L), número de parâmetros (P) associados às matrizes G e R, critério de informação de Akaike (AIC) e acurácias seletivas referentes a três modelos para análise de medidas repetidas em cupuaçuzeiro.

Modelo	Estrutura da matriz R [†]	L	P		AIC	Acurácia seletiva
			G	R		
1	IDV	-19918,93	3	1	40003,68	0,658
2	IDH	-19581,53	3	9	39187,06	0,650
3	AR1H	-19559,32	3	10	39144,65	0,643

Em negrito, o modelo de melhor ajuste.

O comportamento dos efeitos de ambiente permanente de indivíduo foi melhor explicado pela estrutura IDH, em detrimento da estrutura mais simples, IDV. Já os efeitos de parcela apresentaram menor valor de AIC quando considerou-se a estrutura CORH. Por fim, o efeito de genótipos foi melhor explicado pela estrutura FA1 (Tabela 3). Portanto, o modelo selecionado para estimação dos componentes de variância e predição dos valores genotípicos (Modelo 10) considera a estrutura AR1H para os efeitos residuais, IDH para os efeitos de ambiente permanente, CORH para os efeitos de parcela e FA1 para os efeitos de genótipos. Os parâmetros associados às matrizes G (G_g , G_p , G_i) no modelo 10 (Tabela 3) referem-se aos componentes de variância particularizados para cada safra, isto é, o número de medidas corresponde à quantidade de parâmetros para os efeitos de ambiente permanentes (nove parâmetros) e parcela (nove parâmetros). Para este último efeito, existe, ainda, um coeficiente de correlação associado às covariâncias entre pares de safras (Vide estrutura CORH na seção anterior). Por fim, além das nove variâncias específicas para cada safra, os efeitos de genótipos contam com nove cargas fatoriais. Desta forma, totalizaram-se 37 parâmetros.

Tabela 3. Estruturas das matrizes de covariância dos efeitos de ambiente permanente, parcela e genotípicos; logaritmo da verossimilhança (L), número de parâmetros (P) associados às matrizes G e R, critério de informação de Akaike (AIC) e acurácia seletiva de oito modelos ajustados para análise de medidas repetidas em cupuaçuzeiro.

Modelo	Efeito modelado	Estrutura da matriz G	L	P		AIC	Acurácia seletiva
				G	R		
3	Indivíduo	IDV	-19559	3	10	39144,65	0,643
4		IDH	-19485	11	10	39012,87	0,510
5	Parcela	IDH	-19451	19	10	38959,06	0,641
6		AR1H	-19438	20	10	38937,98	0,550
7 [‡]		CORH	-19434	20	10	38927,70	0,603
8 [‡]	Genótipo	IDH	-	-	-	-	-
9		CORH	-19383	29	10	38843,15	0,621
10		FA1	-19368	37	10	38829,34	0,665

[‡]Não atingiu a convergência. Em negrito, o modelo de melhor ajuste

Segundo o Modelo 1, mais simples e parcimonioso, todos os efeitos aleatórios foram significativos, isto é, todos explicam os fenótipos observados. No Modelo 10, a particularização dos componentes de variância para cada safra permite a visualização da maior ou menor influência de cada componente para a variância fenotípica. Isto é evidente, por exemplo, para os efeitos de parcela, cujo componente de variância foi não significativo nas safras 2, 6 e 8 (Tabela 4).

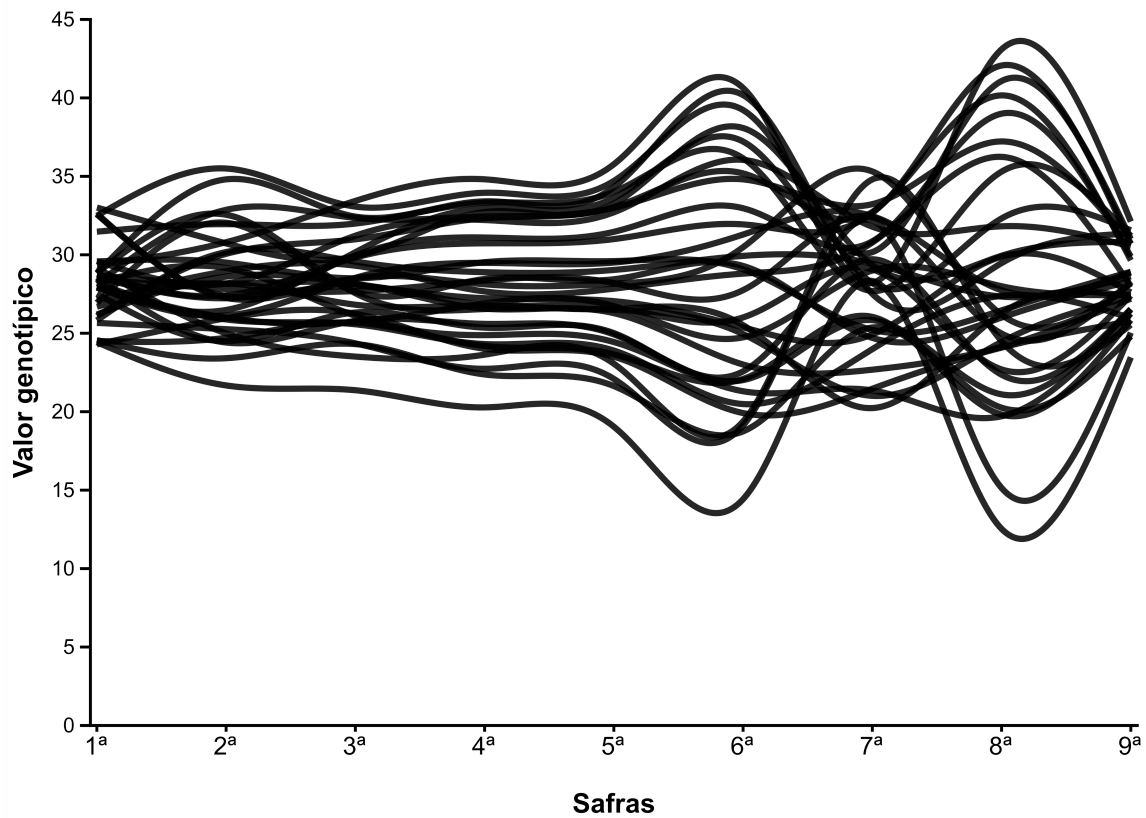
Tal oscilação é melhor retratada pelos coeficientes de determinação. Observa-se que, depois dos efeitos residuais, os efeitos de ambiente permanente de indivíduo são os mais influentes sobre o fenótipo. A herdabilidade média no sentido amplo, de alta magnitude, varia entre 0,55 a 0,82. A acurácia de seleção também se manteve alta (0,55 a 0,89) ao longo de todas as safras, assim como o coeficiente de repetibilidade, que atingiu valores moderados a altos (0,45 a 0,78). A média da produtividade de frutos foi crescente até a quinta safra, oitavo ano após a implantação, período pré-climáx. A partir desta safra, quando as plantas atingem a maturidade produtiva, observou-se uma tendência de bienalidade, isto é, uma safra de alta produtividade seguida de uma safra com baixa produtividade (Figura 1).

Tabela 1. Estimativas dos componentes de variância (\pm intervalo de confiança) e parâmetros genéticos e não genéticos para o caráter produtividade média de frutos (kg/planta) estimados a partir dos modelos 1 (mais simples) e 10 (melhor ajuste) em híbridos triplos de cupuaçuzeiro.

Modelo	Safr	Componentes de variância					Coeficientes de determinação			H_g^2	ρ	r	μ	
		σ_g^2	σ_i^2	σ_p^2	σ_e^2	σ_f^2	c_i^2	c_p^2	c_e^2					
1	1 ^a a 9 ^a	15,07*	52,07*	14,87*	178,32	260,33	0,20	0,06	0,68	0,71	0,25	0,83	28,20	
		$\pm 10,31$	$\pm 9,29$	$\pm 8,33$										
		8,64*	64,36*	7,23*	34,13	114,37	0,56	0,06	0,30	0,64	0,64	0,78	16,58	
		$\pm 5,95$	$\pm 9,92$	$\pm 6,22$										
		14,15*	56,11*	3,98 ^{ns}	49,42	123,66	0,45	0,03	0,40	0,73	0,57	0,84	18,44	
10	5 ^a 6 ^a 7 ^a 8 ^a 9 ^a	$\pm 6,93$	$\pm 9,45$	$\pm 4,85$										
		11,74*	66,95*	13,08*	66,18	157,93	0,42	0,08	0,42	0,70	0,50	0,82	21,10	
		$\pm 4,69$	$\pm 11,32$	$\pm 9,57$										
		18,68*	118,31*	16,52*	66,30	219,81	0,54	0,08	0,30	0,78	0,62	0,87	27,01	
		$\pm 5,18$	$\pm 16,10$	$\pm 13,32$										
		23,94*	97,89*	43,27*	102,64	267,74	0,37	0,16	0,38	0,78	0,45	0,87	30,97	
		$\pm 1,79$	$\pm 17,52$	$\pm 21,45$										
		63,10*	208,24*	0,24 ^{ns}	129,35	400,92	0,52	0,00	0,32	0,80	0,68	0,88	29,56	
		$\pm 17,87$	$\pm 31,86$	$\pm 2,32$										
30,84*	330,83*	69,7*	33,51	464,95	0,71	0,15	0,07	0,55	0,78	0,73	46,61			
$\pm 29,91$	$\pm 41,60$	$\pm 38,99$												
70,06*	131,51*	16,34 ^{ns}	137,56	355,47	0,37	0,05	0,39	0,82	0,56	0,89	30,91			
$\pm 24,12$	$\pm 29,52$	$\pm 16,65$												
6,40*	167,03*	61,58*	50,87	285,88	0,58	0,22	0,18	0,78	0,60	0,87	34,07			
$\pm 1,72$	$\pm 25,39$	$\pm 26,67$												

σ_g^2 : variância genotípica; σ_i^2 : variância dos efeitos permanentes de indivíduo; σ_p^2 : variância dos efeitos permanente de parcela; σ_e^2 : variância residual; σ_f^2 : variância fenotípica; c_i^2 , c_p^2 e c_e^2 : coeficientes de determinação dos efeitos permanentes de indivíduo, de parcela e residuais, respectivamente; H_g^2 : herdabilidade média no sentido amplo; ρ : coeficiente de repetibilidade; r: acurácia seletiva e μ : média fenotípica. * e ^{ns}: significativo e não significativo à $P < 0,05$, respectivamente.

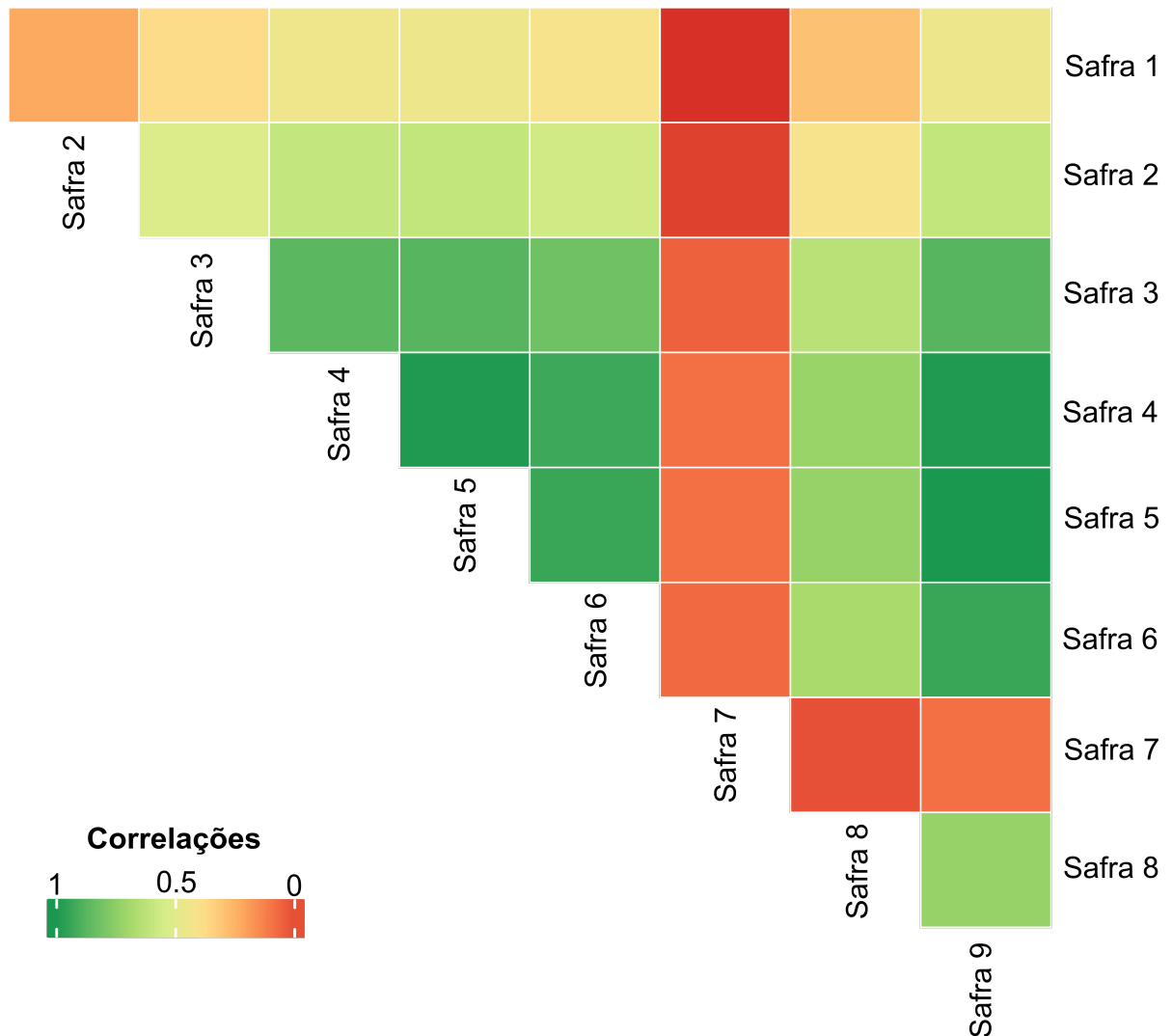
Figura 1. Oscilações dos valores genotípicos dos híbridos triplos de cupuaçuzeiro para produtividade média de frutos, ao longo das safras.



Fonte: O autor (2022).

Com o emprego da estrutura fator analítico para modelar os efeitos genotípicos (Modelo 10), é possível estimar as correlações genéticas entre pares de safras, possibilitando inferências acerca da interação genótipos x medições (GM) (Figura 2). Face à ausência de interação, espera-se que a correlação genética entre medições seja muito alta (aproximadamente 1). Isto não ocorreu no presente experimento, haja vista a presença de fracas associações, atestando que há uma performance diferencial das progênies de acordo com as condições ambientais variáveis ano-a-ano.

Figura 2. Heatmap representativo das correlações genéticas entre safras.

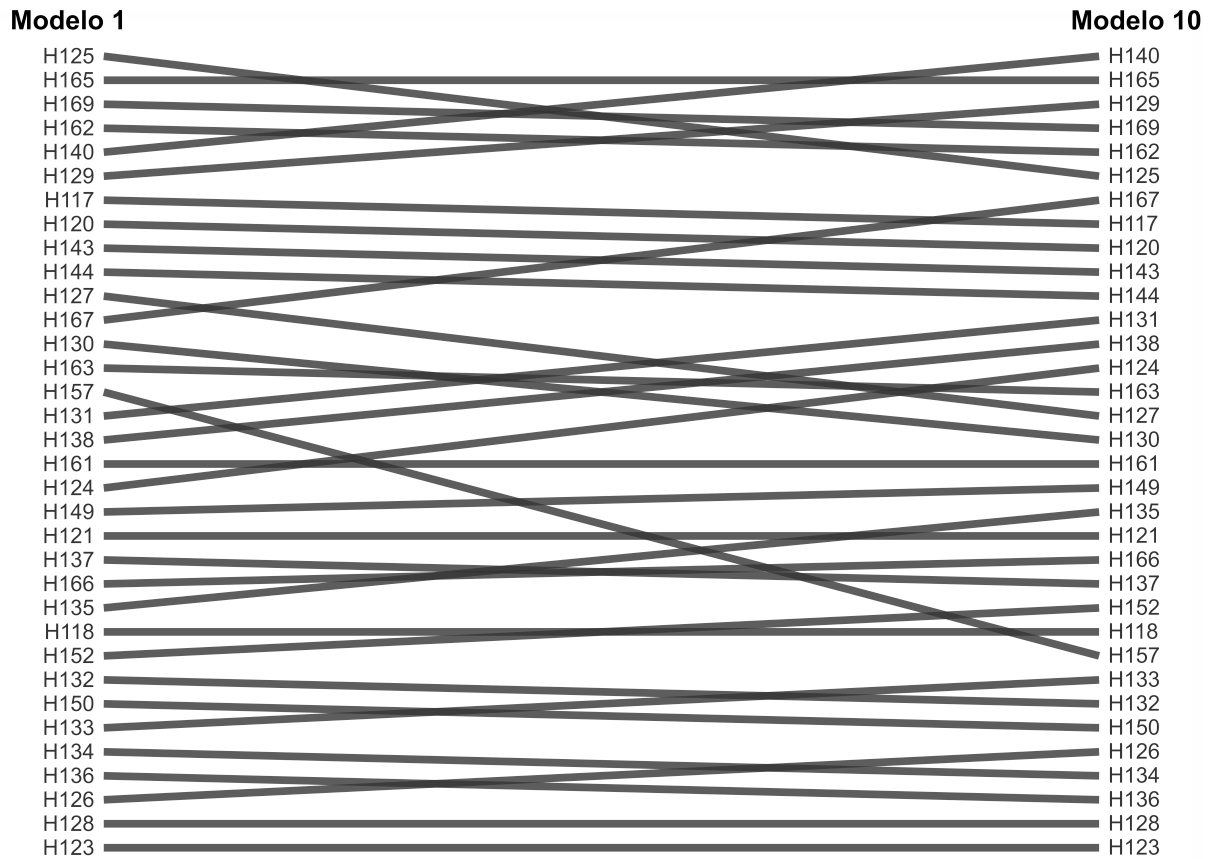


Fonte: O autor (2022).

A predição dos valores genotípicos empíricos (EBLUPs) também é influenciada pelo emprego de diferentes estruturas de covariância dos efeitos aleatórios do modelo. Esta influência ocorre em maior ou menor escala, alterando o ordenamento dos híbridos (Figura 3). Analisando paralelamente os Modelos 1 e 10, observa-se que houve coincidência de 90%. Os híbridos H117, H120, H125, H129, H140, H144, H162, H165 e H169 estão presentes entre os 10 melhores considerando os EBLUPs de ambos modelos. A diferença fica por conta da presença do híbrido H167 no ordenamento derivado do Modelo 10, em substituição ao H143, no ordenamento derivado do Modelo 1. Esta pequena diferença é suficiente para influenciar nos ganhos com a seleção, 13,8% considerando o Modelo 1 e 15,5% considerando Modelo 10.

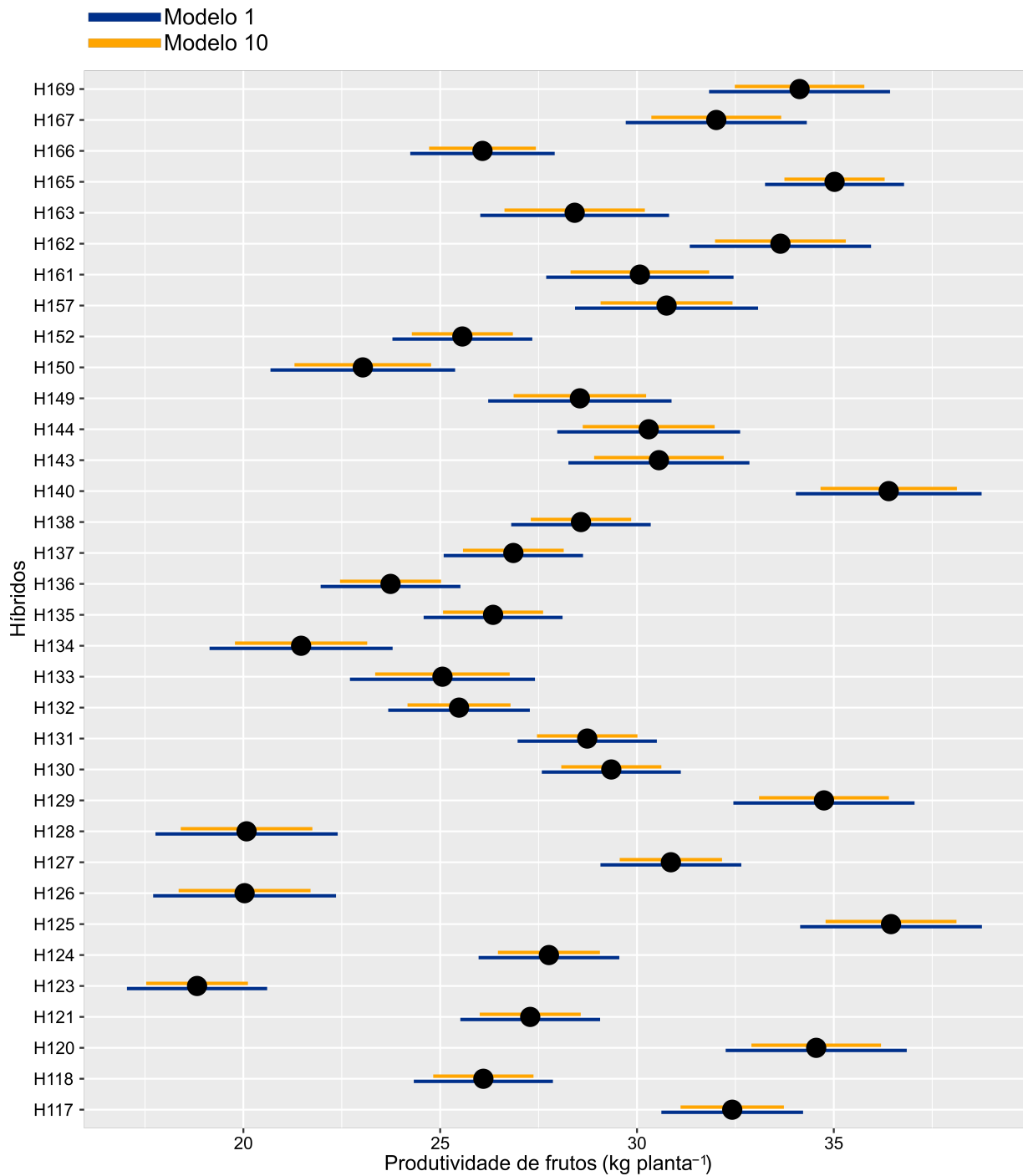
O modelo 10 permitiu, ainda, uma estimação mais acurada dos valores genotípicos, imputando maior precisão à seleção derivada dos resultados deste modelo (Figura 4)

Figura 3. Ordenamentos dos híbridos de cupuaçuzeiro com base nos valores genotípicos empíricos obtidos pelos modelos 1 (mais simples e parcimonioso) e 10 (melhor ajuste).



Fonte: O autor (2022)

Figura 4. Médias fenotípicas e erros-padrão da predição dos valores genotípicos empíricos dos híbridos de cupuaçuzeiro obtidos pelos modelos 1 (mais simples e parcimonioso, em azul) e 10 (de melhor ajuste, em amarelo)



Fonte: O autor (2022).

4 DISCUSSÃO

Com frequência, experimentos envolvendo culturas perenes, tal qual o cupuaçuzeiro, envolvem áreas experimentais extensas e repetidas medições ao longo do tempo (DIAS, 2001).

Os efeitos causados pelas variações ano-a-ano afetam os genótipos de forma diferencial, a depender de sua constituição genética, seu estado de maturação e o micro-ambiente onde a planta está instalada no campo experimental, evidenciando a inter-relação dos fatores (ISIK; HOLLAND; MALTECCA, 2017; LAHIVE; HADLEY; DAYMOND, 2019). Portanto, a heterogeneidade de covariâncias é, na maioria das vezes, regra para estas culturas, e deve ser devidamente considerada por ocasião da escolha do modelo estatístico a ser empregado na avaliação genética, priorizando aqueles de melhor ajuste, capazes de captar as peculiaridades supracitadas (ANDRADE et al., 2016).

Na escolha de modelos, a utilização de um critério de informação é fundamental. A eficácia do AIC neste sentido já foi comprovada em vários estudos (CAVANAUGH; NEATH, 2019; PORTET, 2020; RESENDE; ALVES, 2020). Na presente pesquisa, dois pontos devem ser ressaltados. Há uma oscilação na acurácia seletiva à medida que novas estruturas são modeladas. Para os efeitos residuais, de ambiente permanente de indivíduo e de parcela, as estruturas de melhor adequação não são aquelas que proporcionaram maior acurácia seletiva, o que também foi observado por Andrade et al. (2016). A partir deste comportamento, infere-se que as acurácias maiores nos demais modelos de maior AIC são superestimadas. Porém, uma vez que as matrizes de todos os efeitos foram modeladas e o modelo de melhor ajuste foi identificado, atingiu-se a maior acurácia dentre todos os modelos testados. Portanto, testando diferentes estruturas de covariância dos efeitos aleatórios, foi possível identificar o melhor modelo, pelo seu menor valor de AIC - um critério estatístico - e que também é o mais confiável; e pela maximização da acurácia dentre os modelos avaliados, um critério genético-estatístico (RESENDE; ALVES, 2020).

Para a matriz dos efeitos residuais (R), a estrutura AR1H foi a mais adequada. Isto implica que, no presente estudo, as variâncias residuais são heterogêneas, fato esperado devido as particularidades inerentes à cada safra, e à presença de correlação serial como componente das covariâncias residuais entre safras. A estrutura AR1H é comumente empregada em análises espaciais, considerando que a medida que a distância entre dois indivíduos ou parcelas aumenta, a correlação entre essas diminui (ANDRADE et al., 2020; BERNARDELI et al., 2021). Este conceito pode ser estendido à análise de medidas repetidas, isto é, a correlação entre duas medições diminui à medida que estas possuem maior intervalo (de tempo, no caso presente) entre si (LIU; ROVINE; MOLENAAR, 2012; VERBYLA et al., 2021). Isto ocorre não só pelas peculiaridades de cada safra, mas também pelo estágio de maturação da planta.

As estruturas avaliadas para os efeitos de ambiente permanente de indivíduo e de parcela (IDH para o primeiro e CORH para o último) também consideram variâncias

heterogêneas. No caso da CORH, também há a presença de um coeficiente de correlação considerando as covariâncias. Porém, diferentemente da estrutura AR1H, não há variação entre os coeficientes de correlação à medida que a distância entre medições aumenta. Portanto, infere-se que há heterogeneidade entre as variâncias atribuídas ao referido efeito e as inter-relações são uniformes, independentemente das distâncias entre medições (SHALIZI; ISIK, 2019).

A eficácia da estrutura FA1 para os efeitos genotípicos, já foi comprovada para ensaios em múltiplos ambientes (SHALIZI; ISIK, 2019; SMITH; CULLIS, 2018) e vem sendo adotada para análises de medidas repetidas (FAVERI et al., 2015; VERBYLA et al., 2021). O modelo 10 aproxima-se do modelo não estruturado (o mais completo e parametrizado) ao considerar cada medida como uma variável distinta, permitindo predições particularizadas para cada medição (PIEPHO, 1997). Porém, o modelo 10 é mais parcimonioso, pois considera as correlações entre as variáveis (medidas) para diminuir a dimensionalidade à um fator (SMITH; CULLIS; THOMPSON, 2001; VERBYLA et al., 2021). Não existem registros de modelagem da estrutura FA1 em uma fruteira perene, tal qual o cupuaçuzeiro, sendo o presente estudo pioneiro neste sentido.

A aptidão dos híbridos para seleção é corroborada pela significância da variância genotípica, tanto no Modelo 1, mais simples, quanto no 10, mais adequado. No entanto, somente por meio da estruturação da matriz dos efeitos genotípicos, no modelo de melhor ajuste, foi possível observar uma variação na grandeza desta variância ao longo das safras. Os diferentes estágios de maturação, associados com a expressão diferencial de genes, a qual também está ligada à fatores ambientais, pode ter relação com este comportamento (MELO et al., 2020; TAHI et al., 2019). Ainda analisando ambos modelos em paralelo, é evidente que a detecção das oscilações dos demais efeitos só pôde ser realizada pelo Modelo 10. A estruturação das matrizes dos referidos efeitos viabilizou, por exemplo, a visualização da importância dos efeitos de ambiente permanente para a variância fenotípica, chegando a ser maior que a variância residual em algumas safras, de forma oposta ao Modelo 1.

Ademais, a influência dos efeitos genotípicos na variância fenotípica, mensurada pela herdabilidade média, pode ser considerada alta ($> 0,50$, RESENDE; ALVES, 2020), indício de que a seleção de famílias será bem-sucedida. Tal comportamento incide sobre a acurácia de seleção, igualmente elevada (RESENDE; ALVES, 2020). Outro parâmetro de alta magnitude foi o coeficiente de repetibilidade. Como relatado no parágrafo anterior, no modelo 1, a maior parte dos efeitos não genéticos da variância fenotípica foi alocada na variância residual, resultando em uma baixa repetibilidade. Ao realizar a modelagem das estruturas de covariância

dos efeitos aleatórios do modelo estatístico, foi possível estimar a magnitude de cada efeito de forma mais acurada, diminuindo a variância residual e resultando em uma alta repetibilidade. A baixa variação na acurácia a partir da quarta safra e os altos coeficiente de repetibilidade observados sugerem a possibilidade de realizar a seleção genética mais precocemente (CHAVES et al., 2021).

Particularizando ao cupuaçuzeiro, a cultura possui alta divergência intrapopulacional, promovida pela autoincompatibilidade (RAMOS et al., 2005); além de um longo período juvenil, aproximadamente três anos, comum em fruteiras perenes (RAI; SHEKHAWAT, 2014). Após este período, a planta inicia o período produtivo pré-clímax, que se estende até, aproximadamente, o oitavo ano após a implantação (ALVES et al., 2020). Neste período, a planta adapta-se às condições edafoclimáticas do pomar, e avança até seu processo de maturação produtiva. Isto significa que há uma manifestação diferencial de genes ao longo da vida do indivíduo, constituindo uma fonte de heterogeneidade genética (FRIEDMAN, 2020). Ademais, no presente estudo, foi detectado indícios de bienalidade, outro fator que favorece a variação entre medidas, haja vista o comportamento da média fenotípica, dos coeficientes de herdabilidade média no sentido amplo e dos valores genotípicos a partir da quinta safra.

As baixas correlações genéticas entre safras indicam a existência de interação GM. Em outras palavras, há desempenho diferencial dos híbridos ao longo das safras. Esta informação é importante para priorizar a seleção de genótipos resilientes, capazes de suportar as variações macro-ambientais, isto é, aquelas de cunho climático, cada vez mais frequentes; e micro-ambientais, relacionadas ao manejo e a competição intra- e interespecífica (FARRELL et al., 2018; LAHIVE; HADLEY; DAYMOND, 2019). Isto é especialmente relevante para o cupuaçuzeiro, que, a exemplo do cacauzeiro (*Theobroma cacao* L.), é capaz de continuar produzindo por muitas décadas (DIAS, 2001). Ademais, as baixas correlações genéticas observadas evidenciam a inadequação do modelo simples de repetibilidade para a análise de medidas repetidas, no caso do presente estudo, haja vista sua pressuposição de que as correlações entre pares de safras são próximas à unidade (RESENDE; SILVA; AZEVEDO, 2014).

Houve uma variação discreta entre os valores genotípicos empíricos preditos (EBLUPs) por cada modelo. Apesar de discreta, a modelagem das estruturas de covariância possibilitou um aumento no progresso genético de 1,7%, bem como um menor erro de predição, conferindo confiabilidade à seleção. Desta forma, ressalta-se a importância da modelagem das estruturas de covariância dos efeitos aleatórios do modelo, a qual possibilita inferências verossímeis e

seleções mais acuradas e, por conseguinte, a otimização dos programas de melhoramento (MELO et al., 2020).

5 CONSIDERAÇÕES FINAIS

Modelar as estruturas de covariância e identificar o modelo de melhor ajuste mostrou-se fundamental para a avaliação e seleção genética do cupuaçuzeiro. O modelo de melhor ajuste gerou resultados mais confiáveis, tanto na estimação dos componentes de variância quanto na predição dos valores genotípicos, possibilitando o aprimoramento de uma das etapas mais importantes do programa de melhoramento: a avaliação genética. No modelo 10, as matrizes dos efeitos residuais, de ambiente permanente, de parcela e genotípicos foram modeladas pelas estruturas auto-regressiva heterogênea de primeira ordem (AR1H), bloco diagonal (IDH), simetria composta heterogênea (CORH) e fator analítico (FA1), respectivamente. Por meio deste modelo, os híbridos G117, G120, G125, G129, G140, G144, G162, G165, G167 e G169 foram selecionados, proporcionando um ganho genético de 15,5%.

REFERÊNCIAS

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v. 19, n. 6, p. 716–723, 1974.
- ALVES, R. M.; SEBBENN, A. M.; ARTERO, A. S.; CLEMENT, C.; FIGUEIRA, A. High levels of genetic divergence and inbreeding in populations of cupuassu (*Theobroma grandiflorum*). **Tree Genetics & Genomes**, v. 3, n. 4, p. 289–298, 2007.
- ALVES, R. M.; CHAVES, S. F. S.; GAMA, M. A. P.; PEDROZA NETO, J. L.; SANTOS, T. G. Simultaneous selection of cupuassu tree and Brazilian mahogany genotypes in an agroforestry system in Pará state, Brazil. **Acta Amazonica**, v. 50, n. 3, p. 183–191, 2020.
- ALVES, R. M.; CHAVES, S. F. S. BRS Careca, BRS Fartura, BRS Duquesa, BRS Curinga, and BRS Golias: new cupuassu tree cultivars. **Crop Breeding and Applied Biotechnology**, v. 20, n. 4, p. e342920413, 2020.
- ALVES, R. S.; TEODORO, P. E.; PEIXOTO, L. A.; ROCHA, J. R. A. S. C.; SILVA, L. A.; LAVIOLA, B. G.; RESENDE, M. D. V.; BHERING, L. L. Multiple-trait BLUP in longitudinal data analysis on *Jatropha curcas* breeding for bioenergy. **Industrial Crops and Products**, v. 130, p. 558–561, 2019.
- AMAZONAS. Instituto de Desenvolvimento Agropecuário e Florestal Sustentável do Estado do Amazonas. Relatório de Atividades IDAM 2019. Disponível em: <<http://www.idam.am.gov.br/wp-content/uploads/2020/07/RAIDAM2019.pdf>>. Acesso em: 15 set. 2021.

- ANDRADE, M. H. M. L.; FERNANDES FILHO, C. C.; FERNANDES, M. O.; BASTOS, A. J. R.; GUEDES, M. L.; MARÇAL, T. S.; GONÇALVES, F. M. A.; PINTO, C. A. B. P.; ZOTARELLI, L. Accounting for spatial trends to increase the selection efficiency in potato breeding. **Crop Science**, v. 60, n. 5, p. 2354–2372, 2020.
- ANDRADE, V. T.; GONÇALVES, F. M. A.; NUNES, J. A. R.; BOTELHO, C. E. Statistical modeling implications for coffee progenies selection. **Euphytica**, v. 207, n. 1, p. 177–189, 2016.
- BERNARDELI, A.; ROCHA, J. R. A. S. C.; BORÉM, A.; LORENZONI, R.; AGUIAR, R.; SILVA, J. N. B.; BUENO, R. D.; ALVES, R. S.; JARQUIN, D.; RIBEIRO, C.; COSTA, M. D.-B. L. Modeling spatial trends and enhancing genetic selection: An approach to soybean seed composition breeding. **Crop Science**, v. 61, n. 2, p. 976–988, 2021.
- BUTLER, D. G.; CULLIS, B. R.; GILMOUR, A. R.; GOGEL, B. J.; THOMPSON, R. **ASReml-R reference manual Version 4**. Hemel Hempstead, UK: VSN International, 2018.
- CAVANAUGH, J. E.; NEATH, A. A. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. **WIREs Computational Statistics**, v. 11, n. 3, p. e1460, 2019.
- CHAVES, S. F. S.; ALVES, R. M.; ALVES, R. S.; SEBBENN, A. M.; RESENDE, M. D. V.; DIAS, L. A. S. *Theobroma grandiflorum* breeding optimization based on repeatability, stability and adaptability information. **Euphytica**, v. 217, n. 12, p. 211, 2021.
- CULLIS, B. R.; SMITH, A. B.; BEECK, C. P.; COWLING, W. A. Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. **Genome**, v. 53, n. 11, p. 1002–1016, 2010.
- CULLIS, B. R.; SMITH, A. B.; COOMBES, N. E. On the design of early generation variety trials with correlated data. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 11, n. 4, p. 381–393, 2006.
- DIAS, L. A. S. **Melhoramento genético do cacauzeiro**. Goiás: FUNAPE-UFG, 2001. 578p.
- DIGGLE, P. J. An approach to the analysis of repeated measurements. **Biometrics**, v. 44, n. 4, p. 959–971, 1988.
- EDWARDS, J. W.; JANNINK, J. Bayesian modeling of heterogeneous error and genotype × environment interaction variances. **Crop Science**, v. 46, n. 2, p. 820–833, 2006.
- EMBRAPA. **Mapas de solos e de aptidão agrícola de áreas alteradas do Pará**. 1. ed. Belém, Pará: Embrapa Amazônia Oriental, 2016. 19p.
- FARRELL, A. D.; RHINEY, K.; EITZINGER, A.; UMAHARAN, P. Climate adaptation in a minor crop species: is the cocoa breeding network prepared for climate change? **Agroecology and Sustainable Food Systems**, v. 42, n. 7, p. 812–833, 2018.
- FAVERI, J.; VERBYLA, A. P.; PITCHFORD, W. S.; VENKATANAGAPPA, S.; CULLIS, B. R. Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. **Crop and Pasture Science**, v. 66, n. 9, p. 947, 2015.

FRIEDMAN, J. The evolution of annual and perennial plant life histories: Ecological correlates and genetic mechanisms. **Annual Review of Ecology, Evolution, and Systematics**, v. 51, n. 1, p. 461–481, 2020.

GILMOUR, A. R.; GOGEL, B. J.; CULLIS, B. R.; WELHAM, S. J.; THOMPSON, R. **ASReml User Guide Release 4.1 Structural Specification**. Hemel Hempstead, UK: VSN International, 2015.

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

IBGE. **Instituto Brasileiro de Geografia e Estatística. Censo Agropecuário**.

Disponível em: <<https://censoagro2017.ibge.gov.br/resultados-censo-agro-2017.html>>. Acesso em: 15 set. 2021.

ISIK, F.; HOLLAND, J.; MALTECCA, C. **Genetic Data Analysis for Plant and Animal Breeding**. Cham: Springer International Publishing, 2017.

LAHIVE, F.; HADLEY, P.; DAYMOND, A. J. The physiological responses of cacao to the environment and the implications for climate change resilience. A review. **Agronomy for Sustainable Development**, v. 39, n. 1, p. 5, 2019.

LEMOS, O. F.; TREMACOLDI, C. R.; POLTRONIERI, M. C. **Boas práticas agrícolas para aumento da produtividade e qualidade da pimenta-do-reino no Estado do Pará**. Brasília, DF: Embrapa, 2014. 54p.

LI, Y.; SUONTAMA, M.; BURDON, R. D.; DUNGEY, H. S. Genotype by environment interactions in forest tree breeding: review of methodology and perspectives on research and application. **Tree Genetics & Genomes**, v. 13, n. 3, p. 60, 2017.

LIU, S.; ROVINE, M. J.; MOLENAAR, P. C. M. Selecting a linear mixed model for longitudinal data: Repeated measures analysis of variance, covariance pattern model, and growth curve approaches. **Psychological Methods**, v. 17, n. 1, p. 15–30, 2012.

MELO, V. L.; MARÇAL, T. S.; ROCHA, J. R. A. S. C.; ANJOS, R. S. R.; CARNEIRO, P. C. S.; CARNEIRO, J. E. S. Modeling (co)variance structures for genetic and non-genetic effects in the selection of common bean progenies. **Euphytica**, v. 216, n. 5, p. 77, 2020.

MORAES, J. R. S. C.; ROLIM, G. S.; MARTORANO, L. G.; APARECIDO, L. E. O.; OLIVEIRA, M. S. P.; FARIAS NETO, J. T. Agrometeorological models to forecast açai (*Euterpe oleracea* Mart.) yield in the Eastern Amazon. **Journal of the Science of Food and Agriculture**, v. 100, n. 4, p. 1558–1569, 2020.

MRODE, R. A. **Linear models for the prediction of animal breeding values**. 3rd ed ed. Boston, MA: CABI, 2014.

PARÁ. **Secretaria de Estado de Desenvolvimento Agropecuário e da Pesca. Dados Agropecuários**. Disponível em:

<<http://www.sedap.pa.gov.br/dadosagropecuarios/agropecuaria>>. Acesso em: 15 set. 2021.

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

- PEIXOTO, M. A.; EVANGELISTA, J. S. P. C.; COELHO, I. F.; ALVES, R. S.; LAVIOLA, B. G.; SILVA, F. F.; RESENDE, M. D. V.; BHERING, L. L. Multiple-trait model through Bayesian inference applied to *Jatropha curcas* breeding for bioenergy. **PLoS ONE**, v. 16, n. 3, p. e0247775, 2021.
- PEREIRA, A. L. F.; ABREU, V. K. G.; RODRIGUES, S. Cupuassu—*Theobroma grandiflorum*. In: RODRIGUES, S.; SILVA, E. O.; BRITO, E. S. (Eds.). **Exotic Fruits**. [s.l.] Academic Press, 2018. p. 159–162.
- PIEPHO, H. P. Analyzing genotype-environment data by mixed models with multiplicative terms. **Biometrics**, v. 53, n. 2, p. 761-766, 1997.
- PIEPHO, H. P.; MÖHRING, J.; MELCHINGER, A. E.; BÜCHSE, A. BLUP for phenotypic selection in plant breeding and variety testing. **Euphytica**, v. 161, n. 1–2, p. 209–228, 2008.
- PIEPHO, H. P.; ECKL, T. Analysis of series of variety trials with perennial crops. **Grass and Forage Science**, v. 69, n. 3, p. 431–440, 2014.
- PORTET, S. A primer on model selection using the Akaike Information Criterion. **Infectious Disease Modelling**, v. 5, p. 111–128, 2020.
- R CORE TEAM. **R: A Language and environment for statistical computing**. Viena, Áustria: R Foundation for Statistical Computing, 2022.
- RAI, M. K.; SHEKHAWAT, N. S. Recent advances in genetic engineering for improvement of fruit crops. **Plant Cell, Tissue and Organ Culture (PCTOC)**, v. 116, n. 1, p. 1–15, 2014.
- RAMOS, A. R.; VENTURIERI, G. A.; CUCO, S. M.; CASTRO, N. M. The site of selfincompatibility action in cupuassu (*Theobroma grandiflorum*). **Brazilian Journal of Botany**, v. 28, p. 569–578, 2005.
- RESENDE, M. D. V.; ALVES, R. S. Linear, generalized, hierarchical, bayesian and random regression mixed models in genetic/genomics in plant breeding. **Functional Plant Breeding Journal**, v. 2, n. 2, p. 1–31, 2020.
- RESENDE, M. D. V.; SILVA, F. F.; AZEVEDO, C. F. **Estatística matemática, biométrica e computacional: modelos mistos, multivariados, categóricos e generalizados (REML/BLUP), inferência bayesiana, regressão, aleatória, seleção genômica, QTL, GWAS, estatística espacial e temporal, competição, sobrevivência**. Viçosa, MG, Brazil: UFV, 2014. 881p.
- SHALIZI, M. N.; ISIK, F. Genetic parameter estimates and GxE interaction in a large cloned population of *Pinus taeda* L. **Tree Genetics & Genomes**, v. 15, n. 3, p. 46, 2019.
- SMITH, A. B.; CULLIS, B.; THOMPSON, R. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. **Biometrics**, v. 57, n. 4, p. 1138-1147, 2001

SMITH, A. B.; GANESALINGAM, A.; KUCHEL, H.; CULLIS, B. R. Factor analytic mixed models for the provision of grower information from national crop variety testing programs. **Theoretical and Applied Genetics**, v. 128, n. 1, p. 55–72, 2015.

SMITH, A. B.; CULLIS, B. R. Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. **Euphytica**, v. 214, n. 8, p. 143, 2018.

SMITH, K. F.; SPANGENBERG, G. Forage breeding for changing environments and production systems: an overview. **Crop and Pasture Science**, v. 65, n. 11, p. i - ii, 2014.

SOUZA, A. G. C.; RESENDE, M. D. V.; SILVA, S. E. L.; SOUSA, N. R. The cupuaçu genetic improvement program at Embrapa Amazônia Ocidental. **Cropp Breeding and Applied Biotechnology**, v. 2, n. 3, p. 471–478, 2002.

SOUZA, A. G. C. **Boas práticas agrícolas da cultura do cupuaçuzeiro**. Manaus: Embrapa Amazônia Ocidental, 2007.

STRINGER, J. K.; ATKIN, F. C.; GEZAN, S. A. Statistical Approaches in Plant Breeding: Maximising the Use of the Genetic Information. In: CAMPOS, H.; CALIGARI, P. D. S. (Eds.). **Genetic Improvement of Tropical Crops**. Cham: Springer International Publishing, 2017. p. 3–17.

TAHI, M.; TREBISSOU, C.; RIBEYRE, F.; GUIRAUD, B. S.; POKOU, D.; CILAS, C. Variation in yield over time in a cacao factorial mating design: changes in heritability and longitudinal data analyses over 13 consecutive years. **Euphytica**, v. 215, n. 6, p. 106, 2019.

VERBYLA, A. P.; FAVERI, J.; DEERY, D. M.; REBETZKE, G. J. Modelling temporal genetic and spatio-temporal residual effects for high-throughput phenotyping data. **Australian & New Zealand Journal of Statistics**, v. 63, n. 2, p. 284–308, 2021.

WARD, B. P.; BROWN-GUEDIRA, G.; TYAGI, P.; KOLB, F. L.; VAN SAFORD, D. A.; SNELLER, C. H.; GRIFFEY, C. A. Multienvironment and multitrait genomic selection models in unbalanced early-generation wheat yield trials. **Crop Science**, v. 59, p. 491–507, 2019.